David McKenzie
Susana Puerto
Frank Odhiambo

Unpacking the determinants of entrepreneurship development and economic empowerment for women in Kenya

April 2019

Impact Evaluation Report 95

Private Sector Development

3ie International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) promotes evidence-informed equitable, inclusive and sustainable development. We support the generation and effective use of high-quality evidence to inform decision-making and improve the lives of people living in poverty in low- and middle-income countries. We provide guidance and support to produce, synthesise and quality-assure evidence of what works, for whom, how, why and at what cost.

3ie impact evaluations

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design, and uses the most appropriate mix of methods to capture outcomes and are useful in complex development contexts.

About this report

3ie accepted the final version of the report, *Unpacking the determinants of entrepreneurship development and economic empowerment for women in Kenya*, as partial fulfilment of requirements under grant OW4.1213 awarded through Open Window 4. The content has been copy-edited and formatted for publication by 3ie.

The 3ie technical quality assurance team comprises Tara Kaul, Diana Lopez-Avila, Emmanuel Jimenez, an anonymous external impact evaluation design expert reviewer and an anonymous external sector expert reviewer, with overall technical supervision by Marie Gaarder. The 3ie editorial production team for this report comprises Sahib Singh and Akarsh Gupta, with Beryl Leach providing overall editorial supervision.

All of the content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are also the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Please direct any comments or queries to the corresponding author, David McKenzie, at dmckenzie@worldbank.org.

Funding for this impact evaluation was provided by 3ie's donors, which include UK aid, the Bill & Melinda Gates Foundation and the Hewlett Foundation. A complete listing of all of 3ie's donors is available on the 3ie website.

Suggested citation: McKenzie, D, Puerto, S and Odhiambo, F, 2019. *Unpacking the determinants of entrepreneurship development and economic empowerment for women in Kenya*, 3ie Impact Evaluation Report 95. New Delhi: International Initiative for Impact Evaluation (3ie). Available at: https://doi.org/10.23846/OW4IE95

Cover photo: hodag / Flickr

Unpacking the determinants of entrepreneurship development and economic empowerment for women in Kenya

David McKenzie
World Bank

Susana Puerto
International Labour Organization

Frank Odhiambo
Innovations for Poverty Action

Impact Evaluation Report 95

April 2019

Acknowledgements

This report was written by David McKenzie from the World Bank, Susana Puerto from the International Labour Organization, and Frank Odhiambo from Innovations for Poverty Action. The principal investigators thank Valerie Breda, Faizan Diwan, Bilal Hakeem, Mary Kamore, Virginia Losada, Jane Maigua, Grace Makana, Silvia Paruzzolo, and Miriam Wekesa for their top contributions to the study design and implementation. The International Center for Research on Women (ICRW) provided valuable input to the study planning and the qualitative work. The authors wish to acknowledge the helpful comments and suggestions from participants at various seminars and conferences and gratefully acknowledge funding provided from the International Initiative for Impact Evaluation (3ie), Private Enterprise Development in Low-Income Countries, the International Labour Organization (ILO), the Strategic Research Programme, and the World Bank Jobs Umbrella multi-donor trust fund.

Human Subjects Approval was obtained from Innovations for Poverty Action (IPA, 13February-002) and the Maseno University Ethics Review Committee (MSU/DRPC/MUERC/000006/13). Authority to conduct research in Kenya was provided by the Kenyan Ministry of Science and Technology (NCST/RCD/14/013/553B). A pre-analysis plan and the associated trial were registered on the American Economic Association's Social Science Registry on February 21, 2014 (AEARCTR-0000287).

Summary

The International Labour Organization's (ILO)'s Gender and Entrepreneurship Together training programme (GET Ahead) seeks to enhance women's opportunities in entrepreneurship through knowledge and skills development in business and management. The programme brings a unique gender perspective and interactive approach in order to expose women, primarily from low-income settings, to business and working environments, development of business ideas, product design and management practices. The programme began in Thailand in 2001 and has now been carried out in at least 21 countries, serving approximately 400,000 women. This study assesses an intervention implemented in Kenya with the core objective of facilitating business growth. It offered women with small businesses in rural markets a five-day training course. Its delivery followed ILO's implementation manuals and relied on local training providers. A year after training, some women who had received GET Ahead were offered complementary group and individual support services over a five-month period.

This impact evaluation focuses on understanding the impact of GET Ahead – and subsequent mentoring – on profitability, sales and business survivorship for treated firms. The second research question explores changes in secondary outcomes among treated firms. Next, the study focuses on the intervention's impact on primary outcomes of non-treated firms; lastly, it seeks to understand the causal chain of process and mechanisms. It departs from the hypothesis that training and mentoring should lead to an increase in business knowledge, greater deployment of business practices and a potential increase in profitability or sales.

We worked with a sample of 3,537 women in 157 markets in four counties of Kenya: Kakamega, Kisii, Embu and Kitui. These regions are largely rural, with most of the population below the poverty line. These markets are typically small and remote, largely consisting of women operating a limited variety of businesses.

We used a two-stage randomised experiment to allow for estimation of the causal impact of training participation. First, we allocated markets to treatment or control. Then within the treatment market we randomly selected 1,172 women to invite to training. A comparison of women invited to training in the treatment markets and women in the control markets enables us to estimate the impact of training. Comparing the women in treatment markets who were not invited to training with women in the control markets enables a measurement of spillover impacts of operating in a market where others are trained. We conducted four rounds of follow-up surveys – two after one year, and two after three years – to measure impacts.

After three years, treated women are earning higher profits than the pure control group, with no spillover impact on other women in their markets. They also are more likely to have surviving businesses and higher weekly sales. This increase in business income is accompanied by improvements in mental health and subjective wellbeing. Examining impacts at market level shows that the treated markets have more customers each week, as well as higher sales. There is no change in the rate of new business entry into these markets. Treated firms show an increase in the proportion of good business practices used. A key mechanism for this market growth appears to be the training they received: business owners began keeping more reliable opening hours and diversifying

their range of products. Both factors make the market more attractive for customers and allow overall sales to grow, rather than reallocating sales from one business to another.

Treatment impacts are stronger after three years than after one year. Few firms receive finance, and training does not increase the use of credit. As a result, it seems that firm owners have had to slowly build their inventories by reinvesting profits over time. Mentoring intervention was costlier per woman and has not had a significantly different impact than training alone.

Many evaluations of business training programmes have struggled to find significant impacts. A key reason for this has been the use of relatively small samples with heterogeneous firms. This lack of statistical significance has been interpreted by some as evidence that training seldom works, rather than the correct interpretation that there is a lack of evidence as to whether it works or not. The treatment impacts here are not larger than those found in prior studies, but they are much more precise and show that training can have a positive impact. The impact of US$2.60 per week is not transformative, but it represents an important increase in income for poor women. Moreover, and importantly, this benefit to trained women does not come at the expense of other women operating in their same markets. It appears that training passes a cost-benefit test, since gains would need to last 1.5 years to offset the costs.

Contents

List of figures and tables

Abbreviations and acronyms

3ie	International Initiative for Impact Evaluation
GET Ahead	Gender and Enterprise Together
ICRW	International Center for Research on Women
ILO	International Labour Organization
IPA	Innovations for Poverty Action
ITT	Intent to treat
KIE	Kenya Industrial Estates
KSH	Kenyan shillings
LATE	Local Average Treatment Effect
PAP	Pre-Analysis Plan
WEF	Women Enterprise Fund

1. Introduction

Governments, non-governmental organisations, multilateral organisations and private sector funders around the world promote and offer a variety of direct support to small businesses to help them grow. These investments recognise the important role small enterprises play in providing incomes and livelihoods, especially where there is a lack of sufficient wage employment. Business training is one of the most common services through which to foster the growth of micro, small- and medium-sized enterprises. It seeks to enhance knowledge and management skills, strengthen creditworthiness and business networks, and ultimately boost performance, growth and sustainability of existing small enterprises.

In 2013, the World Bank and the International Labour Organization (ILO) set out to design an impact evaluation that would focus on understanding the impact of business training on women entrepreneurs, their businesses, and female-owned businesses around them. This strong attention to women responds to the common interest in improving knowledge about what works, how and why with regard to improving women's economic empowerment and enterprise development (ILO 2015). Female entrepreneurship is skewed towards smaller and informal enterprises and is often concentrated in less productive firms and less profitable sectors (ILO 2015; World Bank 2012). Microenterprises run by women account for a large share of marketplace activities in many African countries. However, the majority of these businesses are very small in scale, yielding subsistence amounts of earnings for their owners and rarely growing to the point where they create above-subsistence incomes for their owners or jobs for others. Accordingly, efforts to improve their business outcomes have the potential to trigger individual and social gains conducive to gender equality, decent work, productive, good-quality employment and inclusive labour markets (Sustainable Development Goals 5 and 8).

The business training package evaluated in this study is the ILO's Gender and Enterprise Together (GET Ahead) programme, which promotes enterprise development among women in poverty who want to begin, or are already engaged in, small-scale businesses. GET Ahead is a modular training package aimed at supporting primarily semi-illiterate women, with or without an income-generating activity, in order to build their basic business skills and competencies to ensure they start, run and/or grow their small businesses effectively.

The study was set in rural markets in Kenya with the objective of testing how the GET Ahead programme affects the profitability, growth and survival of female-owned businesses, and to evaluate whether any gains in profitability come at the expense of other business owners. A year-and-a-half after the training had taken place, a mentoring intervention was randomly assigned among trained women to test whether additional group-based and in-person support strengthens the impacts of training on intended outcomes. See evaluation questions in Box 1.

Box 1: Key evaluation questions

The impact evaluation aimed to measure the causal impact of GET Ahead on profitability, growth and survival of female-owned businesses; to test whether mentoring services strengthen the impacts of business training on intended outcomes; and to evaluate whether any gains in profitability came at the expense of other business owners. Specific questions included:

1. Mechanisms: Does training lead to changes in business knowledge and practices? To what extent does training work through teaching new skills and to what extent through increasing women's self-confidence to make changes and risks and willingness to negotiate? Do business knowledge and practices change from complementary mentoring services?
2. Impacts on treated businesses: Does implementing the practices taught in training lower costs, raise revenues and increase profits? What do better business outcomes mean in terms of how women run and operate businesses, the control they have over their business and how they balance work and their daily lives? Do profits, revenues and costs change from complementary mentoring services?
3. Impacts on other businesses: Do firms that were not offered training suffer a reduction in sales and profits when firms around them are trained? Does having firms in a marketplace that received mentoring lead to positive or negative impacts on the profitability, sales and business survivorship of non-treated firms?

The study relied on a randomised experiment in 157 rural markets in four counties in Kenya's Western and Eastern regions. With a sample of 3,537 firms, the design followed a two-stage randomisation process, first randomising at market level and then randomising the offer of training to individuals within treated markets. The follow-up mentoring intervention was randomly assigned to surviving businesses within markets comprising 16 or more firms. Four rounds of follow-up surveys with low attrition were used to measure impacts at one and three years after training.

The unique value added by the study stems from the enormous popularity of business training as a policy option for small business development around the world, coupled with three key gaps in the existing literature related to both programme content and analysis. These gaps are:

- **Gender-blind training content:** Most standard business training programmes attempt to provide basic business skills, but typically ignore the role of gender in business operations, constraints and solutions. In response to this perceived gap, GET Ahead addresses the practical and strategic needs of low-income women in enterprise by strengthening their basic business and management skills. It shows women how to develop their personal entrepreneurial traits and obtain support through groups, networks and institutions in the enterprise development field (Bauer et al. 2004). GET Ahead was first launched in Thailand in 2001 and has so far reached 18 countries around the developing world. One impact evaluation of the same programme took place in Vietnam, showing firm growth 12 months after training (Bulte et al. 2016).

GET Ahead seeks to create a 'business mind' among low-income women through a participatory and experiential approach that uses group and play exercises. Through such exercises, women learn about product design, the production process and underlying costs and sales strategies. The programme also discusses household and societal constraints faced by women in business and strategies for overcoming them.

- **Untested value of mentoring in boosting training effects:** By intensifying the treatment effect of GET Ahead through mentoring, the study sought to contribute to the limited evidence base on the role of mentorship and its combination with business training in developing contexts. Mentoring exhibits great potential to enhance entrepreneurial skills and business success. Its pedagogy fits well with the preferred learning styles and psychosocial needs of entrepreneurs, but its impact on mentees remains unclear. Karlan et al. (2012) find that business advice from international consultants does not lead to improvements in the performance of small tailoring enterprises in Ghana, while Valdivia (2012) finds that individual and group visits during a three-month follow-up after training led to improved business practices among Peruvian women, vis-à-vis women who received training only.
- **Systemic design failures in experiments of business training programmes:** The recent rapid growth in impact evaluations of business training programmes and subsequent reviews (McKenzie 2011; Grimm and Paffhausen 2015; McKenzie and Woodruff 2015) have identified three key limitations in study design:
    - A significant number of studies suffer from low statistical power and are therefore unable to rule out large positive or negative impacts of the training (McKenzie and Woodruff 2013);
    - Studies are often unable to trace sources of improvement in revenues and profits through the causal chain; and
    - There are no measurements of the extent to which treated firms derive gains from reallocation versus market growth. A partial exception is Calderón et al. (2012), who attempt to measure spillovers across markets in rural Mexico. The study randomised 17 markets into 7 treated and 10 control groups, but the small sample size coupled with high attrition left little power to test the hypothesis.

As public and private investments in enterprise development increase, the debate about spillovers from business training gains relevance and calls for more and better evidence (Rotemberg 2014). This concern is particularly apparent when working with microenterprises in rural markets in developing countries, where it is easy to believe that if firms are all selling similar products in a small market, any extra sales made by trained firms must come from competing to draw these sales away from neighbouring untrained firms. Conversely, if better management practices are best thought of as a technology (as in Bloom et al. 2015), then there might be positive spillover benefits as untrained firms copy practices adopted by trained firms, and therefore become more productive. Moreover, if training leads firms to expand the variety of products offered and customer services provided, the overall market size may increase, particularly in underdeveloped marketplaces.

This study builds on the existing literature and above-mentioned gaps to test, through a gender lens, the impact of business training and mentorship on entry, exit and reallocation at the market level, as well as impacts on individual firms within markets and the subjective wellbeing, confidence and other soft skills of female entrepreneurs.

This report addresses all questions presented in the Pre-Analysis Plan (PAP) submitted to 3ie in February 2014. In addition to the introduction, the report is organised as follows: a description of the GET Ahead programme and the mentorship offer, with underlying theories of change and research hypotheses; the context of the study and timeline; and the evaluation's design, methods and implementation. It also details an impact analysis, a discussion of the results and implications for policy and practice. Further information is displayed in the appendices.

2. Intervention, theory of change and research hypotheses

2.1 The business training intervention: The Get Ahead Programme

The training package provided is the ILO's Gender and Entrepreneurship Together programme: Get Ahead for Women in Enterprise. In contrast to conventional business training materials, GET Ahead brings a gender lens to enterprise development, exploring and addressing the needs of low-income women in enterprise by strengthening their basic business and management skills. The programme seeks to develop trainees' personal entrepreneurial traits, as well as their ability to work in and benefit from their groups, local networks and institutions (Bauer et al. 2004). The programme began in Thailand in 2001 and has been carried out in at least 21 countries, serving approximately 400,000 women.

The programme has four key modules, with the following themes:

- Module 1: Basics on gender and entrepreneurship. The module introduces GET Ahead and delivers basic concepts on the promotion of gender equality between men and women and the life cycle of people and enterprises.
- Module 2: The business woman and her environment. The module focuses on raising awareness among women about their strengths and weaknesses, as well as the current or future working environment and its effect on the business.
- Module 3: The business project. The module offers trainees information on the development of business ideas, opportunities and challenges, as well as basic information about marketing, production, services and technology, and financing, including costing and bookkeeping.
- Module 4: People, organisation and management. The module focuses on soft skills involved in managing a business (or a family business), with information on management of self and others, business support and networking, and action planning.

Some of the topics covered throughout these modules are not often emphasised in general business training programmes; for example, conversations about the difference between sex and gender and the role of cultural constraints in shaping women in business; dividing household and business tasks; and how to network with other women and the role of women's associations. In addition, it covers topics more typical of standard programmes, such as record-keeping and bookkeeping; separating business

and household finances; marketing; financial concepts; costing and pricing; generating and fine-tuning new business ideas; setting smart objectives; and traits required for business success.

The programme seeks to create a 'business mind' among low-income women engaged in small-scale businesses. The training methodology is participatory, with practical exercises to teach concepts. Modules expose women to different situations and environments, simulating real business life. For example, women learn about different costs involved in production and how to account for their own costs through making lemonade. They have role play exercises to practise different sales strategies for customers; and make necklaces to discuss a production process and the importance of different factors in product design.

The course took place over five days. All trainers had at least five years' experience in training small firms, as well as tertiary qualifications. Training took place in two-to-three locations per county. The locations were chosen to be relatively central to clusters of marketplaces, and were typically held in local hotels or church buildings. Training was offered for free and participants were provided transport subsidies of approximately US$6 per day to cover the costs of travelling from their residences to these locations (an average of 14 kilometres). The cost of providing the training is estimated between US$222 and US$333 per woman trained.[1]

Training took place immediately after the baseline surveys in each county, between June and November 2013.[2] Of the 1,172 individuals assigned to training, 77.7 per cent attended at least one day of training. Of the individuals who attended at least one day, 94.6 per cent attended all five days. Diwan et al. (2014) report on a choice structure experiment linked to this study, which intended to increase training attendance. They discuss the correlates of attendance and find that age and marital status are strong and statistically significant predictors of attendance. All else being equal, women aged above 35 are 35 percentage points more likely to attend training than those below 35, while married women are 24 percentage points less likely to attend than unmarried women. This potentially reflects the competing demands on their time from other household tasks. Women are also more likely to attend if they have previously participated in training (perhaps reflecting greater perceived benefits from attending), have a large household (potentially providing more people to undertake household and business tasks in their absence) and are located closer to the training venue (reducing travel time). Women who earn more profits are less likely to attend, perhaps reflecting a higher opportunity cost of time, or that they think there is less need to improve.

2.2 The mentoring intervention

After the one-year follow-up results, and emerging evidence suggesting that many business training programmes may be too short to show sizeable impacts (McKenzie and Woodruff 2013, 2015), the principle investigators decided to add a mentoring component for half the sample assigned to training.

[1] The smaller number reflects workshops with 20 attendees, while the larger number is for workshops with 30 attendees.
[2] The baseline data and a baseline report and paper based on encouraging take-up: <http://microdata.worldbank.org/index.php/catalog/1985>.

The mentoring intervention provided personalised, hands-on problem-solving support and peer learning to women who had previously participated in the GET Ahead programme, with the goal of reinforcing intended business training outcomes – from improved management skills to business growth. The intervention targeted 446 women who had been exposed to the business training in 2013 and expressed interest in further support through mentorship. Two local public service providers, Kenya Industrial Estate (KIE) and the Women Enterprise Fund (WEF), were identified as partners in delivery of the mentoring services. KIE implemented mentoring in Kakamega and Kisii, while WEF focused on Embu and Kitui.

The design of the mentoring intervention combined group and individual sessions for a period of five months (July-November 2015). Each female-owned firm (or mentee) received 15 mentoring sessions: 10 through group sessions and 5 through one-to-one meetings with the mentor. Group sessions occurred twice a month, every two weeks, while individual sessions took place once a month. Each mentor was assigned a group of five mentees. The table below summarises the topics covered during the 10 group mentoring sessions. Individual sessions deepened discussions on the above topics based on the needs of the mentee and her business.

The ILO contracted a mentoring expert for programme development and curriculum design, which were subsequently discussed and delivered to mentors from KIE and WEF.

- A total of 110 mentors were recruited; only 100 were subsequently trained, and 89 selected for the programme.
- The programme reached out to 446 women who had participated in the GET Ahead programme; 392 signed up for the programme, were inducted and received training. However, only 361 women stayed in the programme throughout the five months.
- Therefore, while 89 mentors were originally linked to 392 mentees, by the end of the programme the number of mentees per mentor ranged from 3 to 6.
- Reasons for dropping out included: (1) in most cases, a lack of interest after realising no grants were involved in the offer; and (2) in a few cases there were considerations about the mentoring programme not being helpful in enhancing skills and business growth.

Monitoring and evaluation of the mentoring was done through: (1) monthly meetings with mentors and implementing partners; (2) monthly reports per mentor; (3) tracking of journals filled out by mentors after each session; (4) visits and phone calls with mentors for follow-up and support; (5) two meetings between the ILO and implementing partners throughout the duration of the programme; and (6) one closing forum in each county to gather feedback from all parties.

Mentors were female business owners of a similar average age to the study sample, with 75 per cent having had post-secondary education and 68 per cent having studied business or accounting, and with a median of 5.5 years of business experience. They were recruited and trained by WEF and KIE.

Out of the 392 women who signed up for the mentoring programme, 388 attended at least one session. Conditional on attending, the median went to 77 per cent of all

sessions. Take-up is thus 388/524 (74%) of those assigned, and 388/446 (87%) of those in business at the time of the intervention. The cost of providing mentoring is estimated at US$553 per women trained.[3]

2.3 The theory of change and research hypotheses

The aim of business training programmes is to help trainees and mentees grow their businesses. Consider a standard firm production decision, in which a firm owner with entrepreneurial ability θ, and production function $f(.)$, chooses inputs of labour L and capital K to maximise profits:

$$\pi = pf(\theta, K, L) - wL - rK \quad \text{s.t. } K \leq \tau A \qquad (1)$$

Where p is the market price of their product, w and r are the prevailing market input prices for labour and capital, and τ reflects how binding credit constraints on capital acquisition are, given wealth level A.

Business training and mentoring can then potentially increase profit levels through several channels. Increasing entrepreneurial ability may allow the business owner to produce more with the same inputs, giving the business more to sell. Better upkeep of records and financial accounts may make it easier for the owner to obtain financing, thereby alleviating credit constraints and again allowing the business to produce more. If markets are perfectly competitive, firm owners will be able to sell all they produce at market price p and there will be no spillover impact on other firms in the market.

However, if firms are competing with one another for customers, then the price p need not remain constant as firms produce more. Better marketing practices, better customer service, and more efficient cost control may enable trained firms to compete away customers from other firms in the market. As a consequence, any gains to treated firms could come at the expense of lower sales and lower profits for non-treated firms in these markets.

Conversely, one of the hopes of the training providers is the possibility of positive spillovers to other firms in the market. A first potential channel is that trained business owners might teach learned skills to others in the market, or that others in the market may be able to observe and adopt the new practices, irrespective of whether they are directly taught. Second, the GET Ahead training emphasises the role of networks and women working together to seek financing or purchase products together to obtain bulk discounts, which could result in lower costs and higher profits for both treated and untreated firms. Implicit in this discussion is the assumption that the production function $f(.)$ is itself unchanged. But business training may encourage firm owners to consider producing new products or offering new services. This offers another means for treated firms to grow without negatively impacting non-treated firms – they can expand the market through increasing variety.

[3] This estimate is based on the number of women who attended at least one session.

The theory of change tested by the impact evaluation of the business training programme is as follows:

Step 1: When training is available, individuals take it up. Take-up is an important element across various types of training interventions and can strongly affect the feasibility of experimental designs. Studies with randomised invitations to attend training often struggle with low take-up and widespread heterogeneity of firms, which dramatically reduces power (McKenzie and Woodruff 2015).

Step 2: Women stay in the training and subsequently, as a result of training, improve their business knowledge.

Step 3: The acquired business knowledge prompts women to implement some of the ideas and practices learnt.

Step 4: Implementation of these practices improves business profits and revenues.

Step 5: Improved business outcomes lead to real improvements in women's lives. Some studies have suggested that assisting women's economic advancement may have positive impacts on empowering them in their homes and communities, but could also have some unintended negative consequences, such as increased conflict in the household, increased time pressure or loss of business control.

The mentoring offer followed a similar logic assuming an underlying causality between mentorship and business knowledge, with a subsequent boost in business practices, followed by increased profitability or sales.

The theory of change assumes the availability of necessary inputs to carry out the training – from financial to human resources (including ILO staff and trainers).

The corresponding research hypotheses[4] are:

1. *Impact on primary outcomes for treated firms:* GET Ahead training and subsequent exposure to mentoring may have positive average impacts **on profitability, sales and business survivorship** for treated firms. Related underlying assumptions include:
    a. Treatment leads to higher survivorship;
    b. Treatment leads to business owners making higher sales and earning greater profits; and
    c. Treatment does not affect reporting errors.
2. *Impact on secondary outcomes for treated firms:* GET Ahead training and subsequent exposure to mentoring may have positive effects on **overall employment, empowerment, subjective wellbeing and household asset ownership**. The evaluation relied on a women's economic empowerment framework in order to examine broader aspects of women's agency, control over resources, and livelihoods. This hypothesis implies assumptions on treatment leading to:
    a. Individuals being more likely to be employed and earning higher income from labour;
    b. Greater empowerment of women in terms of decision-making around finances and business;
    c. Increase in subjective wellbeing and improvement in mental health; and

[4] The PAP expands further on evaluation features designed to test specific research hypotheses.

   d. Increase in household durable asset ownership.
3. *Impact on primary outcomes for non-treated firms:* Other firms in your marketplace receiving training and mentoring may have negative impacts on the **profitability, sales and survivorship of non-treated firms**.
4. *Causal chain of process and mechanisms:* Training and mentoring should lead to increases in business knowledge and greater deployment of business practices, and potentially increase profitability or sales through increased marketing, better responsiveness to down periods, greater access to finance, better inventory management, better avoidance of household demands and increased self-confidence. Better business practices are strongly associated with better firm performance across a range of countries in both the cross-section and over time (McKenzie and Woodruff 2015). Some assumptions supporting this causality include:
   a. Treatment leads to an increase in business knowledge and increased use of standard business practices.
   b. Increased marketing and better presentation of business results in treated firms gaining more customers, potentially from non-treated firms.
   c. Better record-keeping enables businesses to keep better track of business conditions and react more quickly to down periods (Drexler et al. 2011; Karlan and Valdivia 2011).
   d. Treated firms are able to obtain more financing from banks and microfinance organisations and build up higher levels of inventories and capital stock.
   e. Better inventory management and better purchasing practices enable firms to reduce spoilage of stock and take advantage of bulk discounts, thereby lowering business expenses.
   f. Trained women have greater self-efficacy in entrepreneurial tasks and devote more time to working in their business.
5. *Heterogeneity of impacts:* The study focuses on three main dimensions of treatment effect heterogeneity with regard to how the effect of treatment on spillovers varies with the scope for competition, and how effectiveness on individual firm owners varies with their existing skill levels and empowerment.
   a. The first hypothesis assumes that the negative spillover effect of treatment on non-trained firms will be stronger for firms facing more competition and firms producing a higher share of their goods for the local market.
   b. The second assumption is that treatment is more effective for women who initially have higher skills.
   c. It is lastly assumed that treatment is more effective for women with higher initial levels of empowerment.

3. Context

Small enterprises are a major job creation engine. There are approximately 420 to 510 million small- and medium-sized enterprises worldwide, of which 80-95% are in low- and middle-income countries (IFC 2010). The smallest enterprises in developing countries are often low-productivity and informal, but an important source of income vital for livelihood and coping strategies. Similarly, with regard to female entrepreneurs, the incidence of small and informal businesses increases vis-à-vis male entrepreneurs (ILO 2015; World Bank 2012).

In Kenya, the situation is no different. Across the country, women face severe disadvantages compared to men. They are less likely to find a job in the formal economy, and earn lower wages when they do. In rural areas, women have less access to productive assets (e.g. land and credit) and earn lower incomes than men for comparable farm work (World Bank 2014). Furthermore, female entrepreneurs in Kenya cite a lack of training opportunities as a barrier to business growth (Lock and Lawton Smith 2015).

In 2012, the ILO launched the Women's Enterprise Development programme in Kenya with the objective of providing training and support services for women who wished to start and run businesses, while employing a gender perspective that contributed to their empowerment. The GET Ahead training was a contributing component to the programme.

Once the ILO identified GET Ahead as a relevant and likely subject for an impact evaluation, a multi-stakeholder retreat in October 2012 was used to pre-select 10 of Kenya's 47 counties as possible locations for the study.

A subsequent, more detailed participatory review process involving the ILO, government ministries and organisations serving female entrepreneurs[5] was used to select four counties for the study: Kakamega and Kisii in Western region; and Embu and Kitui in Eastern region. These regions are largely rural, each with an average population of approximately 1 million, the majority of whom are below the poverty line.

In each of the four counties, field staff from Innovations for Poverty Action (IPA) Kenya mapped out all market centres deemed as 'medium' or 'large' outside of the main cities and conducted a market census of all female-owned businesses between June and November 2013 (see Timeline). Altogether, 6,296 female-owned businesses in 161 markets were listed. After the census, three markets in Kakamega county were dropped due to the small number of women in these markets. An eligibility filter was applied to determine which women to include in the baseline survey.

This filter required the women to:

- Have reported profits, and not to have reported profits that exceeded sales;
- Have a phone number that could be used to invite them for training;
- Be 55 years or younger in age;
- Not be running a business that only dealt with phonecards or M-Pesa;
- Not be attending school; and
- To have at least one year of schooling.

It also required that:

- The person responding not be an employee;
- The business not have more than three employees;

[5] Stakeholders consulted included the Department of Micro and Small Enterprise Development of the Ministry of Labour, the Ministry of Youth Affairs, the Ministry of Cooperative Development and Marketing, the Ministry of Youth, the Federation of Women Entrepreneurs Associations, the Women Enterprise Fund (WEF), the Youth Employment Development Fund, business development service providers and Inoorero University.

- The business have profits in the past week of between 0 and 4,000 KSH; and
- Sales in the past week were less than or equal to 50,000 KSH.

These criteria were chosen to reduce the amount of heterogeneity in the sample (thereby increasing the ability to detect treatment effects), and to increase the odds of being able to contact and find individuals again. Applying this eligibility filter reduced the 6,296 individuals to 4,037 individuals (a 64% reduction). Baseline surveys took place soon after the listing surveys in each county, between June and November 2013. Out of a target of 4,037 individuals, enumerators were able to interview 3,537 (87.6%) in time to consider inviting them for the training.

The 3,537 individuals were located in 157 separate markets, which were typically small, remote, and largely consisting of women operating a limited variety of businesses, such as selling fruits, vegetables, grains and dried fish products from tables, as well as offering services, such as hairdressing, dressmaking and small food kiosks. The market is a designated place in the village, and authorities typically build several stalls for the sellers to display their goods and services and charge them a fee to sell in the market. The average market in the study has 22 firms surveyed in it, with 75 per cent in retail trade and 25 per cent in services.[6]

In order to address concerns about external validity, the study relied on existing evidence and meta-reviews (McKenzie and Woodruff 2013) to provide a sense of the range of outcomes of business training programmes in different locations around the world, as well as key design insights for the next generation of studies. Key factors influencing the design are: (1) studies typically measure outcomes only, with one follow-up; (2) they often lack statistical power; and (3) they are mostly unable to trace the source of any improvements in profits and revenues through the causal chain – in particular whether they arise from spillovers. The study is designed to build on this existing literature in part by overcoming these issues.

The design also attempts to tackle concerns about applicability and replicability of results, particularly the fact that many business training evaluations evaluate new programmes that are designed just for microfinance clients and have not been implemented in many other countries. The evaluated business programme (GET Ahead) has been implemented in over 12 countries and is expected to expand further; therefore, results from this evaluation are directly applicable to programme design at global level.

Finally, few studies look at contextual and social dimensions to explain economic empowerment. As explained in Golla et al. (2011), women face many social constraints to economic advancement and realising the benefit of their economic activities. The inclusion of a women's economic empowerment framework in this study sought to help frame the measurement of economic outcomes within a broader empowerment context.

[6] In addition, markets typically have a market day once a week, where outside vendors come to sell goods, such as clothing, plastic housewares and shoes. These vendors, who are only in the market one day a week, are not included in our study.

4. Timeline

- Listing and baseline survey: June-November 2013
- Business training intervention: June-November 2013
- Round 2 follow-up: June-October 2014 (one year post training)
- Round 3 short follow-up: November 2014-February 2015
- Mentoring intervention: July-November 2015
- Round 4 follow-up: February-July 2016 (three years post training; 6-10 months post mentoring)
- Round 5 follow-up: May-October 2016

5. Evaluation: Design, methods and implementation

5.1 Ethical measures

This evaluation was reviewed for ethical concerns by the Internal Review Board of IPA, Maseno University Ethics Review Committee, the National Commission for Science, Technology and Innovation, and ICRW for the qualitative part of the study. Ethical considerations avoiding harm to participants were strictly respected at all stages. These considerations, along with special gender considerations for interviewing women, were included in training for all survey staff. All staff involved in the project with access to personal identifying information all took a human subjects training course.

Potential risks of the study included people spending time on the study that might be better spent addressing other issues. This risk was addressed by securing informed consent and clarifying that study participants could drop out at any time, even in the middle of an interview or group discussion.

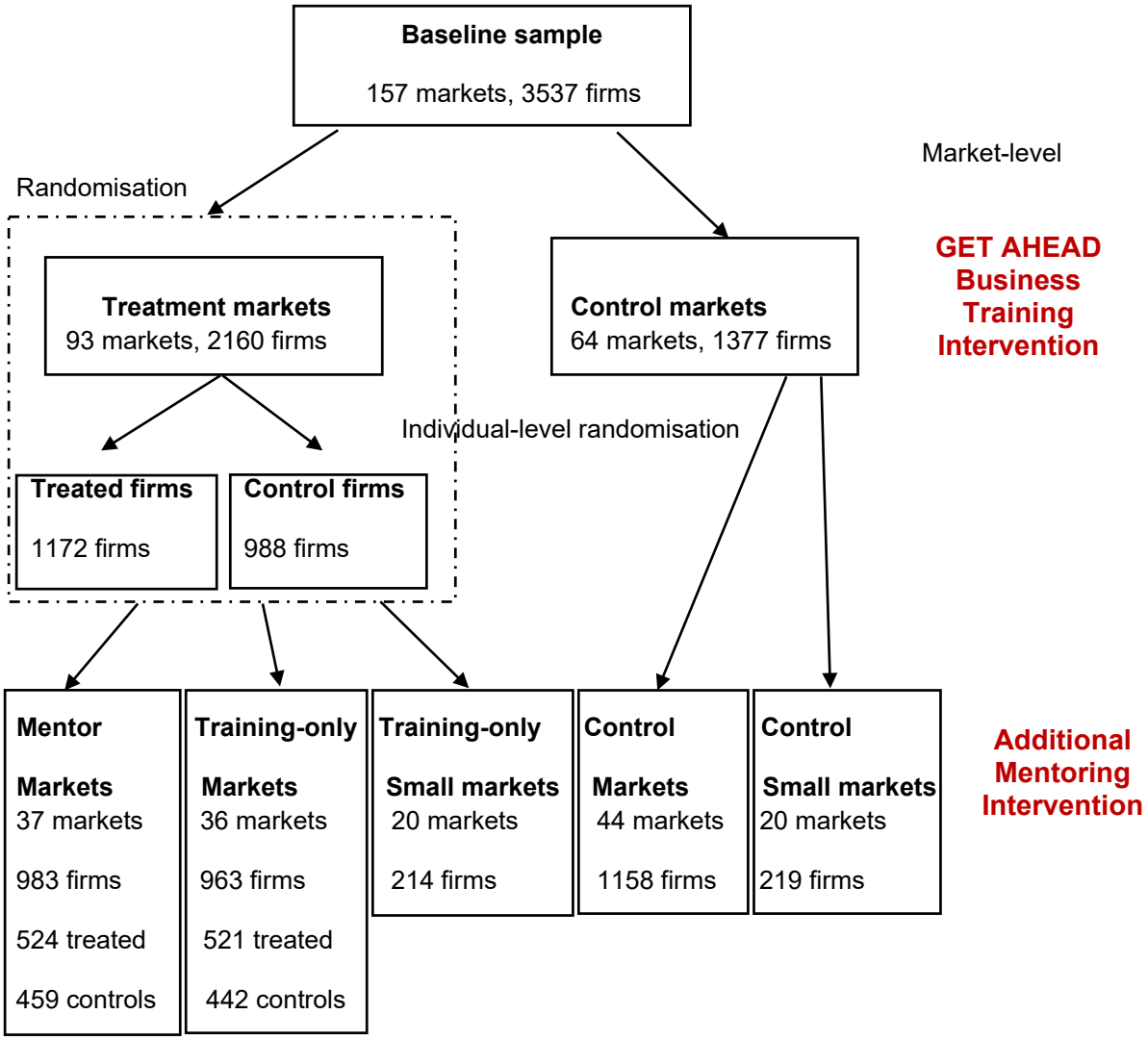
A key concern was confidentiality of sensitive business data. Women may not wish to share this information with other women or spouses. For this reason, questions on individuals' specific business outcomes were asked only in private interviews. Women's names have not been included in the final data set; instead they are identified by a unique ID number. IPA takes a number of precautions to ensure the confidentiality of all information collected from subjects in the studies it conducts. The names of focus group participants were not collected. For structured interviews and household surveys, subjects' names were recorded only on the first page of the survey form, which was detached before the survey booklets were given to the data entry firm for digitisation. Interviewers were trained to keep data confidential.

Ethical concerns could be raised regarding village-level randomisation of the training assignment, which could be interpreted as 'denying services' to women in need. The randomisation process was conducted with great transparency to ensure it appeared as a fair allocation of limited available spots in the training. The team also believed that the randomised approach would allow better understanding of the ethical implications of the spillover effects possibly generated by training programmes that may have increased sales for one trained business owner at the expense of a non-trained business owner.

5.2 Evaluation strategy

Individuals were assigned to treatment and control samples for the GET Ahead business training intervention in a two-stage process (Figure 1).

Figure 1: Randomisation design



Notes: The graphic shows assignment to treatment. The first intervention of business training assigned firms to the GET Ahead business training programme using a two-stage randomisation: first, markets were assigned as treatment or control; then, within markets, firms were randomly selected for training. Then at the second stage, markets with fewer than 16 firms at baseline ('small markets') were dropped, and half the training treatment markets were assigned to also have those assigned to training within the market receive mentoring. Mentoring was only offered on the condition that the firm survived.

First, **markets** were assigned to treatment (some individuals in them were invited to training) or control (no one in the market invited to training) status. Randomisation was performed within 35 strata defined by geographical region (within county) and the number of women surveyed in the market. The need to ensure sufficient numbers for training meant that more than half the markets were allocated to treatment: 93 were assigned to treatment and 64 were assigned to control.

Then, within each market **individuals** were assigned to treatment (invited to training) or control (not invited to training) within treated markets by forming four strata, based on quartiles of weekly profits from the census (≤450, 451-800, 801-1,500, 1,501-4,000), and then assigning half the individuals within each stratum to training. This resulted in 1,172 of the 2,160 individuals in treated markets being assigned to treatment and 988 assigned to control (Table 1).[7]

Table 1: Summary of random assignment

		Markets				Individuals		Control
County	#of Strata	**Total**	Treatment	Control	**Total**	Treatment in treated market	Control in treated market	in control market
Kakamega	9	**31**	19	12	**782**	257	220	305
Kisii	9	**34**	20	14	**844**	274	237	333
Embu	6	**36**	20	16	**715**	231	189	295
Kitui	11	**56**	34	22	**1196**	410	342	444
Total	**35**	**157**	**93**	**64**	**3537**	**1172**	**988**	**1377**

Table 2 provides some basic characteristics of firms at the market level for treatment and control markets. Randomisation succeeded in generating markets with comparable characteristics, with a test of joint orthogonality unable to reject the null hypothesis that average characteristics of the markets are unrelated to treatment assignment. The typical marketplace in the study has 22 surveyed firms, and the average firm earns 1,100 KSH (US$13) per week in profits on sales of 4,500 KSH. In the typical market, 75% of study firms are retail-based and 25% are service-based; 46% of firms in the average market claim to have a business licence, and only 8.5% say they have previously undertaken any form of training to help their business.

[7] When the number of firms within a stratum was odd, the additional unit was also randomly allocated to training.

Table 2: Verification of randomisation at the market level

Means of	Treated markets	Control markets	Test of equality p-value
Number of individuals in baseline	23.2	21.5	0.162
Mean weekly profits from census for baseline respondents	1133	1103	0.299
Mean weekly sales from census for baseline respondents	4648	4307	0.055
Share of firms in retail	0.75	0.75	0.934
Share of firms in services	0.25	0.25	0.959
Share of firms registered with city council	0.47	0.45	0.368
Share of firms with any form of previous business training	0.09	0.08	0.481
Test of joint orthogonality			0.403
Number of markets	93	64	

For the follow-up mentoring intervention, markets with 15 or fewer firms in the study were dropped, since these would have too few firms to make mentoring cost-effective. This left 73 of the 93 treated markets; 37 of these were randomly assigned (within the original randomisation strata) to mentoring, and the remaining 36 were assigned to training only. Within the mentoring markets, all individuals assigned to training were also to be offered mentoring services; however, in practice only surviving businesses were offered these services. A comparison was drawn between mentoring and training: only markets to the 44 markets in the control group with 16 or more firms.

To verify the random assignment, Table 3 provides key characteristics of the women selected in the sample and their firms, by treatment assignment.[8] The sample looks similar across the three groups; and again, the joint orthogonality of baseline characteristics cannot be rejected when comparing the treatment group to the spillover group, the treatment group to the pure control or the spillover group to the pure control.

The average woman in the sample is 36 years old, has 9 years of schooling, and has been running her firm for just over 6 years. Two-thirds of the women are currently married. The modal firm has no employees (only 20% have one or more employees). The mean firm earns 1,100 KSH (US$13) per week in profits on sales of 5,500 KSH (US$65) and has capital stock of 31,000 KSH (US$370).[9] One quarter have received financing from a bank or microfinance organisation; and 45 per cent are registered at local level. Only 35 per cent of firms keep business records at baseline; and on average firms are using just over half of the 26 business practices in the McKenzie and Woodruff (2015) index. This suggests scope for improvement resulting from business training.

[8] Table 1 in the Appendix offers information on the means by treatment status for mentoring.
[9] The exchange rate was approximately 1US$ = 84 KSH in 2013 at the time of baseline.

Table 3: Individual characteristics and verification of randomisation

	Means by treatment assignment						
	Treatment group	Spillover group	Pure control	Pure control	p-value (1) vs (2)	p-value (1) vs (3)	p-value (2) vs (3)
	(1)	(2)	(3)	S.D.			
Age	36.0	35.6	35.7	9.05	0.482	0.454	0.690
Years of education	8.92	8.91	9.09	2.92	0.910	0.569	0.515
Married	0.67	0.66	0.67	0.47	0.404	0.638	0.518
Household size	4.97	4.85	4.85	2.10	0.188	0.262	0.499
Age of firm	6.39	6.57	6.27	6.92	0.574	0.741	0.403
Number of employees	0.27	0.27	0.27	0.63	0.747	0.989	0.953
Weekly profits	1128	1140	1091	834	0.987	0.395	0.322
Weekly sales	5220	5885	5401	8048	0.065	0.266	0.288
Capital stock	30571	34092	29370	60530	0.248	0.863	0.101
Ever received bank/MFI loan	0.24	0.25	0.24	0.43	0.891	0.482	0.934
Keeps records	0.37	0.34	0.34	0.47	0.235	0.094	0.974
Business practices Score	0.53	0.53	0.52	0.19	0.934	0.487	0.598
Retail firm	0.77	0.76	0.75	0.43	0.446	0.121	0.474
Registered with city council	0.44	0.45	0.42	0.49	0.328	0.447	0.100
Joint orthogonality test p-value					0.597	0.435	0.691
Sample size	1172	988	1377				

Notes: Tests of treatment versus spillover group control for individual-level randomisation strata are based on robust standard errors. Tests of the treatment or spillover group compared to the pure control group control for market-level randomisation strata use standard errors clustered at the market level.

5.3 Power calculations: Business training

Many existing business training programmes and other private sector development programmes in Africa have very low statistical power (McKenzie 2011; McKenzie and Woodruff 2013). Lessons from existing literature supported the design of evaluation features aiming at increasing the study's power.

- McKenzie and Woodruff (2013) show that the coefficient of variation (the key metric for how varied firms are) ranges from 0.5 to 0.8 in two studies that explicitly screened firms on size (Berge et al. 2014; De Mel et al. 2014) to between 2 and 3 in most other studies. The study's listing and screening approach intended to bring the coefficient of variation down to at least 0.8, which is the first key parameter used for power calculations.
- To increase power, the study relied on multiple survey rounds, with the gain depending on the autocorrelation of profits. In the power calculations, rho was set to 0.4, following McKenzie's (2012) finding of an autocorrelation of 0.3-0.5 being common for business profits and revenues.
- Lastly, to increase power the study design aimed to secure high intervention take-up.

The above parameters supported power calculations to detect a given treatment effect for individual-level randomisation. The outcome of interest was the following: 'What is the impact of training on firms allocated to training as compared with firms not allocated to training in the same markets?' The calculation suggested a treatment and control group size of 207 firms each.[10] With 1,172 individuals in treated markets assigned to treatment and 988 to control, the study had plenty of power to detect the effect. In fact, the minimal detectable effect at 80 per cent power is a 10.9 per cent increase in profits. Allowing for an anticipated attrition rate of 10 per cent, this leads to a minimum detectable effect of an 11.6 per cent increase in profits.

The power for estimating the following questions – 'What is the impact of the training on firms allocated to training relative to a situation where no one is treated?' and 'What is the spillover impact of trained firms on non-trained firms?' – depends on a further parameter, which is the intra-cluster correlation of profits or revenues within a market. One advantage of the low autocorrelation and heterogeneity of profits for the same firm over time is that the intra-cluster correlation tends to be quite low – much lower than the 0.1-0.2 typically found in education and health studies.

Based on work in Egypt, Ghana and Sri Lanka, the intra-cluster correlation likely falls between 0.02 and 0.05. The study conservatively uses 0.05 and finds a sample size of 352 per group.[11] These power calculations do not account for any reduction in variances from the matched triplet randomisation or the stratification by sector within sample; coupled with the somewhat conservative assumption on the intra-cluster correlation, this suggests that the study would have reasonable power to detect at least a 15 per cent change in profits (or revenues).

5.4 Power calculations: Mentoring

Survival and profits were the focus of the power calculation for the subsequent mentoring intervention. The one-year survival rate in the control group is 0.90. Assuming that the failure rate will be 10 per cent per year for the next two years, then the control mean after three years would be 70 per cent survival, leading to:

- 90.8 per cent power to detect a 10 percentage point increase in the 3-year survival rate when comparing treatment to control within mentoring markets;
- 86 per cent power to detect a 10 percentage point increase in the 3-year survival rate when comparing treatment to the pure controls, using a clustered randomisation; or when examining spillovers by comparing the spillover controls in mentoring markets to the pure controls in pure control markets.

These calculations were conservative in that they are not conditional on randomisation strata fixed effects; if this were the case, the power would increase (Bruhn and McKenzie

[10] Based on STATA's sampsi command [sampsi 100 118.75, pre(1) post(2) sd1(80) r01(0.4) r1(0.4)] with ANCOVA to obtain the sample size needed to detect a 25 per cent increase in profits.

[11] Using the sampclus command in STATA after the sampsi command above, we find that the sample size needed to detect a 25% increase in profits under the same assumptions, as above, is 352 in each group (using the command sampclus, obsclus(15) rho(0.05) after the command given above), while the minimal detectable effect with 10% attrition and rho of 0.05 is approximately 20%.

2009). These calculations were based on intent to treat (ITT) effects; since the mentoring intervention is conditional upon individuals who have already agreed to be in the study for two years, mentoring take-up was expected to be at 90% or higher, in which case these ITTs of 10% correspond to local average treatment effects (LATE) of 11.1%.

With regard to profits, enterprises are purposively screened at baseline to make them more homogeneous in size than most business training studies. As a result, the baseline coefficient of variation for profits was 0.8, compared to coefficients of variation of 1.5-2 in many existing studies (McKenzie and Woodruff 2013). The intra-cluster correlation is low at 0.03 so that clustering does not reduce power much compared to pure randomisation. Two rounds of follow-up surveys were conducted within close proximity of one another and averaged for power (McKenzie 2012). The autocorrelation between profits at baseline and these follow-up profits was assumed to be 0.3, while the autocorrelation in profits one month apart is 0.7. Based on weekly profits of 1,100 KSH for the controls at baseline, the following was calculated:

- 83 per cent power to detect a 13 per cent increase in profits when comparing treatment to control within mentoring markets;
- 81.9 per cent power to detect a 13 per cent increase in profits when comparing treatment to the pure controls using the clustered randomisation, or when examining spillovers by comparing the spillover controls in mentoring markets to the pure controls in pure control markets.

Again, actual power would be greater after allowing for the use of randomisation strata fixed effects. As a result, the experiment continued with confidence on the study's power to detect effects that were of economically meaningful size – in particular, the size seen in other studies in the literature (McKenzie and Woodruff 2015).

5.5 Follow-up surveys

Four rounds of follow-up surveys were conducted to measure outcomes approximately one year and three years after training occurred (see Timeline). Two types of survey were used. A comprehensive long-form survey collecting data on a wide range of business outcomes was used in rounds 2 and 4. These were supplemented by much shorter surveys in rounds 3 and 5. These short surveys were conducted two or three months after the long surveys, and were intended to provide a second observation on volatile business outcomes, such as sales and profits, as well as an additional opportunity to gather data from individuals who could not be found at the time of the long survey.

Table 4 details response rates. Overall, 95 per cent of the sample were interviewed in at least one of rounds 2 or 3, and 92.3 per cent were interviewed in at least one of rounds 4 or 5. In addition, in cases of refusal, travel, death or for other reasons, information was collected from other household members or close contacts on whether the individual in the sample was currently operating a business. This offered survival status for 99.3 per cent of the sample at one year and 97.2 per cent at three years. There is no significant difference in data availability with treatment status at the three-year horizon, although those assigned to treatment are 1-2 percentage points more likely to have data available at the one-year horizon.

Table 4: Data availability by treatment status

	Round 2	Round 3	R2 or R3	Round 4	Round 5	R4 or R5
Panel A: Interviewed						
Assigned to training	0.034***	0.030***	0.023***	0.009	0.023*	0.010
	(0.012)	(0.011)	(0.008)	(0.013)	(0.013)	(0.011)
Spillover group	0.014	-0.013	-0.003	-0.026*	0.002	-0.009
	(0.013)	(0.014)	(0.010)	(0.014)	(0.015)	(0.012)
Pure control mean	0.886	0.889	0.943	0.894	0.876	0.923
Panel B: Data on survival available						
Assigned to training	0.011*	0.016**	0.002	0.005	0.010	0.003
	(0.006)	(0.006)	(0.003)	(0.009)	(0.010)	(0.007)
Spillover group	0.015**	-0.003	0.002	0.000	0.008	0.004
	(0.006)	(0.008)	(0.004)	(0.010)	(0.010)	(0.007)
Pure control mean	0.968	0.962	0.991	0.947	0.924	0.970
Panel C: Data on weekly sales and profits available						
Assigned to training	0.027**	0.031***	0.016**	0.013	0.031**	0.011
	(0.012)	(0.009)	(0.006)	(0.013)	(0.013)	(0.010)
Spillover group	0.015	-0.006	-0.003	-0.007	0.011	0.006
	(0.013)	(0.012)	(0.008)	(0.014)	(0.014)	(0.011)
Pure control mean	0.907	0.913	0.964	0.903	0.881	0.939
Sample size	3537	3537	3537	3537	3537	3537

Notes: Robust standard errors in parentheses, clustered at the market level.
*, **, and *** indicate significance at the 10, 5, and 1 per cent levels, respectively.
Data on weekly sales and profits availability codes data as available if firm is known to be closed (since then sales and profits are known to be zero).

In addition to the survey data, two other sources of information aided in assessing impact. The first were photos of the businesses' inventories, which were taken at the time of the baseline survey and in rounds 2 and 4. Two independent field staff valued these inventories based on the market prices of the different items and averaged these values to get a photo-based measure of the size of the firm. They also visually compared the firms in photos to determine whether the business looked bigger in relative terms to what it was at baseline. Second, intensive qualitative work about impacts over the first year of the study was carried out by ICRW in 2015 (ICRW 2015).

5.6 Measurement of key outcomes:

- All nominal values were converted into real Kenyan shillings (August 2013) using the consumer price index for the midpoint of each survey round.
- 'Firm survival' is measured as whether the owner still operates a business, regardless of whether they have changed the business line. For individuals who could not be interviewed, survival was measured by asking family members and neighbours about the status of the owner.
- 'Daily sales' are sales of the business in the last day, coded as zero if the business was closed that day or is closed for good. It is truncated at the 99th

percentile (i.e. observations above the 99th percentile were replaced with the 99th percentile).

- ‘Weekly sales’ are total sales of the business in the last week, coded as zero if the business was closed that week or is closed for good. It is truncated at the 99th percentile. When aggregated across all firms in the market, this forms total market sales.
- ‘Main product sales’ are obtained by multiplying the number of units of the main product sold in the last week by the unit price, truncated at the 99th percentile and coded as zero if the business is closed.
- ‘Business profits’ are measured using the direct question of de Mel et al. (2009), using the previous week as a reference period since pre-testing found that a weekly rather than monthly recall was easier for business owners to answer: ‘What was the total income the business earned during last week after paying all expenses including wages of employees, but not including any income you paid yourself? That is, what were the profits of your business in the last week?’ This is coded as zero if the business is closed, and truncated at the 99th percentile. When aggregated across all firms in the market, this forms total market profits.
- ‘Main product profits’ are obtained by multiplying the mark-up on the main product sold by the number of its units sold in the last week, truncated at the first and 99th percentile, and coded as zero if the business is closed.
- ‘Photo inventories’ are the value of inventories as assessed by valuing a photograph of the business inventories. A common set of market prices is used to aggregate products. Two independent enumerators would count the number of each product they could see in the photo (e.g. 53 tomatoes) and then aggregate by the price per product. If the two valuations differed by more than 5,000 KSH, they would iterate again until they agreed on a valuation.
- ‘Aggregate index of profits and sales’ is the average of standardised z-scores of the primary profits and sales measures.
- ‘Employed for pay’ is coded as 1 if they are self-employed or worked for wages in the last week.
- ‘Income from work’ is the sum of weekly profits and income gained from wage work in the last week.
- ‘Empowerment index’ is the sum of the following outcomes: compelled to spend money on husband or family (coded 1 if answer is no); not the only person with access to their firms’ money (coded as 1 if only they have access); has some money that they have sole control over and can spend how they like; and do not need anyone’s permission to visit a friend, travel to sell a business asset, travel to a new location to work, stay overnight in a different town, work later than usual hours, take out a loan or spend money on an investment for their business.
- ‘Life ladder today’ and ‘life ladder five years’’ are measured by a standard 10-step Cantril ladder, where individuals are asked to imagine the best (step 10) and worst (step 1) possible lives for themselves, and then say which step represents their current position and where they will be in five years.
- ‘Mental health’ is measured by the MHI-5 index of Veit and Ware (1983), coded so higher scores denote better mental health.
- ‘Household durables index’ is the first principal component of dummy variables for ownership of 10 household assets (iron and heaters, fridge or freezer, fan,

sewing machine, radio or CD player, TV or DVD player, motorcycle or scooter, car or van, oven, and gas cooker) and of the number of cows and goats owned.

- ‘Number of new entrants’ is the number of firms operating in the market outside of the experimental sample that have opened since the baseline survey and training intervention.
- ‘Total other firms’ is the total number of other firms operating in the market outside of the experimental sample (new entrants plus those pre-existing firms that were not included due to absence from the market at the time of listing, or being dropped by the eligibility filters).
- ‘Other firm profits’ is the total profits in the market of the other firms outside the survey sample.
- ‘Weekly customers’ is the number of customers the firm has had in the last week, truncated at the 99th percentile. This is coded as zero for firms that are closed. When aggregated to the market level, this forms total market customers.
- ‘Business knowledge’ is the number of correct answers out of seven questions intended to measure whether the firm owner can calculate sales, expenses and profits.
- ‘Business practices’ is the proportion of 26 practices in marketing, record-keeping, buying and stock control, and financial planning used by the firm (McKenzie and Woodruff 2015). This is only measured for firms surviving at the time of the survey.
- ‘Worked with a mentor’ is a dummy variable for whether they have worked with a mentor to try to improve their business in the last year (only asked in round 4).
- ‘Entrepreneurial self-efficacy’ is the number out of 10 of business activities in which the owner rates themselves as ‘very confident’ in their abilities (only asked in round 2). This includes entrepreneurial tasks, such as ‘Estimate customer demand for a new product’, ‘Persuade a bank to lend you money’ and ‘Identify good employees’.
- ‘GET Ahead attitudes’ is the sum of scores from 11 questions designed to measure the attitudes that GET Ahead training is meant to encourage. These are scored 1 to 5, where 1 = strongly disagree and 5 = strongly agree. Questions will be coded so that higher scores indicate better entrepreneurial attitudes. Examples include, ‘Even when my business is going well, I keep my eyes open in case I find a way to improve it’; ‘I don’t worry about where my business will be in the future – I just plan week-to-week based on what comes up’ (negatively coded); and ‘My business provides about the same as others/is doing about the same as others, so there’s no need to make it better.’ (negatively coded).

6. Programme: Design, methods and implementation

Since 2004, the ILO has introduced women's enterprise development interventions through technical cooperation projects and a dedicated group of trainers in over 12 countries. GET Ahead is a management training programme designed by the ILO with a strong gender component that focuses on helping women in poverty who want to begin or are already engaged in business to strengthen their basic management skills. The programme:

> differs from conventional business training materials by highlighting essential entrepreneurial skills from a gender perspective...It addresses the practical and strategic needs of low-income women in enterprise by strengthening their basic business and people management skills. It shows women how to develop their personal entrepreneurial traits and obtain support through groups, networks and institutions dealing with enterprise development (Bauer et al. 2004).

GET Ahead is often complemented by other measures aimed at helping women entrepreneurs to access additional business services as a means of starting or consolidating businesses. These services have in the past included supporting access to finance, financial literacy training and markets, and building the capacity of women entrepreneurs' associations to help them better defend and advocate for their needs and rights, etc.

The following table provides further details on the activities delivered:

Table 5: Model for a GET Ahead five-day workshop for entrepreneurs

Day 1	Day 2	Day 3	Day 4	Day 5
• Opening • Gender equality promotion: life cycle of people and enterprises • The business woman: she can do it	• The business environment: she is not alone • Business ideas • Marketing	• Production, services and technology • Marketing	• Finance • Management of self and others • Business support and networking	• Management or networking • Action Planning • Closing

The programme's mentors were female business owners of a similar average age to the study sample, with 75 per cent having had post-secondary education and 68 per cent having studied business or accounting, and a median of 5.5 years of business experience. They were recruited and trained by implementing partners WEF and KIE.

Each mentor was assigned to a group of 3-6 mentees. The programme then entailed 10 group sessions in which the mentor met the mentees, with 1 session held every 2 weeks for a period of 5 months, following a standardised structure (Table 6). In addition, the mentor would meet with the individual mentees once a month over this period to provide individualised guidance. The mentoring process aimed to reinforce the business training by asking the business owner to identify core goals for their business, consider where gaps exist, explore options to fill these gaps and then take action towards meeting these goals.

Table 6: Mentoring approach and activities

Session	Objective/Module	Outcome	Time
Session 1: Introduction, objectives and agreement	• Introduction of mentor/mentees • Definition of objectives • Understanding the mentoring agreement	• Knowing each other • Being clear on the mentorship objective and expectations • Understanding and signing the mentoring agreement	2 hrs
Session 2: GROW goal	Goals review	Mentees' goals established • How the mentees will appear when they are successful • How the business will appear after growing, with regard to sourcing, production, marketing, packaging, sales, networking, access to finance and general business management	2 hrs
Session 3: GROW reality	Reality analysis	Current status established • Where is the mentee right now in light of her personal and business goals? • Business performance established in terms of sourcing, production, marketing, packaging, sales, networking, access to finance and general business management	2 hrs
Session 4: GROW gap analysis	Identifying key gaps (goals vis-à-vis current status)	Gaps identified by exploring the difference between the goal and the current reality	2 hrs
Session 5: GROW exploring options	Exploring options to address identified gaps	A list of options to close each of the gaps identified (in terms of capabilities, skills and assets)	2 hrs
Session 6: Feasibility assessment	Financial analysis	• Exploring the feasibility of filling in the gaps • What is the cost of implementing the options? • What will be the source of finance required to close the gaps?	2 hrs
Session 7: GROW way forward/action points	Plan of action to address identified gaps	• Analysing options and agreeing way forward • Drawing up an action plan with clear timelines and budget	2 hrs
Session 8: Implementing the way forward	Evaluating progress	• Exploring what the mentee has done on the agreed action points • Identifying successes, challenges, way forward	2 hrs
Session 9: Implementing the way forward	Evaluating progress	• Exploring what the mentee has done on the agreed action points • Identifying successes, challenges, way forward	2 hrs

Session	Objective/Module	Outcome	Time
Session 10: Evaluation, review of action plan and sustainability	End of programme evaluation, review of action plan and agreement on sustainability plan	• Documenting the impact of the mentorship (i.e. how the mentee and business have benefited from the mentoring sessions) • Reviewing the action plan based on outcome of sessions 8 and 9 • Agreeing on a sustainability plan after the end of formal mentorship	2 hrs

The evaluated intervention sought to identify 1,200 female small business entrepreneurs in 2-3 counties in Kenya. As the evaluation design advanced, a participatory process was established to support county selection. A total of 10 out of 47 counties across Kenya were pre-selected as possible locations for the study. A more detailed review of these 10 counties was then used to select four counties in which training was ultimately provided: Kakamega and Kisii in Western region; and Embu and Kitui in Eastern region. Ultimately, 1,172 individuals were assigned to business training in treated markets and 988 were assigned to control.

The final beneficiaries of the interventions were women entrepreneurs. Other beneficiaries included business development service providers who were trained to provide gender-responsive services to these women. All intermediaries and providers were aware that the intervention was the subject of a rigorous impact evaluation.

The evaluation involved close collaboration between researchers and programme implementers in the field, which had the enormous advantage of allowing the results to have immediate and strong links to policy and practice.

Roles and responsibilities:

- The research was led and managed jointly by the ILO and the World Bank. The lead principal investigator was David McKenzie, an expert in impact evaluation of firm programmes, having published over 100 papers and completed randomised experiments in more than 15 countries.
- The training and mentoring programmes were managed by the ILO team (Jane Maigua, Virginia Rose and Susana Puerto) with support from Valerie Breda.
- The local implementing partners in the delivery of the mentoring programme were WEF and the KIE, semi-autonomous government agencies with a focus on women's entrepreneurship development. The partners were identified on the basis of their common mandate for promoting enterprise development, industrialisation and facilitating the graduation of micro- and small-sized enterprises into medium-sized enterprises, as well as their ability to influence public action in entrepreneurship in Kenya.
    - WEF is a semi-autonomous government agency in the Ministry of Devolution and Planning. It was established in August 2007 to provide accessible and affordable credit as a means of supporting women in starting and/or expanding their businesses for wealth and employment creation. WEF also provides business support services, such as capacity building, marketing, promotion of linkages and infrastructure support. It is a flagship project under the social pillar in Vision 2030, and therefore serves as a demonstration of the

Kenyan government's commitment to the realisation of the Millennium Development Goal on Gender Equality and Women Empowerment (MDG 3).
    - KIE is a government parastatal in the Ministry of Industrialisation. Established in 1967, KIE seeks the promotion of industrialisation through development and incubation of small- and medium-sized enterprises.
- Data collection tasks were the responsibility of IPA Kenya. IPA supported the implementation of the evaluation from its inception in 2013, including data collection and coordination and logistical support, to the implementation of the training and mentoring programmes.

There were no deviations from the study PAP. Strong monitoring measures were taken throughout programme implementation and the rollout of the evaluation to ensure design fidelity. IPA and ILO teams kept records of attendance at the trainings and mentoring sessions.

7. Impact analysis and results of the key evaluation questions

7.1 Estimation approach

A pre-analysis plan and the associated trial were registered on the American Economic Association's Social Science Registry on 21 February 2014. The assigned registry number is AEARCTR-0000287.[12] This plan pre-specified the primary and secondary outcomes of the study, the estimation approach and the causal chain to be investigated prior to the collection of any follow-up data. Subsequent funding enabled the implementation of the short follow-up surveys to collect additional data on profits and sales immediately following the long-form surveys, and to add the mentoring intervention.

Following McKenzie (2012), the analysis pools this short- and long-run follow-up data to obtain average effects at one and three years and uses an analysis of covariance specification where the baseline data are available in order to maximise power. Individual firm-level outcomes are examined by estimating for firm *i* in market *j* at time *t=2,..,5*:

$$Y_{i,j,t} = \beta_0 + \beta_1 T_{i,j} * 1Year_t + \beta_2 T_{i,j} * 3Years_t + \beta_3 S_{i,j} * 1Year_t + \beta_4 S_{i,j} * 3Years_t + \pi Y_{i,j,t=1} + \gamma BM_{ij,t=1} + X'_{k,i,j}\theta + \sum_{s=2}^{5} \delta_s 1(s = t) + \varepsilon_{i,j,t} \quad (2)$$

where $Y_{i,j,t}$ is the given outcome variable measured in round t; $Y_{i,j,t=1}$ is its baseline value and $BM_{i,j,t=1}$ is a dummy variable indicating whether this baseline value is missing; $T_{i,j}$ is an indicator for being in a treatment market and assigned to treatment; $S_{i,j}$ is the spillover term measuring whether firm *i* is a control firm in a market assigned to treatment; $1Year_t$ is a dummy taking the value one in follow-up rounds 2 and 3; $3Years_t$ is a dummy taking the value one in follow-up rounds 4 and 5; X_k is a vector of randomisation strata dummy variables (geographic region*market size*profit range) (following Bruhn and McKenzie 2009); δ_s are survey round dummies; and $\varepsilon_{i,j,t}$ is the error term, which is clustered at the market level to account for the market-level random assignment. β_1 and β_2 provides the

[12] <http://www.socialscienceregistry.org/trials/287>.

intent-to-treat effects at one- and three-year horizons, which is the effect of being assigned to treatment relative to being a firm in the control markets. Note that the 3-year impact represents the impact of having been assigned to receive business training 3 years earlier, followed by having a 50 per cent chance of being assigned to receive mentoring approximately 9 months earlier. The test β_1=β_2 determined whether the impacts differ over time, and β_1=β_2=0 determined whether the hypothesis that there is no treatment effect in any period can be rejected.

One of the main questions of interest in this study is whether there are spillovers from treated individuals to other individuals within the same villages. These spillovers could be positive (e.g. treated women share knowledge with control women in the same markets) or negative (e.g. treated women compete away the sales of control women from the same markets). β_3 and β_4 measure these spillover effects at the one- and three-year horizons by comparing control firms in treated markets to control firms in control markets.

To test whether assignment to the mentoring treatment had a differential impact from training alone, the analysis relied on the following regression, estimated using only rounds 4 and 5:

$$Y_{i,j,t} = \gamma_0 + \gamma_1 Mentor_{i,j} + \gamma_2 Trainonly_{i,j} + \gamma_3 Spillovermentor_{i,j} + \gamma_4 SpilloverTrainonly_{i,j} + \pi Y_{i,j,t=1} + \gamma BM_{ij,t=1} + X'_{k,i,j}\theta + \delta_4 1(t=4) + \varepsilon_{i,j,t} \quad (3)$$

where $Mentor$ indicates being in a market assigned to mentorship and being assigned a mentor, $Trainonly$ indicates being in a market assigned to training only and being assigned training, and $Spillovermentor$ and $SpilloverTrainonly$ are the spillover groups in the mentor and training-only markets. The question is only estimated for markets eligible for the mentoring intervention (those with more than 15 firms at baseline). The standard errors are again clustered at the market level. The test $\gamma_1 = \gamma_2$ was conducted to determine whether mentoring has a different effect from training alone.

The main analysis focuses on the ITT. Table 7 reports the LATE of receiving training on the primary outcomes, by instrumenting training attendance with training assignment in equation (2). None of the control group attended training, so the LATE is the same as the average treatment effect on the treated. The LATE impacts are approximately 24 per cent higher than the ITTs.

Table 7: LATE impacts for primary outcomes

	Firm survival	Daily sales	Weekly sales	Main product sales	Weekly profits	Main product profits	Photo inventories	**Aggregate index**
Impact of receiving training, allowing for spillovers								
Received Training*1 year	0.007	202**	355	187	97	65	597	0.061*
	(0.012)	(85)	(318)	(370)	(86)	(113)	(421)	(0.032)
Received training* 3 years	0.038**	212**	1283***	562*	273**	192	1017	0.108***
	(0.016)	(104)	(428)	(329)	(108)	(119)	(1087)	(0.039)
Spillover group*1 year	0.002	32	-476*	157	-65	-16	336	-0.011
	(0.011)	(70)	(268)	(327)	(66)	(93)	(343)	(0.026)
Spillover group*3 years	0.013	1	25	181	-30	5	670	0.002
	(0.014)	(85)	(328)	(269)	(79)	(94)	(867)	(0.029)
Sample size	13508	12943	12909	12064	12881	11985	5598	12923

Notes: Robust standard errors in parentheses, clustered at the market level.
All regressions control for randomisation strata fixed effects, the baseline value of the outcome and survey round fixed effects.
Receipt of training instrumented with assignment to training.

7.2 Impacts

7.2.1 Impacts on primary outcomes at the individual level

The primary outcomes are whether the business training succeeds in helping the firm survive and any increase in firm sales, profits and size (as measured by the photo value of inventories). Several measures of these variables were considered, as specified in the PAP, along with an aggregate index, which is the average of standardised z-scores of these outcomes and provides an overall measure of whether the intervention has succeeded in increasing firm performance while also controlling for multiple hypothesis testing. Table 8 provides the results.

Table 8: Impacts on primary outcomes

	Firm survival	Daily sales	Weekly sales	Main product sales	Weekly profits	Main product profits	Photo inventories	**Aggregate index**
Panel A: Impact of assignment to training, allowing for spillovers								
Assigned to training*1 year	0.006	161**	284	150	78	52	482	0.049*
	(0.010)	(68)	(255)	(297)	(69)	(91)	(343)	(0.026)
Assigned to training* 3 years	0.030**	171**	1038***	454*	221**	155	818	0.088***
	(0.013)	(85)	(349)	(268)	(88)	(97)	(879)	(0.032)
Spillover group*1 year	0.002	32	-476*	157	-65	-16	337	-0.011
	(0.011)	(70)	(270)	(328)	(66)	(93)	(345)	(0.026)
Spillover group*3 years	0.013	1	27	182	-29	5	671	0.002
	(0.014)	(85)	(330)	(270)	(79)	(95)	(874)	(0.030)
Mean of pure control group	0.852	1173	5763	3368	1439	1137	8567	0.005
Sample size	13508	12943	12909	12064	12881	11985	5598	12,923
p-value: Training 1 year = 3 years	0.056	0.889	0.006	0.380	0.051	0.399	0.679	0.148
p-value: Spillover 1 year = 3 years	0.403	0.726	0.103	0.943	0.650	0.866	0.688	0.647
p-value: Training effect zero both years	0.068	0.047	0.008	0.239	0.043	0.269	0.344	0.024
Panel B: Impact of mentoring compared to training alone, allowing for spillovers								
Assigned to mentoring	0.011	172	1216**	586	279**	133	1909	0.098**
	(0.018)	(115)	(486)	(366)	(116)	(129)	(1339)	(0.041)
Assigned to training alone	0.052***	209*	954**	523	300***	236*	256	0.104**
	(0.016)	(110)	(427)	(357)	(114)	(132)	(941)	(0.042)
Spillover group to mentoring	0.024	-75	138	391	-16	27	1972	-0.001
	(0.018)	(112)	(451)	(340)	(99)	(120)	(1299)	(0.037)
Spillover group to training alone	0.003	122	-3	231	53	69	-192	0.034
	(0.020)	(105)	(401)	(380)	(97)	(129)	(1095)	(0.037)
Sample size	5822	5625	5606	5535	5591	5530	2655	5,608
p-value: Mentoring = Training alone	0.030	0.765	0.605	0.886	0.877	0.522	0.184	0.901

Notes: Robust standard errors in parentheses, clustered at the market level.

*, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

All regressions control for randomisation strata fixed effects, for the baseline value of the outcome and for survey round fixed effects.

Panel B uses only survey rounds 4 and 5, since the mentoring intervention was carried out after round 3.

Aggregate index is the average of standardised z-scores of the other variables. See data appendix for variable definitions.

Panel A shows the one- and three-year impacts of assignment to training, while panel B separates the three-year results by whether the firm is in a market also assigned to mentoring or one assigned to training only. Over the one-year horizon the direct impacts are all positive, but only the impact on daily sales is statistically significant at the 5 per cent level, and there is a 0.05 standard deviation increase in the aggregate index, significant at the 10 per cent level. These estimates are all larger in magnitude by year three, with statistically significant impacts on survival (3 percentage points), daily sales (171 KSH, or 14.6% of the control mean), weekly sales (1,038 KSH, or 18% of the control mean), weekly sales of the main product the business sells (454 KSH, or 13.4% of the control mean), and weekly profits (221 KSH, or 15.4% of the control mean). The overall aggregate index shows a 0.088 standard deviation increase, which is significant at the 1 per cent level. The year three results are statistically different from year one results for key outcomes of survival, weekly sales and weekly profits, and the null hypothesis of no treatment effect in either period can be rejected.

The impact on the objective and independent measure of business size; i.e. the value of inventories as photographed in the firm, is positive (818 KSH, or 9.5% of the control mean), but not statistically significant. We carry out three robustness checks on our primary outcomes. First, we use the objective visual photographic evidence as to whether the business looks bigger than it was at baseline (Table 9). Second, we report results conditional on survival (Table 10). Third, we examine transformations of profits and sales, which are less susceptible to outliers (Table 11).

Table 9: Robustness checks: Is the business visually larger after three years?

	All firms	Survivors	All firms	Survivors
Assigned to training	0.042**	0.037		
	(0.021)	(0.023)		
Spillover group	0.019	0.019		
	(0.022)	(0.023)		
Assigned to mentoring			0.059**	0.058*
			(0.028)	(0.031)
Spillover group for mentoring			0.022	0.017
			(0.030)	(0.032)
Assigned to training alone			0.034	0.022
			(0.028)	(0.029)
Spillover group for training			0.028	0.023
			(0.026)	(0.027)
Sample size	2864	2571	2527	2265
p-value: Mentoring = Training alone			0.440	0.286

Notes: Robust standard errors in parentheses, clustered at the market level.
*, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Outcome is whether photographs of business inventories in the round 4 survey shows the business to be larger than in the baseline photograph.
Columns 2 and 4 are conditional on the firm surviving; columns 1 and 3 code closed firms as not being bigger.
Columns 3 and 4 are for the ‘non-small’ markets for which mentoring was randomised.

The analysis and robustness checks based on the photographs compares the round 4 business to the baseline business and classifies it as bigger, smaller or the same size.[13] Treated businesses are 4.2 percentage points more likely to be classified as bigger, which is significant at the 5 per cent level (Table 9).

Table 10: Robustness checks: Impact on primary outcomes conditional on survival

	Daily sales	Weekly sales	Main product sales	Weekly profits	Main product profits	Photo inventories	**Aggregate Index**
Panel A: Impact of assignment to training, allowing for spillovers							
Assigned to training*1 year	165**	294	126	69	46	604	0.051*
	(77)	(287)	(329)	(75)	(102)	(392)	(0.029)
Assigned to training* 3 years	145	967**	380	182*	124	487	0.080**
	(94)	(382)	(294)	(98)	(107)	(1017)	(0.035)
Spillover group*1 year	30	-557*	168	-79	-23	357	-0.012
	(78)	(290)	(368)	(70)	(105)	(404)	(0.028)
Spillover group*3 years	-24	-46	169	-61	-17	607	-0.001
	(95)	(354)	(303)	(86)	(108)	(982)	(0.033)
Mean of pure control group	1386	6818	3993	1702	1348	10694	0.107
Sample size	11339	11305	10460	11277	10381	4775	11319
p-value: Training 1 year = 3 years	0.804	0.027	0.509	0.177	0.567	0.902	0.324
p-value: Spillover 1 year = 3 years	0.574	0.138	0.998	0.840	0.964	0.797	0.733
p-value: Training effect zero both years	0.097	0.034	0.435	0.180	0.494	0.307	0.068
Panel B: Impact of mentoring compared to training alone, allowing for spillovers							
Assigned to mentoring	171	1289**	628	295**	132	2081	0.107**
	(132)	(529)	(395)	(132)	(142)	(1560)	(0.045)
Assigned to training alone	165	792*	360	233*	183	-658	0.087*
	(118)	(456)	(383)	(119)	(140)	(1093)	(0.044)
Spillover group to mentoring	-124	-28	355	-68	1	2020	-0.015
	(120)	(467)	(367)	(102)	(133)	(1414)	(0.039)
Spillover group to training alone	139	26	243	62	66	-460	0.046

[13] The analysis uses the objective visual photographic evidence as to whether the business looks larger than it was at baseline.

	Daily sales	Weekly sales	Main product sales	Weekly profits	Main product profits	Photo inventories	**Aggregate Index**
	(118)	(438)	(441)	(110)	(150)	(1220)	(0.042)
Sample size	4862	4843	4772	4828	4767	2193	4845
p-value: Mentoring = Training alone	0.967	0.354	0.559	0.674	0.761	0.058	0.693

Notes: Robust standard errors in parentheses, clustered at the market level.
*, **, and *** indicate significance at the 10, 5, and 1 percent levels respectively.
All regressions control for randomisation strata fixed effects for the baseline value of the outcome and for survey round fixed effects. Panel B uses only survey rounds 4 and 5, since the mentoring intervention was carried out between round 3 and round 4. Aggregate index is the average of standardised z-scores of the other variables. See data appendix for description of the different outcome variables.

Our main analysis codes profits and sales as zero for firms that have closed down (since they are not earning profits or sales). Table 10 shows that firms assigned to training have significantly higher weekly profits and weekly sales, and a higher aggregate index even when conditioned on survival. Table 11 also shows the results remain significant after using an alternative definition of firm survival, and when using alternative transformations of profits and sales, such as the inverse hyperbolic sine and logarithmic transformations. Finally, it also addresses the potential concern that business training affects how profits and sales are reporting by showing no treatment effect on the number of reporting errors, nor on the difference between the reported inventories and the size based on valuing photographs of inventories.

Table 11: Robustness of primary outcome results

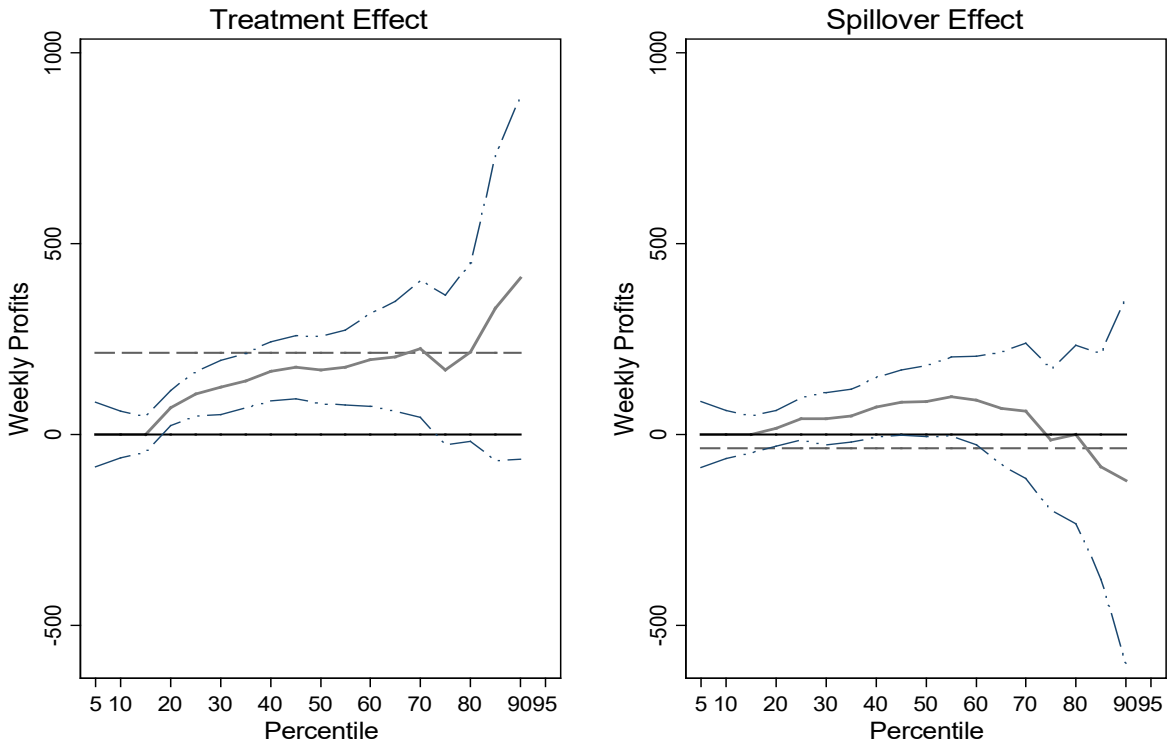
	Alternative survival	Inverse Hyperbolic sales	profits	Log sales	Log profits	Number errors	Inventory reporting
Assigned to training*1 year	0.006	0.124	0.098	0.094**	0.068	-0.014	0.060
	(0.010)	(0.116)	(0.101)	(0.037)	(0.047)	(0.026)	(0.063)
Assigned to training*3 years	0.031**	0.569***	0.428***	0.172***	0.138***	0.006	0.078
	(0.014)	(0.136)	(0.119)	(0.048)	(0.051)	(0.030)	(0.063)
Spillover group*1 year	0.001	-0.015	-0.008	-0.027	0.010	0.004	-0.053
	(0.011)	(0.120)	(0.103)	(0.040)	(0.045)	(0.029)	(0.066)
Spillover group*3 years	0.013	0.237	0.182	0.009	0.043	-0.014	0.028
	(0.014)	(0.151)	(0.130)	(0.047)	(0.047)	(0.029)	(0.074)
Mean of pure control group	0.835	7.270	6.240	8.265	6.889	0.542	0.849
Sample size	13666	12909	12881	10420	10790	11541	4391
p-value: Training 1 year = 3 years	0.046	0.001	0.003	0.036	0.112	0.607	0.836
p-value: Spillover 1 year = 3 years	0.374	0.069	0.110	0.385	0.463	0.638	0.375
p-value: Training effect zero both years	0.074	0.000	0.001	0.002	0.027	0.847	0.344

Notes: Robust standard errors in parentheses, clustered at the market level.
*, **, and *** indicate significance at the 10, 5, and 1 percent levels respectively.
All regressions control for randomisation strata fixed effects, the baseline value of the outcome where available, and for survey round fixed effects. See test for variable definitions

Figure 2 and Figure 3 show the quantile treatment effects on profits and sales for the three-year time horizon. The quantile effects increase by quantile, showing that the impacts were larger at the top of the distribution than the bottom, but are statistically significant from the 20th percentile onwards until at least the 80th percentile.

Consider next the evidence for spillover effects. Over a one-year horizon, there are negative spillovers on weekly profits and weekly sales, with this effect significant at the 10 per cent level for sales. However, there are also positive and insignificant spillover estimates for several of the other profits and sales measures and, as a result, the impact on the aggregate index is small, with the point estimate being a 0.01 standard deviation reduction, which is not statistically significant. Over the three-year horizon, none of the outcomes show a significant spillover impact, and the aggregate index has a 0.002 standard deviation increase, which is small and not statistically significant. Likewise, the quantile treatment effects in Figure 2 and Figure 3 show the spillover impacts to be much smaller than the direct effects, and not statistically significant. As a result, it is not possible to reject the prospect that there are no spillover effects (either positive or negative) on untreated firms operating in the same markets as treated firms.

Figure 2: Quantile treatment effects on profits after three years

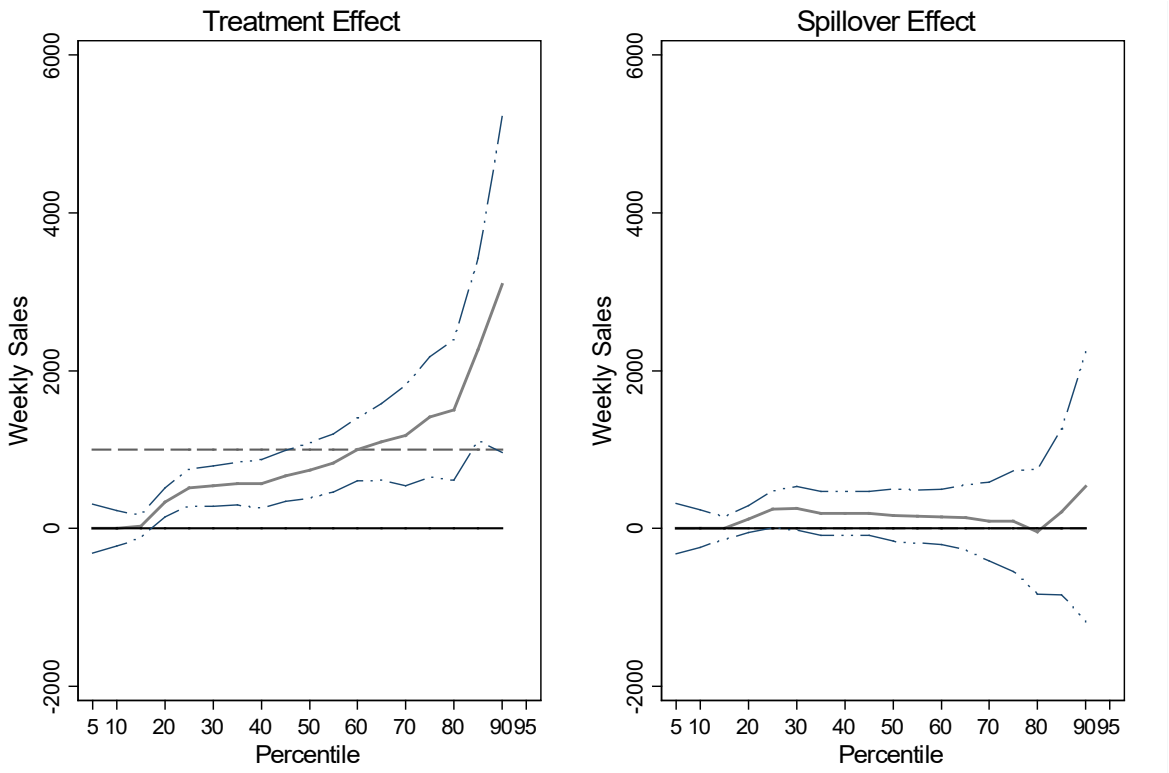


Notes: ITT effects from quantile regressions controlling for randomisation strata and baseline profits, combining round 4 and round 5 survey data. Straight dashed line shows OLS estimate; 95 per cent confidence intervals displayed allowing for clustering at market level.

This picture of either small or no spillovers in sales and profits is consistent with the qualitative work, with some respondents saying they saw no change in market competition as a result of some women being trained, while others noted a mix of positive (passing on knowledge to other firms) and negative (treated women providing better customer service to attract customers) spillovers.

Panel B (Table 8) examines whether the impacts differ for those assigned to mentoring versus training alone. The impacts are similar in magnitude for most outcomes, with survival being the one exception, where training alone appears to have a larger impact than training followed by mentoring. When the overall index is considered, which accounts for multiple testing, there is a 0.098 standard deviation increase for mentoring and 0.104 standard deviation increase for training alone. These impacts are very similar in magnitude and the equality hypothesis cannot be rejected ($p = 0.901$).

Figure 3: Quantile treatment effects on sales after three years



Notes: ITT effects from quantile regressions controlling for randomisation strata and baseline sales, combining round 4 and round 5 survey data. Straight dashed line shows OLS estimate; 95 per cent confidence intervals displayed allowing for clustering at market level.

7.2.2 Impacts on secondary outcomes at the individual level

In order to examine the extent to which these improvements in profits and sales translate into broader individual wellbeing, the study's PAP specified four domains of secondary outcomes to examine (see section 2.3 Theory of change and research hypotheses). The impacts on these outcomes are shown in Table 12.

Table 12: Impacts on secondary outcomes

	Employed for pay	All work income	Empowerment index	Life ladder today	Life ladder 5 years	Mental health	Household durables index
Panel A: Impact of assignment to training, allowing for spillovers							
Assigned to training*1 year	0.009	177*	0.109	0.254***	0.253***	0.265*	0.108*
	(0.008)	(96)	(0.110)	(0.050)	(0.064)	(0.151)	(0.058)
Assigned to training*3 years	0.020*	389***	-0.037	0.221***	0.322***	0.356**	0.091
	(0.011)	(141)	(0.094)	(0.049)	(0.054)	(0.162)	(0.065)
Spillover group*1 year	0.003	-112	0.013	-0.011	0.000	0.115	0.048
	(0.009)	(93)	(0.113)	(0.056)	(0.068)	(0.143)	(0.059)
Spillover group*3 years	0.011	24	0.070	0.062	0.029	0.272	0.073
	(0.012)	(133)	(0.105)	(0.046)	(0.058)	(0.174)	(0.070)
Mean of pure control group	0.885	2.144	7.034	5.015	7.982	17.707	-0.063
Sample size	13508	12881	5873	12609	12608	5873	5823
p-value: Training 1 year = 3 years	0.323	0.114	0.289	0.581	0.420	0.659	0.813
p-value: Spillover 1 year = 3 years	0.533	0.283	0.689	0.297	0.740	0.460	0.721
p-value: Training effect zero both years	0.169	0.019	0.543	0.000	0.000	0.034	0.137
Panel B: Impact of mentoring compared to training alone, allowing for spillovers							
Assigned to mentoring	0.005	488**	-0.062	0.283***	0.417***	0.477**	0.226*
	(0.015)	(187)	(0.133)	(0.063)	(0.057)	(0.206)	(0.121)
Assigned to training alone	0.035**	477***	0.133	0.204***	0.271***	0.169	0.132
	(0.014)	(181)	(0.115)	(0.056)	(0.064)	(0.217)	(0.106)
Spillover group to mentoring	0.017	30	-0.019	0.192***	0.093	0.236	0.306***
	(0.015)	(180)	(0.142)	(0.052)	(0.067)	(0.232)	(0.117)
Spillover group to training alone	0.005	178	0.293**	0.042	0.064	0.049	-0.000
	(0.016)	(153)	(0.128)	(0.052)	(0.062)	(0.221)	(0.128)
Sample size	5822	5591	2470	5511	5511	2470	2465
p-value: Mentoring = Training alone	0.062	0.958	0.174	0.270	0.037	0.237	0.450

Notes: Robust standard errors in parentheses, clustered at the market level.
*, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.
All regressions control for randomisation strata fixed effects, for the baseline value of the outcome where available and for survey fixed effects. Panel B uses round 4 and 5 data only.

The first set of secondary outcomes concerns employment, whether self-employment or wage work. Column 1 of Table 12 shows that 88.5 per cent of the pure control group are engaged in some form of employment for pay at the time of the three-year follow-up surveys, and an increase in firm survival also translates into an increase in employment. Column 2 looks at all income from work, which combines profits from self-employment with any earnings from wage labour. Higher profits are not offset by lower wage earnings, as total income from work increases by 389 KSH per week, or 18.1 per cent of the control mean.

The second set of outcomes concerns topics and approaches emphasised in the training that aimed to empower women in decision-making around finances and business. The study measures 10 different outcomes in this domain (e.g. are they compelled to spend money on their husband or family, do they need someone's permission to travel to sell a business asset, do they have money they have sole control over, etc.). The average individual in the control group is able to do seven out of these 10, and column 3 shows that training is not found to have any sizeable or significant impact on this measure of empowerment at either time horizon. This is consistent with the qualitative assessment, which noted that training did not appear to change individual or household decision-making dynamics (ICRW 2015).

The third domain examined is subjective wellbeing and mental health. Subjective wellbeing today and the anticipated subjective standard of living in five years' time is measured on a Cantril ladder, with mental health using the MHI-5 index of Veit and Ware (higher scores indicate better mental health). Respondents showed a great deal of optimism about the future, seeing themselves on step 5 out of 10 on the life ladder currently, but expecting to be on step 8 in five years' time. Training increases both current and future subjective wellbeing by 0.2-0.3 steps. The impact on mental health is positive and also statistically significant in both time periods. Finally, the impact on ownership of household durable assets is measured. There is a positive, but marginally significant impact in year one, and a positive but insignificant impact in year three. The only evidence of a significant difference between groups receiving mentoring or training alone is observable in the 'any employment' outcome, which is consistent with the differential survival effect. Taken together, these results show that higher profits and sales do appear to have translated into higher overall wellbeing for women provided with training.

7.2.3 Market-level impacts

Our results show that firms assigned to treatment have grown, with no negative spillover for the untrained in the same market. This suggests that the overall markets have grown. To formally test this hypothesis, the outcomes are aggregated to the market level, and estimated through the following equation at the level of the 157 markets:

$$TotalY_{j,t} = \beta_0 + \beta_1 MarketTrained_j * 1Year_t + \beta_2 MarketTrained * 3Years_t + \\ + \pi TotalY_{j,t=1} + X_j'\theta + \sum_{s=2}^{5} \delta_s 1(s = t) + \varepsilon_{j,t} \quad (4)$$

Where $MarketTrained$ indicates market j was assigned to training, $TotalY$ is the total of outcome Y (profits, sales, or customers) in market j at time t, and the standard errors are clustered over time at the market level.

Table 13 shows the results of estimating equation (4). By year three, overall profits have grown by 2,554 KSH per week (9.2% of the control mean, not statistically significant), overall sales in the market by 15,353 KSH per week (13.7% of the control mean, significant at the 5% level), and the total number of customers in the market per week by 318 (16.7% of the control mean). These results are consistent with the overall market growing, at least over the longer term.

A market census was conducted in rounds 2 and 4 to determine whether there had been new entrants to the market since the time of training, and whether these new entrants, in addition to existing firms that did not pass the eligibility filters, had cut back on profits due to some firms being trained. Table 13 shows positive but not statistically insignificant impacts on the numbers of new entrants in these markets, the total number of firms not in the survey and the weekly profits of these firms. The success of those trained, therefore, also does not come from crowding out new entrants.

Table 13: Impacts on market level

	Total market profits	Total market sales	Total market customers	# New entrants	Total other firms	Other firm profits
Market assigned to training*1 year	792	2672	131	0.193	1.473	4439
	(1378)	(5211)	(104)	(0.509)	(1.375)	(2940)
Market assigned to training*3 years	2554	15353**	318**	0.833	1.290	2662
	(1568)	(6613)	(124)	(0.785)	(1.784)	(3685)
Mean of control markets	27651	110944	1916	8.59	22.48	34870
Sample size	628	628	628	301	301	301
p-value: 1 Year = 3 Years	0.201	0.021	0.038	0.517	0.933	0.684

Notes: Standard errors in parentheses, clustered at the market level. Unit of observation is market survey round.
*, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
Regressions include controls for randomisation strata, survey round and baseline value of outcome where available.

8. Discussion

8.1 Why do these impacts arise? Causal chain and mechanisms

The surveys and PAP enable tracing out of the causal chain from providing training through to changes in business outcomes, and examination of the different mechanisms through which training may or may not have an effect. Since different mechanisms are examined, the study provides sharpened q-values that hold constant the false discovery rate when reporting results for specific outcomes (Table 14). All outcomes at the three-year horizon that have p-values below 0.05 also have sharpened q-values below 0.05, whereas only the one-year impacts that are significant at the 1 per cent level have sharpened q-values below 0.10.

Table 14: Sharpened q-values for mechanism impacts

	1-year		3-years	
Measure	**p-value**	**Sharpened q**	**p-value**	**Sharpened q**
Business knowledge	0.839	1.000	0.175	0.096
Business practices	0.000	0.001	0.000	0.001
Worked with mentor			0.000	0.001
Weekly customers	0.648	1.000	0.001	0.004
Gained new customer	0.082	0.387	0.040	0.039
Lost new customer	0.025	0.127	0.006	0.011
Sales per customer	0.396	0.847	0.138	0.087
Open set time			0.010	0.015
Introduce new product	0.000	0.001	0.000	0.001
Profit ratio	0.131	0.537	0.530	0.249
Monitors sales trends	0.000	0.001	0.000	0.001
Fraction stock spoiled	0.620	1.000	0.521	0.249
Received bulk discount	0.700	1.000	0.141	0.087
Production cost change	0.320	0.847	0.466	0.249
Received loan	0.323	0.847	0.506	0.249
Business bank account			0.010	0.015
Inventory value	0.246	0.847	0.050	0.046
Capital stock	0.953	1.000	0.003	0.007
Own labour	0.957	1.000	0.000	0.001
Self-efficacy	0.838	1.000	0.293	0.152
Get Ahead attitudes	0.396	0.847	0.032	0.036
Women's association	0.870	1.000	0.002	0.006
Number discuss business	0.388	0.847	0.037	0.039
Works with others	0.154	0.544	0.195	0.102

Note: To control for multiple hypothesis testing when examining multiple mechanisms, we construct sharpened q-values following Anderson (2008) and Benjamini et al. (2006). This process uses a two-stage procedure to control the false discovery rate when reporting results for specific outcomes. Table 14 reports the original p-values and corresponding sharpened q-values. We see that all three-year outcomes that have p-values below 0.05 also have sharpened q-values below this level. In contrast, over the one-year horizon, only the impacts on business practices, introducing new products and monitoring sales trends are significant after this adjustment.

8.1.1 Changes in business knowledge, practices and use of a mentor

The first step in the causal chain is for training to lead to changes in business knowledge and business practices among women receiving training. Business knowledge was assessed by giving respondents a description of a business and asking them seven questions that involve calculating revenue, value of stock on hand, variable costs, total expenses, profits, fixed costs and break-even point. This proved very difficult for most participants, with the median respondent only answering two out of seven questions correctly and only 0.5 per cent answering all correctly. These questions were only asked in the first long follow-up survey and was asked of both those with surviving businesses as well as those whose business had closed down.

Column 1 of Table 15 shows that there is no significant treatment effect or spillover effect on business knowledge. This is consistent with the financial literacy results of Carpena et al. (2011), who find that financial literacy training does not improve performance on questions involving numerical calculations.

Business practices are measured through a set of 26 questions that capture information about marketing, record-keeping, buying and stock control, and financial planning of the firm. These questions are only measured in the long follow-up survey rounds for firms that survive. These questions have been shown to correlate strongly with business performance in a range of countries by McKenzie and Woodruff (2015) and to predict future survival and growth of the firm. The mean firm in the pure control group employs 53 per cent of these practices.

Column 2 shows that the impact of being invited to training is a statistically significant increase in the use of business practices at both the one-year and three-year horizons: treated firms have a 0.05-0.07 increase in the proportion of practices used. This is approximately a 10-13% increase in the control mean. There is a marginally significant positive spillover in business practices to untreated firms in the same marketplaces over one year; however, the magnitude is very small (0.01) and the three-year impact is smaller still and not significant. The increase in practices is three times as large for those assigned to mentoring (0.096) as those assigned to training alone (0.033), suggesting that mentoring helped to reinforce business practices.

Table 15: Impact on business knowledge, business practices and use of a mentor

	Business knowledge	Business practices	Worked with mentor
Panel A: Impact of assignment to training, allowing for spillovers			
Assigned to training*1 year	-0.018	0.049***	
	(0.089)	(0.008)	
Assigned to training* 3 years		0.068***	0.335***
		(0.009)	(0.037)
Spillover group*1 year	0.124	0.014*	
	(0.091)	(0.008)	
Spillover group*3 years		0.006	0.018
		(0.009)	(0.020)
Mean of pure control group	2.000	0.533	0.106
Sample size	3059	5404	2520
p-value: Training 1 year = 3 years		0.067	
p-value: Spillover 1 year = 3 years		0.389	
p-value: Training effect zero both years		0.000	
Panel B: Impact of mentoring compared to training alone, allowing for spillovers			
Assigned to mentoring		0.096***	0.666***
		(0.012)	(0.028)
Assigned to training alone		0.033***	0.108***
		(0.010)	(0.020)
Spillover group to mentoring		-0.006	0.009
		(0.011)	(0.023)
Spillover group to training alone		0.004	0.047*
		(0.012)	(0.026)
Sample size		2,236	2,212
p-value: Mentoring = Training alone		0.000	0.000

Notes: Robust standard errors in parentheses, clustered at the market level.
*, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.
All regressions control for randomisation strata fixed effects.
Business knowledge only collected in round 2, mentoring only collected in round 4.
Business practices collected in rounds 2 and 4, with regression controlling for survey round and for baseline business practices.
Outcomes are conditional on operating a firm.

Finally, column 3 confirms that those assigned to the mentoring treatment are indeed much more likely to say they have used a mentor. However, there is also a significant impact on using a mentor from those firms assigned only to training, suggesting that training did lead to a doubling in the use of mentor, even without the separate mentorship intervention.

8.1.2 Impact on dealing with customers

Increased marketing and better presentation of the business may enable a firm to increase sales as a result of gaining new customers and being better able to retain existing customers. This is examined in the first three columns of Table 16.

Table 16: Impact on dealing with customers

	Weekly customers	Gained new customer	Lost regular customer	Sales per customer	Open business at set time	Introduced new product
Panel A: Impact of assignment to training, allowing for spillovers						
Assigned to training*1 year	1.7	0.023*	-0.038**	25.5		0.116***
	(3.8)	(0.013)	(0.017)	(30.0)		(0.022)
Assigned to training*3 years	14.5***	0.034**	-0.041***	38.7	0.062***	0.089***
	(4.4)	(0.016)	(0.015)	(25.9)	(0.024)	(0.021)
Spillover group*1 year	-2.2	0.009	-0.020	0.7		0.007
	(3.9)	(0.014)	(0.018)	(20.6)		(0.021)
Spillover group*3 years	-0.7	0.008	-0.009	2.2	0.030	0.008
	(4.6)	(0.018)	(0.016)	(17.1)	(0.023)	(0.021)
Mean of pure control group	100	0.767	0.804	134	0.342	0.269
Sample size	12867	12977	12978	11186	2398	6216
p-value: Training 1 year = 3 years	0.004	0.586	0.878	0.734		0.327
p-value: Spillover 1 year = 3 years	0.753	0.964	0.666	0.951		0.970
p-value: Training effect zero both years	0.003	0.049	0.003	0.247		0.000
Panel B: Impact of mentoring compared to training alone, allowing for spillovers						
Assigned to mentoring	12.1*	0.004	-0.048***	50.3	0.103***	0.102***
	(6.1)	(0.021)	(0.017)	(38.0)	(0.033)	(0.028)
Assigned to training alone	12.0**	0.054***	-0.054***	24.6	0.030	0.084***
	(5.2)	(0.020)	(0.017)	(24.6)	(0.031)	(0.028)
Spillover group to mentoring	0.8	0.000	-0.035*	-5.3	0.026	0.021
	(6.4)	(0.022)	(0.020)	(19.1)	(0.030)	(0.026)
Spillover group to training alone	-4.3	0.009	-0.002	-0.0	0.039	0.022
	(6.0)	(0.023)	(0.019)	(20.5)	(0.032)	(0.029)
Sample size	5,602	5,636	5,636	4,801	2,090	2,636
p-value: Mentoring = Training alone	0.992	0.020	0.758	0.558	0.071	0.597

Notes: Robust standard errors in parentheses, clustered at the market level.

*, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

All regressions control for randomisation strata fixed effects.

Opening at set time only asked in round 4 survey and is conditional on business operating.

The women in the study sample serve a median of 60 customers and mean of 100 customers per week in the control markets (after top-coding at the 99th percentile to reduce the influence of outliers). Column 1 shows that those assigned to training are serving an additional 14 customers a week after three years, significant at the 1 per cent level. Columns 2 and 3 show that treated firms are more likely to have gained a new customer during the last three months, and less likely to have lost a regular customer during the same time period. There is no significant spillover impact, which is consistent with previous evidence that the total number of customers in the market increased.

Total sales of the business increased, as did the number of customers. Column 4 considers the sales per customer. The point estimates are positive, suggesting more revenue per customer, but is not statistically significant. If outcomes are instead considered in log form, the identity is:

$$\log(sales) = log(sales/customers) + log(customers) \quad (5)$$

Then the treatment effects are 17.6% on sales (p = 0.006), 9% on sales per customer (p = 0.107) and 8.5% on the number of customers, suggesting that the increase in sales volume is coming approximately evenly from expanding the number of customers as well as from obtaining more revenue per customer.

Why might business owners gain more customers? The qualitative work in part points to a better customer experience, with cleaner shops that are more likely to open on time, and owners being more pleasant to customers. For example, the qualitative work noted one woman saying:

> *There are those who didn't know how to attract customers, but now I can see they have been able to attract customers…They are talking to them nicely unlike before where they would talk rudely', and another noting that now, 'I try to talk [with] customers in a nice and professional way. I try to make delicious meals and make my place of work look attractive to them…I make sure [my cooking] it's up to the customers' standard, not watery food* (ICRW2015)

In qualitative interviews with those who had gone through mentoring, one participant said, '*I used to not care much about my business, I could open late, and sometimes the shop would remain closed if I don't feel like opening. Now I know that it is important to put effort into my business*.'

This potential channel is incorporated into the three-year follow-up survey, which asks whether the business always opened at a set time each day. Column 5 shows that this only occurs among 34 per cent of the control group and training increases this by a significant 6 percentage points. Mentoring appears to have a greater impact than training alone on having regular opening hours.

The second way in which firms were able to attract more customers was diversifying the range of products they sold. Banerjee and Duflo (2008) note the preponderance of small businesses in developing countries all selling similar things, with no reason for customers to seek them out in terms of either product line or shopping environment. Column 6 shows that firms assigned to training were 9-11 percentage points more likely to have introduced a new product to their business, significant at the 1 per cent level.

There was tremendous variety in the new products introduced, with almost every business that introduced a new product providing a different answer as to what this product was. Examples included avocados, oranges, carrots, garlic, fertiliser, body oil, cowpeas, belts, hair oil, tea, tobacco and Weetabix. Since most firms were selling only a handful of products, this diversification could have attracted customers that would not have otherwise purchased from them.

8.1.3 Impact on financial control, access to finance and capital levels

The analysis continues by examining the extent to which training enables firm owners to produce more, or earn higher profits, by managing resources more effectively and overcoming financial constraints. Several studies have emphasised the possibility that business training may have its strongest impact on sales during a bad month by helping participants identify strategies to reduce downward fluctuations in sales through diversifying the products they offer, as well as being more proactive about alternative activities during slow months. McKenzie and Woodruff (2013) note, however, that evidence for this has been mixed in existing studies. This channel is examined in columns 1 and 2 of Table 17, showing that although businesses are more likely to regularly use business records to determine whether sales of a particular product are increasing or decreasing, this does not translate into an impact on reducing fluctuations in profits in bad months relative to usual months.

Columns 3, 4, and 5 of Table 17 examine the firms' inventory management and purchasing outcomes. There are no significant impacts of treatment on the fraction of stock lost to spoilage, the receipt of bulk discounts in purchasing or in the cost of producing the firm's most profitable item. The last is measured with considerable noise, so the insignificant point estimate admits a wide confidence interval.

Better business practices may enable firms to obtain more financing. This may work through either the demand or supply side for formal finance. On the demand side, individuals who have received training may have more plans for expansion or feel more confident approaching banks. On the supply side, banks may be more willing to lend to firms that keep better records or have higher profitability and sales. However, column 6 of Table 17 shows that access to finance is limited in the sample, with only 17 per cent of the pure control group receiving a loan from a bank or microfinance organisation in the last year, and that treatment does not have a significant impact on this. This is consistent with the qualitative work, which found a number of challenges on both the demand and supply side for formal finance: women were often afraid of what might happen if they fell behind on payments, found the loan application process cumbersome and time-consuming, and faced challenges in terms of ability to provide documentation and/or collateral in some cases (ICRW 2015).

Table 17: Impacts on financial control, access to finance and capital levels

	Ratio of worst/current profit	Monitors sales trends	Fraction stock lost to spoilage	Received a bulk discount	Production cost change (%)	Received loan from bank/MFI	Has business bank account	Inventory value	Capital stock value
Panel A: Impact of Assignment to Training, Allowing for Spillovers									
Assigned to training*1 year	0.014	0.112***	-0.007	0.009	15.536	0.017		2475	-52
	(0.009)	(0.023)	(0.014)	(0.023)	(15.561)	(0.017)		(2125)	(882)
Assigned to training* 3 years	-0.007	0.141***	-0.014	0.034	13.954	0.013	0.058**	5889**	3219***
	(0.011)	(0.022)	(0.021)	(0.023)	(19.084)	(0.019)	(0.022)	(2978)	(1085)
Spillover group*1 year	-0.004	0.025	0.037**	-0.028	7.825	0.014		-339	-828
	(0.009)	(0.024)	(0.017)	(0.025)	(17.121)	(0.018)		(1966)	(917)
Spillover group*3 years	-0.011	0.000	-0.005	0.013	-5.558	-0.017	-0.003	3350	672
	(0.012)	(0.023)	(0.024)	(0.024)	(20.019)	(0.021)	(0.023)	(2829)	(1065)
Mean of pure control group	0.450	0.269	0.194	0.423	160.4	0.172	0.354	19918	13410
Sample size	10609	6214	4989	5879	5541	5404	3005	6183	6203
p-value: Training 1 year = 3 years	0.136	0.263	0.766	0.447	0.935	0.850		0.107	0.002
p-value: Spillover 1 year = 3 years	0.622	0.355	0.102	0.255	0.526	0.139		0.093	0.174
p-value: Training effect zero both years	0.252	0.000	0.767	0.308	0.569	0.583		0.137	0.005
Panel B: Impact of mentoring compared to training alone, allowing for spillovers									
Assigned to mentoring	-0.024*	0.194***	-0.038	0.029	-8.074	0.004	0.048	15371***	3180**
	(0.012)	(0.029)	(0.026)	(0.030)	(23.297)	(0.026)	(0.034)	(4523)	(1504)
Assigned to training alone	-0.015	0.070***	-0.009	0.002	-15.671	0.044*	0.099***	4705*	2561*
	(0.013)	(0.024)	(0.028)	(0.030)	(25.575)	(0.026)	(0.025)	(2793)	(1330)
Spillover group to mentoring	-0.010	-0.012	0.002	-0.012	-3.198	-0.023	0.010	9707**	628
	(0.014)	(0.031)	(0.032)	(0.030)	(25.752)	(0.027)	(0.035)	(3770)	(1365)
Spillover group to training alone	-0.029**	-0.014	-0.029	0.010	-41.846	0.006	0.023	2649	-133
	(0.014)	(0.032)	(0.034)	(0.032)	(26.112)	(0.028)	(0.025)	(3198)	(1451)
Sample size	4563	2636	2078	2497	2389	2236	2636	2634	2635
p-value: Mentoring = Training alone	0.515	0.000	0.353	0.432	0.791	0.157	0.137	0.021	0.720

Notes: Robust standard errors in parentheses, clustered at the market level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
All regressions control for randomisation strata fixed effects. Mentoring regressions only for three-year follow-up.
Business bank account not asked in round 2 follow-up survey.

Firms are 5.8 percentage points more likely to be using business bank accounts (column 7). Columns 8 and 9 then look at whether firms have grown in size according to their inventory levels and capital stock. By year three, both the inventory levels and capital stocks are higher in the treatment group, which is consistent with them having more to sell and with building up the size of the firm. The inventory impact is larger and statistically significant (at 1% level) after mentoring than after training alone. Since firms did not receive more finance, this build-up in inventories and capital is likely to reflect re-investment of profits. This may also explain why impacts are stronger after three years than one year, since it takes time for firm owners to slowly reinvest profits into more merchandise to sell.

8.1.4 Impacts on owner hours, attitudes, and social capital

The final set of intermediate outcomes and mechanisms are examined in Table 18. A first analysis looks at whether women changed the amount of time they are devoting to their business. The qualitative work suggested this might be the case, as evidenced by this quote from one participant:

> *Then I used to open [my business] any time I wished…I would open much later…but these days it's better since I constantly open at nine and close at night at around eight. Those days I just used to do a little work, I could not stay for long. Whenever I got some money to pay for my merry-go-round, then I would just close my business for the day. Also, then if I got someone who was buying five bags of maize, then I could just close business and leave for home, but these days I stay until I am convinced that it's time to leave* (ICRW 2015)

This qualitative finding does not appear in the one-year horizon; however, it is there over three years for both mentoring and training alone, with women working four hours more per week.

Table 18: Attitudes and associations

	Own labour hours	Entrepreneurial self-efficacy	GET Ahead attitudes	Women's association	# Discuss business	Works together with other women
Panel A: Impact of assignment to training, allowing for spillovers						
Assigned to training*1 year	0.059	-0.037	-0.018	0.003	0.181	0.035
	(1.106)	(0.181)	(0.021)	(0.017)	(0.210)	(0.024)
Assigned to training* 3 years	4.596***			0.057***	0.584**	0.033
	(1.274)			(0.018)	(0.277)	(0.026)
Spillover group*1 year	-2.060*	-0.180	-0.047**	0.026	0.032	-0.000
	(1.138)	(0.171)	(0.022)	(0.019)	(0.224)	(0.025)
Spillover group*3 years	-0.192			-0.013	0.066	0.010
	(1.356)			(0.017)	(0.308)	(0.024)
Mean of pure control group	42.6	4.457	3.802	0.115	4.672	0.490
Sample size	6215	3059	3059	5400	5398	5403
p-value: Training 1 year = 3 years	0.001			0.042	0.259	0.959
p-value: Spillover 1 year = 3 years	0.189			0.155	0.934	0.731
p-value: Training effect zero both years	0.001			0.005	0.072	0.242
Panel B: Impact of mentoring compared to training alone, allowing for spillovers						
Assigned to mentoring	4.777***			0.055**	0.920**	0.027
	(1.642)			(0.023)	(0.362)	(0.039)
Assigned to training alone	4.300**			0.062***	0.303	0.030
	(1.677)			(0.023)	(0.329)	(0.032)
Spillover group to mentoring	-0.429			-0.029	0.040	-0.016
	(1.758)			(0.020)	(0.417)	(0.032)
Spillover group to training alone	0.101			0.002	-0.003	0.041
	(1.659)			(0.025)	(0.390)	(0.031)
Sample size	2635			2234	2235	2236
p-value: Mentoring = Training alone	0.814			0.795	0.135	0.940

Notes: Robust standard errors in parentheses, clustered at the market level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
All regressions control for randomisation strata fixed effects. Mentoring regressions only for three-year follow-up.
Self-efficacy and Get Ahead attitudes only asked in round 2.

The training was also intended to increase women's confidence in their ability to perform business tasks. Entrepreneurial self-efficacy is measured through 10 questions that capture information on the owner's confidence in their ability to perform key business activities, such as formulating ideas for new products, selling a product to a customer they are meeting for the first time and persuading a bank to lend them money for their business. The mean control group individual is very confident in their ability to do 4.5 out of these 10 tasks, and Column 2 shows no significant effect of training on self-efficacy. In addition, 11 questions were intended to measure the types of attitudes that GET Ahead tries to encourage. These include willingness to take risks to get ahead in business, planning for the future, feeling confident that one can find solutions to problems that arise and keeping eyes open for ways to improve the business. These are scored on a five-point scale, where five indicates greatest agreement. The mean score across these questions is 3.8 for the pure control group. Column 3 shows no significant effect of treatment on these.

The training also emphasised and encouraged cooperation with other women in the marketplace. The remaining columns of Table 18 examine aspects of this. Columns 4, 5 and 6 consider membership of women's associations, discussing business with other women in the market, and working with other women to obtain bulk discounts or purchase goods together. There are no significant treatment effects on any of these measures over the one-year horizon, but treated individuals were found to be more likely to belong to associations and discuss business with others over the three-year horizon. Despite the trainers' hope that treated women would band together with untreated women to purchase inputs at discounts together, there is no direct effect nor spillover effect on this outcome.

8.2 Cost benefit

The analysis shows that, over a three-year horizon, the training has grown these underdeveloped markets, increasing the total volume of sales and number of customers. As a result, gains to the treated firms do not come at the expense of lost customers for untreated firms in the same markets. The cost of training was approximately US$200 per person assigned to training and the gain in weekly profits was 221 KSH (US$2.60). The gain in profits would therefore need to last at least 76 weeks, or approximately 1.5 years, for the benefits of the programme to exceed the costs. Although there is no long-term means of examining these benefits, they seem greater at three years than one year, thereby suggesting that this outcome appears plausible. This cost-benefit calculation is most suitable for deciding whether it would be individually profitable for a business owner to undertake training and to measure producer surplus given the lack of spillovers. Measurement of the social cost benefit would need to also take into account the change in consumer surplus arising from increased product variety and from any price changes (which we do not measure precisely).

The additional mentoring treatment cost approximately US$553 per individual assigned to mentoring. Since it is not possible to reject the notion that impacts on primary outcomes are the same for mentoring as for training alone, the evidence here suggests that adding mentoring does not pass a cost-benefit test.

9. Implications

Firms assigned to training are 3 percentage points more likely to survive after 3 years, earn 18 per cent higher sales and make 15 per cent higher profits. Their owners have better mental health and a higher subjective standard of living. These gains are greater at three years than at one year after the training and are similar for firms assigned to training only as for firms also assigned to a mentor. These gains come with no significant spillover effects on untreated firms operating in the same markets, and total sales and the total number of customers is higher in the treated markets than control markets. There is also no reduction in new entry into these markets after training. This market growth appears to stem from better customer service, better business practices and the introduction of new products, with no significant impacts on access to finance or input management. The conclusion is that, in underdeveloped markets, microenterprise growth need not come at the expense of competitors and business training can help the overall market grow.

Factoring in the cost of training, the evaluation shows that the gain in profits would need to last for at least 76 weeks, or approximately 1.5 years, for the benefits of the programme to exceed the costs. The data at three years suggest that this is plausible. Adding mentoring does not increase cost effectiveness.

9.1 The findings in today's research context

Better business practices are strongly associated with better firm performance across a range of countries in both the cross-section and over time (McKenzie and Woodruff 2015). However, in their review of the growing body of experimental literature assessing the impact of business training in developing countries, McKenzie and Woodruff (2013) note that the evidence on effectiveness is mixed, in part because many studies have low statistical power and measure impacts over short durations.

Three strands of this literature are of particular relevance to this study. The first is evidence that business training may be less effective for female business owners, either because they work in sectors with very low efficiency scales or because they face many other constraints that limit the ability of their businesses to grow (e.g. de Mel et al. 2014; Berge et al. 2014; Giné and Mansuri 2016). GET Ahead, the training programme examined in this study, was designed especially for women with low education levels. Bulte et al. (2016) evaluate the impact of this same programme on female microfinance clients in Vietnam and find some evidence of firm growth 12 months after training. The results show stronger and longer-term evidence that this training can help female-owned businesses to grow.

Second, several studies examine whether augmenting standard business training with mentoring can enhance its effectiveness. Valdivia (2015) and Giné and Mansuri (2016) find that individualised hand-holding or mentoring after training does not have sustained impacts relative to training alone. Brooks et al. (2016) find, also in Kenya, that assigning a mentor to a young firm does increase profits in the short-run, but the effect fades over time. The results are consistent with mentoring not delivering additional gains compared to training alone.

Third, almost all of the existing literature to date has been unable to measure impacts at the market level and therefore cannot determine whether any gains are experienced by untrained firms. The one exception is Calderón et al. (2013), who worked with 17 villages in rural Mexico, assigning 7 to treatment and 10 to control. They find no significant spillovers, although this may in part reflect low statistical power given the small number of villages and because they lose 18 per cent of their sample to attrition and 41 per cent to closure by their second follow-up. The paper builds on this work and not only shows no significant spillovers, but significant growth at the market level.

9.2 Implications for policymakers and practitioners

Many markets in developing countries are underdeveloped, with poorly managed firms selling a narrow variety of products. The results of this paper show that training can be one way to develop these markets.

Governments and practitioners investing in and developing measures to support women in business can rely on training interventions to improve the survival, profitability, and growth of female-owned businesses that receive training. The positive impacts and long-term horizon found in Kenya suggest there is scope to expand small businesses run by women. Furthermore, the lack of negative spillovers among competitor businesses implies that business training has the potential to create cost-effective social gains.

The evaluation also prompts development actors to improve programme design. With increased attention to enterprise development, implementers are called on to boost their interventions by: (1) integrating gender perspectives and empowering women to take on training opportunities; (2) targeting the development of an array of skills – from hard to soft, in a participatory manner; and (3) paying further attention to mechanisms to improve business survival, sales and profits as key triggers of individual wellbeing.

Provided the proximity of the government, social partners and other key stakeholders to the programme and the evaluation, the study findings can have tangible impacts in Kenya. At national level, the government and development partners can internalise the results and expand/improve implementation through already-trained service providers, such as WEF, KIE and the Kenya Youth Enterprise Development. At the local level, local governments and civil society organisations serving women can improve the design and delivery of their interventions.

Appendix

Table A1: Means by treatment status for mentoring intervention

	Assigned mentor (1)	Training alone (2)	Spillover for mentor (3)	Spillover for training alone (4)	Pure control (5)	p-value (1) vs (2)	p-value (3) vs (4)
Age	35.73	36.39	35.38	35.70	35.80	0.262	0.807
Years of education	9.08	8.76	9.00	8.89	9.08	0.019	0.321
Married	0.69	0.65	0.68	0.64	0.68	0.166	0.999
Household size	4.96	4.99	4.93	4.83	4.86	0.962	0.841
Age of firm	6.36	6.82	6.53	6.73	6.31	0.310	0.491
Number of employees	0.27	0.26	0.27	0.27	0.26	0.985	0.422
Weekly profits	1124	1115	1132	1155	1085	0.782	0.643
Weekly sales	5657	5054	6028	6016	5247	0.596	0.558
Capital stock	37859	24222	39042	27324	26733	0.001	0.005
Ever received bank/MFI loan	0.26	0.26	0.23	0.27	0.23	0.652	0.061
Keeps records	0.36	0.37	0.33	0.36	0.34	0.967	0.318
Business practices score	0.53	0.52	0.53	0.53	0.52	0.153	0.920
Retail firm	0.77	0.81	0.76	0.79	0.75	0.223	0.115
Registered with city council	0.48	0.36	0.49	0.39	0.40	0.001	0.023
Sample size	524	521	459	442	1158		

Notes: Means and sample sizes are shown for samples in non-small markets, for which mentoring intervention is applied.

Online appendixes

Online appendix A: Kenya female enterprise survey (long)

http://www.3ieimpact.org/sites/default/files/2019-04/ow4.1213-online-appendix-A-KENYA-FEMALE-ENTERPRISE-SURVEY-LONG-SURVEY.pdf

Online appendix B: Kenya female enterprise survey (short)

http://www.3ieimpact.org/sites/default/files/2019-04/ow4.1213-online-appendix-B-KENYA-FEMALE-ENTERPRISE-SURVEY-SHORT-SURVEY.pdf

Online appendix C: Follow-up census of women entrepreneurs

http://www.3ieimpact.org/sites/default/files/2019-04/ow4.1213-online-appendix-C-FOLLOW-UP-CENSUS-OF-WOMEN-ENTREPRENEURS.pdf

Online appendix D: Focus group discussion guide (phase i)

http://www.3ieimpact.org/sites/default/files/2019-04/ow4.1213-online-appendix-D-FOCUS-GROUP-DISCUSSION-GUIDE-PHASE-1.pdf

Online appendix E: Key informant interview guide: implementing organizations

http://www.3ieimpact.org/sites/default/files/2019-04/ow4.1213-online-appendix-E-KEY-INFORMANT-INTERVIEW-GUIDE-IMPLEMENTING-ORGANIZATIONS.pdf

Online appendix F: Key informant interview guide: implementing organizations (trainers)

http://www.3ieimpact.org/sites/default/files/2019-04/ow4.1213-online-appendix-F-KEY-INFORMANT-INTERVIEW-GUIDE-IMPLEMENTING-ORGANIZATIONS_0.pdf

Online appendix G: Initial qualitative results

http://www.3ieimpact.org/sites/default/files/2019-04/ow4.1213-online-appendix-G-INITIAL-QUALITATIVE-RESULTS.pdf

Online appendix H: Phase ii qualitative results

http://www.3ieimpact.org/sites/default/files/2019-04/ow4.1213-online-appendix-H-PHASE-II-QUALITATIVE-RESULTS.pdf

References

Anderson, M, 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), pp.1481-495.

Banerjee, A and Esther D, 2008. What is middle class about the middle classes around the world? *Journal of Economic Perspectives,* 22(2), pp.3-28.

Bauer, S, Finnegan, G and Haspels, N, 2004. Gender and Entrepreneurship Together: GET Ahead for Women in Enterprise: Training Package and Resource Kit. Bangkok, Berlin and Geneva: ILO.

Benjamini, Y, Abba MK, and Daniel, Y, 2006. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3), pp.491-507.

Berge, L, Bjorvatn, K and Tungodden, B, 2014. Human and financial capital for microenterprise development: evidence from a field and lab experiment. *Management Science,* 61(4), pp.707-22.

Bloom, N, Sadun, R, and Van Reenen, J, 2015. *Management as a technology*. Mimeo: Stanford.

Brooks, W, Donovan, K, and Johnson, T, 2016. *The dynamics of inter-firm skill transmission among Kenya microenterprises*. Mimeo.

Bruhn, M and McKenzie, D, 2009. In pursuit of balance: randomisation in practice in development field experiments. *American Economic Journal: Applied Economics*, 1(4), pp.200-32.

Bulte, E, Lensink, R, van Velzen, R, and Thi Hong Vu, H, 2016. *Do gender and business trainings affect business outcomes? Experimental evidence from Vietnam*. Mimeo.

Calderón, G, Cunha, J, and de Giorgi, G, 2013. Business literacy and development: evidence from a randomised trial in rural Mexico. *NBER Working Paper* No. i.

Carpena, F, Cole, S, Shapiro, J, and Zia, B, 2011. Unpacking the causal chain of financial literacy. *World Bank Policy Research Working Paper* No. 5798.

Chandy, R, Anderson-McDonald, S, and Zia, B, 2016. Pathways to profits: identifying separate channels of firm growth through business training. *World Bank Policy Research Working Paper* No. 7774.

De Mel, S, McKenzie, D, and Woodruff, C, 2009. Measuring microenterprise profits: must we ask how the sausage is made? *Journal of Development Economics*, 88(1), pp.19-31.

De Mel, S, McKenzie, D, and Woodruff, C, 2014. Business training and female enterprise start-up, growth, and dynamics: experimental evidence from Sri Lanka. *Journal of Development Economics*, 106, pp.199-210.

Diwan, F, Makana, G, McKenzie, D, and Paruzzolo, S, 2014. Invitation choice structure has no impact on attendance in a female business training programme in Kenya. *PLOS ONE*, 9(10), p.e109873.

Drexler, A, Fischer, G, and Schoar, A, 2011. *Keeping it simple: financial literacy and rules of thumb*. Mimeo: LSE.

Fafchamps, M, McKenzie, D, Quinn S, and Woodruff, C, 2012. Using PDA consistency checks to increase the precision of profits and sales measurement. *Journal of Development Economics*, 98(1), pp.51-7.

Gindling, TH, and Newhouse, D, 2014. Self-employment in the developing world. *World Development,* 56, pp.313-31.

Giné, X, and Mansuri, G, 2016. *Money or management? A field experiment on constraints to entrepreneurship in rural Pakistan*. Mimeo: World Bank.

Grimm, M, and Paffhausen, AL, 2015. Do interventions targeted at micro-entrepreneurs and small and medium-sized firms create jobs? A systematic review of the evidence for low and middle-income countries. *Labour Economics*, 32, pp.67-85.

ICRW, 2015. 'I am determined, no matter how hard it becomes, I will not give up': A qualitative assessment of the ILO's GET Ahead Business Training Programme. ICRW.

IFC, 2010. Scaling-up SME access to financial services in the developing world: Financial Inclusion Experts Group | SME Finance Sub-Group: International Finance Corporation. World Bank Group.

ILO, 2015. Small and medium-sized enterprises and decent and productive employment creation. Report IV. International Labour Conference, 104th Session, 2015. Geneva.

Karlan, D, and Valdivia, M, 2011. Teaching entrepreneurship: impact of business training on microfinance clients and institutions. *Review of Economics and Statistics*, 93(2), pp.510-27.

Karlan, D, Knight, R, and Udry, C, 2012. *Hoping to win, expected to lose: Theory and lessons on microenterprise development*. Mimeo: Yale.

Lock, R, and Lawton Smith, H, 2015. The impact of female entrepreneurship on economic growth in Kenya. CIMR Research Working Paper Series, Working Paper No. 26. Available at: <http://www.bbk.ac.uk/innovation/publications/docs/WP26.pdf>.

McKenzie, D, 2011. How can we learn whether firm policies are working in Africa? Challenges (and solutions?) for experiments and structural models. Journal of African Economies, 20(4), pp.600-625.

McKenzie, D, 2012. Beyond baseline and follow-up: the case for more t in experiments. *Journal of Development Economics,* 99(2), pp.210-21.

McKenzie, D, and Woodruff, C, 2013. What are we learning from business training and entrepreneurship evaluations around the developing world? *The World Bank Research Observer*, 29(1), pp.48-82.

McKenzie, D, and Woodruff, C, 2015. Business practices in small firms in developing countries. *Management Science* (forthcoming).

McKenzie, D, 2017. Identifying and spurring high-growth entrepreneurship: experimental evidence from a business plan competition. *American Economic Review*, 107(8), pp.2278-307

Rotemberg, M, 2014. *Equilibrium effects of firm subsidies*. Mimeo.

Valdivia, M, 2015. Business training plus for female entrepreneurship? Short and medium-term experimental evidence from Peru. *Journal of Development Economics,* 113, pp.33-51.

Veit, CT and Ware, Jnr, JE, 1983. The structure of psychological distress and wellbeing in general populations. *Journal of Consulting and Clinical Psychology*, 51, pp.730-42.

World Bank, 2012. World development report 2012: Gender equality and development. Washington DC: World Bank. Available at: <https://siteresources.worldbank.org/INTWDR2012/Resources/7778105-1299699968583/7786210-1315936222006/Complete-Report.pdf>.

World Bank, 2014. Country Partnership Strategy for Kenya FY2014-2018: Partnership for shared growth and prosperity. Washington DC: World Bank. Available at: <http://documents.worldbank.org/curated/en/173431468284364640/pdf/889400CAS0P1440Kenya0CPS000Volume02.pdf>.

Other publications in the 3ie Impact Evaluation Report Series

The following reports are available from http://3ieimpact.org/evidence-hub/publications/impact-evaluations

Impacts of key provisions in Ghana's Petroleum Revenue Management Act, 3ie Impact Evaluation Report 94. Edjekumhene, I, Voors, M, Lujala, P, Brunnschweiler, C, Owusu, CK and Nyamekye, A, 2019.

Using information to break the political resource curse in natural gas management in Mozambique, 3ie Impact Evaluation Report 93. Armand, A, Costa, AI, Coutts, A, Vicente, P and Vilela, I, 2019.

Harnessing transparency initiatives to improve India's environmental clearance process for the mineral mining sector, 3ie Impact Evaluation Report 92. Pande, R and Sudarshan, A, 2019.

Impacts of removing user fees for maternal health services on universal health coverage in Kenya, 3ie Impact Evaluation Report 91. Abuya, T, Dennis, M, Matanda, D, Obare, F and Bellows, B, 2018.

Impact of voice reminders to reinforce harvest aggregation services training for farmers in Mali, 3ie Impact Evaluation Report 90. Osei, RD, Dzanku, FM, Osei-Akoto, I, Asante, F, Hodey, LS, Adu, PN, Adu-Ababio, K and Coulibaly, M, 2018.

Impacts of Breakthrough's school-based gender attitude change programme in Haryana, India, 3ie Impact Evaluation Report 89. Jayachandran, S, Jain, T and Dhar, D, 2018.

Hotspot interventions at scale: the effects of policing and city services on crime in Bogotá, Colombia, 3ie Impact Evaluation Report 88. Blattman, C, Green, D, Ortega, D and Tobón, S, 2018.

Impact evaluation of the Philippine Special Program for Employment of Students, 3ie Impact Evaluation Report 87. Beam, E, Linden, L, Quimbo, S and Richmond, H, 2018.

Community-based distribution of oral HIV self-testing kits: experimental evidence from Zambia, 3ie Impact Evaluation Report 86. Hensen, B, Ayles, H, Mulubwa, C, Floyd, S, Schaap, A, Chiti, B, Phiri, M, Mwenge, L, Simwinga, M, Fidler S, Hayes, R, Bond, V and Mwinga, A, 2018.

Evaluating the economic impacts of rural banking: experimental evidence from southern India, 3ie Impact Evaluation Report 85. Field, E and Pande, R, 2018.

Direct provision versus facility collection of HIV tests: impacts of self-testing among female sex workers in Uganda. 3ie Impact Evaluation Report 84. Ortblad, K, Musoke, DK, Ngabirano, T, Oldenburg, C and Bärnighausen, T, 2018.

Increasing female sex worker HIV testing: effects of peer educators and HIV self-tests in Zambia, 3ie Impact Evaluation Report 83. Chanda, MM, Ortblad, KF, Mwale, M, Chongo, S, Kanchele, C, Kamungoma, N, Fullem, A, Bärnighausen, T and Oldenburg, CE, 2018.

Community delivery of antiretroviral drugs: a non-inferiority matched-pair pragmatic cluster-randomized trial in Dar es Salaam, Tanzania, 3ie Impact Evaluation Report 82. Francis, JM, Geldsetzer, P, Asmus, G, Ulenga, N, Ambikapathi, R, Sando, D, Fawzi, W and Bärnighausen, T, 2018.

Nourishing the future: targeting infants and their caregivers to reduce undernutrition in rural China, 3ie Impact Evaluation Report 81. Cai, J, Luo, R, Li, H, Lien, J, Medina, A, Zhou, H and Zhang, L, 2018.

Impacts of the World Food Programme's interventions to treat malnutrition in Niger. 3ie Impact Evaluation Report 80. Brück, T, Ferguson, NTN, Ouédraogo, J and Ziegelhöfer, Z, 2018.

Impact evaluation of the World Food Programme's moderate acute malnutrition treatment and prevention programmes in Sudan. 3ie Impact Evaluation Report 79. Guevarra, E, Mandalazi, E, Balegamire, S, Albrektsen, K, Sadler, K, Abdelsalam, K, Urrea, G and Alawad, S, 2018.

Impact evaluation of WFP's programs targeting moderate acute malnutrition in humanitarian situations in Chad. 3ie Impact Evaluation Report 78. Saboya, M, Rudiger, J, Frize, J, Ruegenberg, D, Rodríguez Seco, A and McMillon, C, 2018.

Improving midday meal delivery and encouraging micronutrient fortification among children in India, 3ie Impact Evaluation Report 77. Shastry, GK, Berry, J, Mukherjee, P, Mehta, S and Ruebeck, H, 2018.

Evaluation of infant development centres: an early years intervention in Colombia, 3ie Impact Evaluation Report 76. Andrew, A, Attanasio, O, Bernal, R, Cordona, L, Krutikova, S, Heredia, DM, Medina, C, Peña, X, Rubio-Codina, M and Vera-Hernandez, M, 2018.

Can the wounds of war be healed? Experimental evidence on reconciliation in Sierra Leone. 3ie Impact Evaluation Report 75. Cilliers, J, Dube, O and Siddiqi, B, 2018.

Impact evaluation of the Menabe and Melaky development programme in Madagascar, 3ie Impact Evaluation Report 74. Ring, H, Morey, M, Kavanagh, E, Kamto, K, McCarthy, N, Brubaker, J and Rakotondrafara, C, 2018.

Impact evaluation of the Smallholder Dairy Commercialization Programme in Kenya, 3ie Impact Evaluation Report 73. Bonilla, J, McCarthy, N, Mugatha, S, Rai, N, Coombes, A and Brubaker, J, 2018.

Impact and adoption of risk-reducing drought-tolerant rice in India, 3ie Impact Evaluation Report 72. Yamano, T, Dar, MH, Panda, A, Gupta, I, Malabayabas, ML and Kelly, E, 2018.

Poverty and empowerment impacts of the Bihar Rural Livelihoods Project in India, 3ie Impact Evaluation Report 71. Hoffmann, V, Rao, V, Datta, U, Sanyal, P, Surendra, V and Majumdar, S 2018.

How should Tanzania use its natural gas? Citizens' views from a nationwide Deliberative Poll, 3ie Impact Evaluation Report 70. Birdsall, N, Fishkin, J, Haqqi, F, Kinyondo, A, Moyo, M, Richmond, J and Sandefur, J, 2018.

Impact evaluation of the conditional cash transfer program for secondary school attendance in Macedonia, 3ie Impact Evaluation Report 69. Armand, A and Carneiro, P, 2018.

Age at marriage, women's education, and mother and child outcomes in Bangladesh, 3ie Impact Evaluation Report 68. Field, E, Glennerster, R, Nazneen, S, Pimkina, S, Sen, I and Buchmann, N, 2018.

Evaluating agricultural information dissemination in western Kenya, 3ie Impact Evaluation Report 67. Fabregas, R, Kremer, M, Robinson, J and Schilbach, F, 2017.

General equilibrium impact assessment of the Productive Safety Net Program in Ethiopia, 3ie Impact Evaluation Report 66. Filipski, M, Taylor, JE, Abegaz, GA, Ferede, T, Taffesse, AS and Diao, X, 2017.

Impact of the Uddeepan programme on child health and nutrition in India, 3ie Impact Evaluation Report 65. Kochar, A, Sharma, A and Sharma, A, 2017.

Evaluating oral HIV self-testing to increase HIV testing uptake among truck drivers in Kenya, 3ie Impact Evaluation Report 64. Kelvin, EA, Mwai, E, Romo, ML, George, G, Govender, K, Mantell, JE, Strauss, M, Nyaga, EN and Odhiambo, JO, 2017.

Integration of EPI and paediatric HIV services for improved ART initiation in Zimbabwe, 3ie Impact Evaluation Report 63. Prescott, M, Boeke, C, Gotora, T, Mafaune, HW, Motsi, W, Graves, J, Mangwiro, A and McCarthy, E, 2017.

Increasing male partner HIV testing using self-test kits in Kenya, 3ie Impact Evaluation Report 62. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

Evaluating the impact of community health worker integration into prevention of mother-to-child transmission of HIV services in Tanzania, 3ie Impact Evaluation Report 61. Nance, N, McCoy, S, Ngilangwa, D, Masanja, J, Njau, P and Noronha, R, 2017.

Using HIV self-testing to promote male partner and couples testing in Kenya, 3ie Impact Evaluation Report 60. Thirumurthy, H, Omanga, E, Obonyo, B, Masters, S and Agot, K, 2017.

Increasing male partner HIV self-testing at antenatal care clinics in Kenya, 3ie Impact Evaluation Report 59. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

Impact of free availability of public childcare on labour supply and child development in Brazil, 3ie Impact Evaluation Report 58. Attanasio, O, Paes de Barros, R, Carneiro, P, Evans, D, Lima, L, Olinto, P and Schady, N, 2017.

Estimating the effects of a low-cost early stimulation and parenting education programme in Mexico, 3ie Impact Evaluation Report 57. Cardenas, S, Evans, D and Holland, P, 2017.

The Better Obstetrics in Rural Nigeria study: an impact evaluation of the Nigerian Midwives Service Scheme, 3ie Impact Evaluation Report 56. Okeke, E, Glick, P, Abubakar, IS, Chari, AV, Pitchforth, E, Exley, J, Bashir, U, Setodji, C, Gu, K and Onwujekwe, O, 2017.

*The Productive Safety Net Programme in Ethiopia: impacts on children's schooling, labour and nutritional statu*s, 3ie Impact Evaluation Report 55. Berhane, G, Hoddinott, J, Kumar, N and Margolies, A, 2016.

The impact of youth skills training on the financial behaviour, employability and educational choice in Morocco, 3ie Impact Evaluation Report 54. Bausch, J, Dyer, P, Gardiner, D, Kluve, J and Mizrokhi, E, 2016.

Using advertisements to create demand for voluntary medical male circumcision in South Africa, 3ie Impact Evaluation Report 53. Frade, S, Friedman, W, Rech, D and Wilson, N, 2016.

The use of peer referral incentives to increase demand for voluntary medical male circumcision in Zambia, 3ie Impact Evaluation Report 52. Zanolini, A, Bolton, C, Lyabola, LL, Phiri, G, Samona, A, Kaonga, A and Harsha Thirumurthy, H, 2016.

Using smartphone raffles to increase demand for voluntary medical male circumcision in Tanzania, 3ie Impact Evaluation Report 51. Mahler, H and Bazant, E, 2016.

Voluntary medical male circumcision uptake through soccer in Zimbabwe, 3ie Impact Evaluation Report 50. DeCelles, J, Kaufman, Z, Bhauti, K, Hershow, R, Weiss, H, Chaibva, C, Moyo, N, Braunschweig, E, Mantula, F, Hatzold, K and Ross, D, 2016.

Measuring the impact of SMS-based interventions on uptake of voluntary medical male circumcision in Zambia, 3ie Impact Evaluation Report 49. Leiby, K, Connor, A, Tsague, L, Sapele, C, Koanga, A, Kakaire, J and Wang, P, 2016.

Assessing the impact of delivering messages through intimate partners to create demand for voluntary medical male circumcision in Uganda, 3ie Impact Evaluation Report 48. Semeere, AS, Bbaale, DS, Castelnuovo, B, Kiragga, A, Kigozi, J, Muganzi, A, Kambugu, A and Coutinho, AG, 2016.

Optimising the use of economic interventions to increase demand for voluntary medical male circumcision in Kenya, 3ie Impact Evaluation Report 47. Thirumurthy, H, Omanga, E, Rao, SO, Murray, K, Masters, S and Agot, K, 2016.

The impact of earned and windfall cash transfers on livelihoods and conservation in Sierra Leone, 3ie Impact Evaluation Report 46. Bulte, E, Conteh, B, Kontoleon, A, List, J, Mokuwa, E, Richards, P, Turley, T and Voors, M, 2016.

Property tax experiment in Pakistan: Incentivising tax collection and improving performance, 3ie Impact Evaluation Report 45. Khan, A, Khwaja, A and Olken, B, 2016.

Impact of mobile message reminders on tuberculosis treatment outcomes in Pakistan, 3ie Impact Evaluation Report 44. Mohammed, S, Glennerster, R and Khan, A, 2016.

Making networks work for policy: Evidence from agricultural technology adoption in Malawi, 3ie Impact Evaluation Report 43. Beaman, L, BenYishay, A, Fatch, P, Magruder, J and Mobarak, AM, 2016.

Estimating the impact and cost-effectiveness of expanding access to secondary education in Ghana, 3ie Impact Evaluation Report 42. Dupas, P, Duflo, E and Kremer, M, 2016.

*Evaluating the effectiveness of computers as tutors in China, 3*ie Impact Evaluation Report 41. Mo, D, Bai, Y, Boswell, M and Rozelle, S, 2016.

Micro entrepreneurship support programme in Chile, 3ie Impact Evaluation Report 40. Martínez, CA, Puentes, EE and Ruiz-Tagle, JV, 2016.

Thirty-five years later: evaluating the impacts of a child health and family planning programme in Bangladesh, 3ie Impact Evaluation Report 39. Barham, T, Kuhn, R, Menken, J and Razzaque, A, 2016.

Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection and malnutrition in India, 3ie Impact Evaluation Report 38. Clasen, T, Boisson, S, Routray, P, Torondel, B, Bell, M, Cumming, O, Ensink, J, Freeman, M and Jenkins, M, 2016.

Evaluating the impact of vocational education vouchers on out-of-school youth in Kenya, 3ie Impact Evaluation Report 37. Hicks, JH, Kremer, M, Mbiti, I and Miguel, E, 2016.

Removing barriers to higher education in Chile: evaluation of peer effects and scholarships for test preparation, 3ie Impact Evaluation Report 36. Banerjee, A, Duflo E and Gallego, F, 2016.

Sustainability of impact: dimensions of decline and persistence in adopting a biofortified crop in Uganda, 3ie Impact Evaluation Report 35. McNiven, S, Gilligan, DO and Hotz, C 2016.

A triple win? The impact of Tanzania's Joint Forest Management programme on livelihoods, governance and forests, 3ie Impact Evaluation Report 34. Persha, L and Meshack, C, 2016.

The effect of conditional transfers on intimate partner violence: evidence from Northern Ecuador, 3ie Impact Evaluation Report 33. Hidrobo, M, Peterman, A and Heise, L, 2016.

The effect of transfers and preschool on children's cognitive development in Uganda, 3ie Impact Evaluation Report 32. Gillian, DO and Roy, S, 2016.

Can egovernance reduce capture of public programmes? Experimental evidence from India's employment guarantee, 3ie Impact Evaluation Report 31. Banerjee, A, Duflo, E, Imbert, C, Mathew, S and Pande, R, 2015.

Improving maternal and child health in India: evaluating demand and supply strategies, 3ie Impact Evaluation Report 30. Mohanan, M, Miller, G, Forgia, GL, Shekhar, S and Singh, K, 2016.

Smallholder access to weather securities in India: demand and impact on production decisions, 3ie Impact Evaluation Report 28. Ceballos, F, Manuel, I, Robles, M and Butler, A, 2015.

What happens once the intervention ends? The medium-term impacts of a cash transfer programme in Malawi, 3ie Impact Evaluation Report 27. Baird, S, Chirwa, E, McIntosh, C and Özler, B, 2015.

Validation of hearing screening procedures in Ecuadorian schools, 3ie Impact Evaluation Report 26. Muñoz, K, White, K, Callow-Heusser, C and Ortiz, E, 2015.

Assessing the impact of farmer field schools on fertilizer use in China, 3ie Impact Evaluation Report 25. Burger, N, Fu, M, Gu, K, Jia, X, Kumar, KB and Mingliang, G, 2015.

The SASA! study: a cluster randomised trial to assess the impact of a violence and HIV prevention programme in Kampala, Uganda, 3ie Impact Evaluation Report 24. Watts, C, Devries, K, Kiss, L, Abramsky, T, Kyegombe, N and Michau, L, 2014.

Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts, 3ie Impact Evaluation Report 23. Gine, X, Patel, S, Cuellar-Martinez, C, McCoy, S and Lauren, R, 2015.

A wide angle view of learning: evaluation of the CCE and LEP programmes in Haryana, 3ie Impact Evaluation Report 22. Duflo, E, Berry, J, Mukerji, S and Shotland, M, 2015.

Shelter from the storm: upgrading housing infrastructure in Latin American slums, 3ie Impact Evaluation Report 21. Galiani, S, Gertler, P, Cooper, R, Martinez, S, Ross, A and Undurraga, R, 2015.

Environmental and socioeconomic impacts of Mexico's payments for ecosystem services programme, 3ie Impact Evaluation Report 20. Alix-Garcia, J, Aronson, G, Radeloff, V, Ramirez-Reyes, C, Shapiro, E, Sims, K and Yañez-Pagans, P, 2015.

A randomised evaluation of the effects of an agricultural insurance programme on rural households' behaviour: evidence from China, 3ie Impact Evaluation Report 19. Cai, J, de Janvry, A and Sadoulet, E, 2014.

Impact of malaria control and enhanced literacy instruction on educational outcomes among school children in Kenya: a multi-sectoral, prospective, randomised evaluation, 3ie Impact Evaluation Report 18. Brooker, S and Halliday, K, 2015.

Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua, 3ie Impact Evaluation Report 17. Barham, T, Macours, K, Maluccio, JA, Regalia, F, Aguilera, V and Moncada, ME, 2014.

The impact of mother literacy and participation programmes on child learning: evidence from a randomised evaluation in India, 3ie Impact Evaluation Report 16. Banerji, R, Berry, J and Shortland, M, 2014.

A youth wage subsidy experiment for South Africa, 3ie Impact Evaluation Report 15. Levinsohn, J, Rankin, N, Roberts, G and Schöer, V, 2014.

Providing collateral and improving product market access for smallholder farmers: a randomised evaluation of inventory credit in Sierra Leone, 3ie Impact Evaluation Report 14. Casaburi, L, Glennerster, R, Suri, T and Kamara, S, 2014.

Scaling up male circumcision service provision: results from a randomised evaluation in Malawi, 3ie Impact Evaluation Report 13. Thornton, R, Chinkhumba, J, Godlonton, S and Pierotti, R, 2014.

Targeting the poor: evidence from a field experiment in Indonesia, 3ie Impact Evaluation Report 12. Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R, 2014.

An impact evaluation of information disclosure on elected representatives' performance: evidence from rural and urban India, 3ie Impact Evaluation Report 11. Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M and Mahapatra, B, 2014.

Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10. Duflo, E, Greenstone, M, Pande, R and Ryan, N, 2013.

No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia, 3ie Impact Evaluation Report 9. Ashraf, N, Bandiera, O and Jack, K, 2013.

Paying for performance in China's battle against anaemia, 3ie Impact Evaluation Report 8. Zhang, L, Rozelle, S and Shi, Y, 2013.

Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in the Democratic Republic of Congo, 3ie Impact Evaluation Report 7. Humphreys, M, Sanchez de la Sierra, R and van der Windt, P, 2013.

The impact of daycare on maternal labour supply and child development in Mexico, 3ie Impact Evaluation Report 6. Angeles, G, Gadsden, P, Galiani, S, Gertler, P, Herrera, A, Kariger, P and Seira, E, 2014.

Impact evaluation of the non-contributory social pension programme 70 y más in Mexico, 3ie Impact Evaluation Report 5. Rodríguez, A, Espinoza, B, Tamayo, K, Pereda, P, Góngora, V, Tagliaferro, G and Solís, M, 2014.

Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4. Meenakshi, JV, Banerji, A, Mukherji, A and Gupta, A, 2013.

The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone, 3ie Impact Evaluation Report 3. Casey, K, Glennerster, R and Miguel, E, 2013.

A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana, 3ie Impact Evaluation Report 2. Burwen, J and Levine, DI, 2012.

The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique, 3ie Impact Evaluation Report 1. Martinez, S, Naudeau, S and Pereira, V, 2012.

The International Labour Organization's Gender and Entrepreneurship Together training programme (GET Ahead) seeks to enhance women's opportunities in entrepreneurship through knowledge and skills development in business and management. McKenzie and colleagues evaluated the impact of offering women with small businesses in rural Kenya a five-day training course. A year after training, some women who had received GET Ahead were offered complementary group and individual support services over a five-month period. After three years, women who received training earn higher profits, and are more likely to have surviving businesses and higher weekly sales. A key mechanism for this market growth appears to be the training they received: business owners began keeping more reliable opening hours and diversifying their range of products. The increase in business income is accompanied by improvements in mental health and subjective well-being.

Impact Evaluation Series

International Initiative for Impact Evaluation
202-203, Rectangle One
D-4, Saket District Centre
New Delhi – 110017
India

3ie@3ieimpact.org
Tel: +91 11 4989 4444

www.3ieimpact.org