Short and Longer Run Impacts of *Kaizen* Management Training: The Case in Tanzania[*]

Yuki Higuchi[a] and Tetsushi Sonobe[b]

a. *Nagoya City University, Nagoya, Japan*

b. *National Graduate Institute for Policy Studies, Tokyo, Japan*

February 2016

Abstract

We conducted a randomized experiment among Tanzanian small garment manufacturers to examine the impacts of short-term basic management training. Our training featured a common-sense approach to production management, *Kaizen*, and was taught in two different modes: one in classroom and the other on site. Using panel data collected before and after the training intervention in the span of four years with negligible attrition rate, we find that our training program improved entrepreneurs' management capacity and such improvement remained significant for three years after the intervention. Although the training impact on business performance was not significant shortly after the training, the entrepreneurs who participated in both classroom and on-site training had significantly higher value added three years after the training program. Such business improvement can be explained by the selective adoption of relevant management practices and by the changes in investment behavior of the training participants. The time lag between the training intervention and the business improvement suggests the importance of collecting panel data in a long span to evaluate impacts of management training.

Keywords: Management training, Small enterprise, Randomized controlled trial, Industrial cluster, Tanzania

JEL classification: L2, M1, O1

[*] We would like to thank Keijiro Otsuka, Yasuyuki Todo, Takashi Kurosaki and Shinsuke Uchida for helpful comments. This research was supported by the World Bank Japan PHRD Trust Fund No. TF096317, JSPS KAKENHI Grant Numbers 25101002 and 15H06540, and Grant-in-Aid for Research in Nagoya City University.

1. Introduction

World Bank (2012) emphasizes that a way to reduce poverty is to create decent jobs by developing the industrial sector. Although it is entrepreneurs that are supposed to hold the key to successful industrial development, their managerial capacity, let alone their capacity to spur innovation, is highly limited in developing countries (e.g., Bloom, Genakos, et al. 2012; Bruhn et al, 2010; Sonobe and Otsuka, 2011). Hence, there has been renewed interest in entrepreneurs' capacity building, especially through management training. Using randomized controlled trials (RCTs), an emerging body of literature finds that entrepreneurs' management capacity can be improved by management training (e.g., Berge, Bjorvatn, Juniwaty et al., 2012; Bjorvatn, Juniwaty, Tungodden, 2014; Bloom, Eifert et al., 2013; Bruhn and Zia, 2013; De Mel et al., 2014; Drexler et al., 2014; Field et al., 2010; Karlan and Valdivia, 2011; Udry et al., 2012).

McKenzie and Woodruff (2014), however, point out that this line of research has yet to provide useful information for policy makers. Although these studies confirm that training intervention improves entrepreneurs' management skills, they have not presented clear evidence of favorable impacts on the business performance of the trainees. The limited training impact on business performance could be attributed to noisy data, small sample sizes, or inadequately designed training program, but it is possible that the impacts are evaluated too early. Indeed, the follow-up surveys of the existing studies tend to be conducted within a short period, say, a year after the respective training intervention, while it may take more time for the trained entrepreneurs to choose useful management practices through trial and error or to substantially change their business operation. Such possibilities are recognized by business consultants (e.g., Imai, 2012; Morgan and Liker, 2006); no rigorous empirical studies, however, has been conducted by researchers

with a notable exception by Bloom, Eifert et al. (2013), which analyzed the long-term impacts of intensive consultation to a small number of large-sized enterprises in India.

This paper is an attempt to evaluate longer-term impacts of management training programs than the majority of existing studies. We provided a short-term management training to small garment manufacturers selected randomly from an industrial cluster in Tanzania. The focus on manufacturers, rather than microentrepreneurs, allows us to keep the attrition of samples at a minimum because manufacturers who purchased production equipment stay in the same strand of business for a long period of time. In addition, our focus on the industrial cluster allows us to control for a great deal of variation that would otherwise be introduced if we were to broaden our study across a large number of entrepreneurs in various industries. Our training program introduced *Kaizen*, a common-sense and inexpensive approach to production management, which had greatly contributed to the development of various industries both in developed and developing countries, including the U.S. automobile industry (Higuchi et al, 2015; Sonobe and Otsuka, 2014, Van Biesebroeck, 2003). *Kaizen* approach was taught by experienced experts in two different modes: one in classroom and the other on site.

Analyzing panel data collected before and after the training intervention in the span of four years, we find that the training has positive and significant impacts on the management capacity both in short and longer-run. In the short-run, the trainees adopted many management practices taught in our training, presumably, without knowing whether they are relevant to their business or not. They, however, started to select useful and suitable practices on trial-and-error basis, and selectively adopt the useful practices in the longer run. We also find suggestive evidence of knowledge spillover from the participants to the non-participants. Many of the sample entrepreneurs know each other

and they have talked about *Kaizen* after our training. Such interactions occurred not only between the participants but also between the participants and non-participants. Hence, through learning from the participants, the non-participants became willing to pay for our training and, indeed, improved their management capacity.

With regard to the business performance, the training impacts on value added and sales revenue were not significant shortly after the intervention. The impacts, however, emerged three years after the intervention. The time lag between the training intervention and improved business performance can possibly be explained by two factors. Firstly, as suggested by the differences of training impact on management in the short-term and in the longer-term, training participants needed time to find management practices that are useful to improve their business. Secondly, it took time for the participants to substantially change their business by making investment. Our regression analyses suggest that although the participants are no more likely to invest within a year after the intervention, they are significantly more likely to do so in a longer span. Therefore, the training impact on business performance only emerged in a longer-run. Such time lag between the training intervention and the business improvement suggests the importance of collecting panel data in a long span to evaluate the impacts of management training.

The remainder of this paper is organized as follows. Section 2 describes the experimental design and Section 3 checks the internal validity of our experiment by addressing the issues of balance, attrition and externality. Section 4 presents the impact evaluation results and Section 5 summarizes the findings and discusses implications for future studies.

2. Experimental Design

2.1. Study Sites and Sample Enterprises

The garment industry has widely been observed in developing countries (Pack, 1987). In some developing countries like Bangladesh, the garment industry has developed and a number of large-scale firms mass-produce and export the products (Mottaleb and Sonobe, 2011). In many other developing countries, however, the industry is dominated by small manufacturers, which tend to be clustered in geographical proximity (Sonobe and Otsuka, 2011); our study site is no exception. In Dar es Salaam, the business capital of Tanzania, a cluster of garment industry had started to be formed by female entrepreneurs, who began their business after participating in business training programs by United Nations Industrial Development Organization (UNIDO) in the 1990s. The UNIDO training targeted mostly housewives in Dar es Salaam and covered sewing skill, book keeping, and business planning, and many training participants started their business in their houses. Although not included in our sample, one of the pioneering female entrepreneurs employs more than 30 regular workers and exports her products to Europe. Many of the entrepreneurs in this cluster desire to follow her lead and shift from tailor-type production to mass-production for exporting to developed country.

We focus on the garment or batik producers who are members of one of the following three associations: Tanzania Handicraft Association (TANCRAFT), Handproducts of Tanzania (HOT) and Artisan Development Association of Tanzania (ADAT).[1] We randomly selected 114 entrepreneurs out of about 250 members of these

1 There is another association of garment producers called SHIME, which we excluded their members from our scope since they market their products together in their own marketplace and hence the marketing part of our training program would be little appreciated by them.

associations as our baseline sample. The sample entrepreneurs employ on average four workers while twelve of them are single-person entrepreneurs.

2.2. Experimental Intervention

In most existing studies, management training interventions highlighted basic skills in accounting, marketing, and business strategy (McKenzie and Woodruff, 2014). Production management including quality control was seldom taught in these training programs. By contrast, our interventions featured *Kaizen* approach to production management. *Kaizen* is a Japan-pioneered approach, based on US-born industrial engineering and quality management, and emphasizes productivity improvement by the collaborative and continuous effort of entrepreneurs and their workers (Imai, 2012). We made a contract with a business consulting firm in Japan to dispatch experienced (with at least ten years of experience) *Kaizen* experts to our study site and we also hired local consultants. The *Kaizen* experts taught the local consultants in English, and the latter taught the training participants in local language. Therefore, any failure to find training impact should not be because of low quality of instructors or miscommunication between the instructors and sample entrepreneurs.

As shown in Figure 1, our training intervention had two components: classroom training for 2.5 hours a day, five days a week over four weeks span, and on-site training for minimum two and up to eight days. In the classroom training program, the team of instructors taught standard contents of business development service, such as entrepreneurship, business strategy, marketing, bookkeeping and product design in

addition to *Kaizen* for four weeks.[2] Two classroom training participants were selected to be model enterprises to serve as showcases of *Kaizen* practices. At the two selected enterprises, the team convinced the entrepreneurs to change the layout of their workshops. Subsequently, a one-day seminar was held, as a part of on-site training, to let the model entrepreneurs give presentations about their enterprises' physical changes and the responses from their workers as well as their own opinions. After the seminar, the team visited participants' enterprises for minimum two and maximum eight times to demonstrate how to apply *Kaizen* practices to improve their work environment, productivity, and product quality. The number of visits was determined by the local consultants, depending on the willingness and availability of each trainee. We provided two different modes of training, i.e., classroom and on-site, in order to compare their relative effectiveness and scalability.

2.3. Randomization and Take-up

We randomly assigned the total of 114 sample enterprises in our baseline samples into three treatment groups and a control group. The first treatment group was invited to both the classroom and on-site training programs and labeled as Group TT. The second and third were labeled Group TC and Group CT, respectively, and were invited only to either the classroom (TC) or the on-site program (CT). Group CC, the control group, was invited to neither of the programs. The model enterprises mentioned above, however, were not selected randomly. This is because, to serve as a model, an enterprise

2 The standard contents of business development service is taken from International Labour Organisation (ILO)'s Start and Improve Your Business (SYIB) programme. SYIB programme has introduced in more than 100 countries (ILO, 2003).

had to be willing to accept other training participants' visits and hence had to have enough space. Because they were treated differently from the other enterprises in Group TT or Group TC, we exclude them from our analyses. In addition, we exclude three enterprises from which we were not able to collect reliable data, leaving 109 enterprises for the empirical analyses.

Of the 50 sample enterprises invited to the classroom training programs (Groups TT and TC), 45 participated in the program as shown toward the bottom of Table 1. The sample enterprises were regarded as participants when they attended at least ten days of the training out of the twenty days. The take-up rate for the classroom training program was 90 percent. By contrast, all the 54 enterprises invited to the on-site training program (Group TT and CT) received our consultation, presumably because of the instructors' visit to each enterprise, which would reduce opportunity cost of receiving consultation. We should note, however, that some enterprises stopped receiving consultation in the middle of the course and did not complete the training as recommended by the trainers. While we regard those enterprises as on-site training participants in our main analyses, we provide empirical analyses treating those enterprises as non-participants in the Appendix. For both classroom and on-site training programs, all the participants were those who had been invited to the programs, and thus, there was no defiers.

3. Internal Validity

3.1. Balance

Panel A of Table 1 presents the means of characteristics of the sample entrepreneurs by the treatment status. Middle-aged female entrepreneurs are dominant in our sample as the industry was developed by housewives who had received UNIDO training in the

1990s. Chagga, which is an ethic tribe known for being hardworking and having business network throughout Tanzania, accounts for a quarter of our sample. A point to note from this panel is the educational attainment of our sample entrepreneurs is much higher than the average schooling attainment in Tanzania. Together with the facts that majority of the samples had past business training experience or working experience in the textile industry, the high educational attainment may illustrate that only a selected population could become manufacturer given the economic and business environment in Tanzania. To the extent that none of the *p*-values reported in column 5 is statistically significant, our randomization was successful.

To measure management capacity of our sample entrepreneurs, we constructed a management score based on 27 diagnostic criteria listed in Appendix Table A1. During the surveys, our enumerators visited each sample enterprise and judged whether the enterprise met each criterion based on either their visual inspection or the way in which the entrepreneur responded to their questions.[3] For each enterprise, the management score is the number of the diagnostic criteria that it was found to meet, and, hence, the lowest possible score is zero and the highest is 27. The higher the score, the better management capacity an entrepreneur has. To measure business performance of the entrepreneurs, we use real value added and real sales revenue.[4] Following the lead of McKenzie (2012), who suggests that taking averages over time helps to reduce noise in data on accounting-based performance indicators with low autocorrelation, we take the

3 In order to ensure data quality, all of our enumerators had bachelor degree and had received enumerator training before the beginning of survey.

4 We did not ask the respondents the amount of their value added. Instead, we asked the quantity and price of each items produced, material costs, subcontracting costs, energy costs, and transportation costs. We computed the value added by subtracting aforementioned costs from sales revenue.

average values in 2008 and 2009 as the baseline business performance. In order to compare business performance covering several years we adjust the values by using PPP conversion factor from World Bank's World Development Indicators.[5]

Panel B shows the baseline management score and business performance. While Group TT seems to have higher baseline management score than other groups, presumably reflecting insignificant but higher human capital in terms of schooling and business training, the difference across the groups is not significant as shown by the *p*-values in column 5. In contrast, Group CC have higher baseline value added, together with larger standard deviation, because of two enterprises with exceptionally large value added and sales revenue. The insignificant *p*-values for business performance variables —value added, sales revenue and number of workers—suggests the success of our randomization; we, however, control for the baseline values of these variables when conducting econometric analyses in order to control for heterogeneity of our sample entrepreneurs.

3.2. Attrition

Five rounds of surveys were conducted before and after the training intervention as shown in Figure 1. The baseline survey was conducted before the training programs and data was collected on characteristics of sample entrepreneurs, baseline management score and business performance in 2008 and 2009. The first follow-up survey was conducted soon after the classroom training program, i.e., before the start of on-site training; the second follow-up survey was conducted in April 2011, i.e., soon after the completion of

5 We also used exchange rates and GDP deflators to adjust the values of business performance but the results were essentially the same. These results are available upon request to the first author.

the on-site training. From the baseline survey to the second follow-up survey, there was no incidence of attrition. By the time of the third follow-up survey in September 2012, however, three entrepreneurs had stopped operation and thus data was not collected from them. Six out of the original 109 enterprises had stopped operation by the time of fourth follow-up survey in January 2014, that is, nearly four years after the baseline survey.[6]

In contrast to high attrition rate in the existing studies of management training (McKenzie and Woodruff, 2014), the attrition rate is kept to a minimum in our study, likely because of our focus on manufacturers. The manufacturers tend to stay in the same business longer than microentrepreneurs due to the purchased production equipment. Although not reported, our regression analyses show that the attrition is not systematically correlated with the treatment status at any round of the follow-up surveys.

3.3. Externality

Our sample enterprises are located in the industrial cluster. While the focus on industrial clusters allows us to control for various heterogeneity of entrepreneurs, the geographical proximity is likely to cause externality by knowledge spillover among our samples.[7] Indeed, an entrepreneur knew as many as 31.4 other sample entrepreneurs on average as presented in Table 2. The number further increased to 37.1 after the training programs. As the sample entrepreneurs were excited to learn about newly introduced *Kaizen*, they talked about our *Kaizen* training and what they had learned in the training.

6 One enterprise in Group TC and two in Group CT did not operate at the time of the third follow-up survey and one enterprise in Group TT, three in Group CT and two in Group CC did not operate at the time of the fourth follow-up survey.

7 Knowledge spillover from the treated factory to the non-treated factory owned by the same entrepreneur is documented in Bloom, Eifert et al. (2013). Also, the productivity spillover from the trained worker to the non-trained co-worker is reported in de Grip and Sauermann (2012).

Our sample entrepreneur talked about *Kaizen* with 35.0 other entrepreneurs in the sample, and even those in Group CC talked about *Kaizen* with 23.1 sample entrepreneurs who were either participated in the training or heard about *Kaizen* from the training participants.

The knowledge spillover from the participants to the non-participants is also evident from the changes in the sample entrepreneurs' interest in our training. We measured their interest in our training by collecting information on willingness to pay local currency equivalent of 150USD for our programs.[8] Table 2 shows that the proportion of the sample entrepreneurs who were willing to pay was high even before the training, but it increased to 100 percent after the training. The increase in the willingness to pay among the control group suggests that *Kaizen* gained good reputation not only among the training participants but also among the non-participants.

The presence of the knowledge spillover, or externality, violates the Stable Unit Treatment Value Assumption (SUTVA), an assumption that the control group is not affected by the treatment (Rubin, 1978). In our study site, however, this is not a serious concern for policy purpose when estimating the training impact on the management score. As the non-participants seem to have learned *Kaizen* from the participants, the direction of knowledge spillover is most likely to be positive, and thus, the estimated coefficients indicate conservative training impacts.

8 The question on willingness to pay is prone to reporting bias because of its hypothetical nature. In order to reduce such bias, we followed the lead of Blumenschein et al. (2008). After outlining our training programs, we asked them a hypothetical question, "Would you pay 400,000 Tanzanian shilling to participate in the training program?" which was followed, if the answer was affirmative, by an additional question, "How sure are you about the answer? Are you definitely sure or probably sure?" Based on these set of questions, we constructed a dummy variable that is one if the answer is "definitely sure of the willingness to pay" and zero otherwise.

In contrast, the training impact on business performance may be over-estimated if the training participants stole business from the non-participants, such possibility being pointed out by Bloom, Schankerman et al. (2013). We have two reasons to believe that the degree of market stealing is not substantial in our study context. Firstly, *Kaizen* approach emphasizes cost reduction through elimination of wasted work and material, rather than sales promotion, to increase productivity. Indeed, the training impact on value added is stronger than the impact on sales revenue. Secondly, even if the training participants increased sales, the potential market is much larger than the production capacity of the industrial cluster. Our sample entrepreneurs have unexploited market both inside and outside Dar es Salaam, emerging regional markets in East Africa, or even in other countries beyond Africa. Therefore, the degree of market stealing is likely to be limited and, to the extent that the market stealing is smaller than the knowledge spillover, the estimated coefficients show the conservative training impact.

4. Results

4.1. Descriptive Analyses

In addition to the baseline survey, we collected the data of management score at each of the four follow-up surveys. Although there was no baseline difference between groups as presented in Table 3, Table 4 shows that the management scores were significantly different after the training intervention. Between the two programs, i.e., soon after the classroom training, the scores of Group TT and Group TC were significantly higher than that of Group CC (see *p*-values in columns 5 and 6). Similarly, soon after the on-site training, the score of Group CT became higher than that of Group CC although not significant. Such differences between the treatment groups and control

group were sustained one and half years after the training, suggesting the sustained training impact on management capacity. Comparing the baseline and one and half years after the training, however, the score of the Group CC also sharply increased. Although there may be other factors that influence the management score of our sample entrepreneurs, the substantial increase of management score among group CC can, at least partially, be attributed to the knowledge spillover from the participants to the non-participants.

In the latest follow-up survey conducted three year after the training intervention, the score is lower than that measured one and half years after the intervention while remains higher than the baseline level. Such changes in management score may illustrate that the participants as well as the non-participants became excited about our training programs and started to adopt a number of management practices regardless of their effectiveness or relevance to their business. As time passes, however, they started to select only the useful part of the management practices on the trial-and-error basis. According to our informal interview with a number of sample entrepreneurs, they selectively chose relevant management practices in the longer term and such relevant practices were different from one enterprise to another, depending on the size of business, characteristics of the produced item, and layout of the workshop.

For business performance, we have baseline data—the average of 2008 and 2009—and four data points, that is, 2010, 2011, 2012 and 2013. Although the *p*-values reported in columns 5 to 7 are mostly insignificant, the pattern by groups has changed over time. At the baseline, the enterprises in Group CC had highest business performance but their business performance became lowest in 2012 and 2013. In contrast, the business performance of enterprises in Group TT was lower than that of

Group CC and was at similar level with groups TC and CT at the baseline, but it became highest in 2012 and 2013. In addition, *p*-value reported in column 5 show that the value added of Group TT is statistically (although marginally) higher than that of Group CC in 2013. This may imply that the training, particularly the combination of the two programs, has positive impacts on business performance in a longer span. Table 3 shows the overall declining trend in business performance, particularly from 2011 to 2012. This is partially due to increasing competition with Chinese imports and partially due to the bankruptcy of a major supplier of raw material.

In order to capture the changes in business, we present the data on investment behavior of our sample enterprises. At each follow-up survey, we collected retrospective information on their capital investment. The proportion of sample enterprises that made investment is not large and the amount of investment varied greatly, and thus, we constructed a dummy variable taking one if the enterprise made any investment in each of the periods. We present the data toward the bottom of Table 3. Although the proportion of sample entrepreneurs made investment is not statistically different among Groups in the first two periods, entrepreneurs in Groups TT and TC are significantly more likely to have invested between January 2012 and December 2013. Hence, the different investment behavior in the longer-run, at least partially, explains positive longer-term training impact on business performance.

4.2. Baseline Correlates of Management Score and Business Performance

Before turning to econometric analyses of training impacts, it seems useful to examine how the baseline management score and business performance are correlated with the variables representing the entrepreneurs' characteristics. Appendix Table A2

presents the results of the regressions for this purpose. Bloom and van Reenen (2010) find in a large sample of medium-sized firms both in developed and emerging countries that their measure of management practices is closely associated with the human capital of managers. Consistently, column 1 shows that the management score is closely associated with years of schooling and past business training experience. In addition, the management score is associated with years of operation even after controlling for age and its squared term, because the entrepreneurs improve their management capacity with experience.

Columns 2 and 4 show that business performance, measured by value added and sales revenue, is associated with years of schooling, past training experience and years of operation. While admitting that it is potentially endogenous, we regressed the baseline business performance on the baseline management score in columns 3 and 5. The business performance is significantly associated with the management score and the coefficients of variables representing years of education, past training experience and years of operation become smaller and statistically weaker. This can be interpreted as being that the part of association between business performance and human capital of entrepreneurs is explained by their management capacity. The correlation between our management score and the entrepreneurs' human capital and that between the management score and business performance suggest that our measurement of management capacity is reliable to some extent.

4.3. Econometric Specification

Let y_{it} be an outcome variable, which can be the management score, value added or sale revenue of enterprise i in period t. Our basic specification is written as;

$$y_{it} = \sum_s \alpha_s Z_i T_{st} + \boldsymbol{X_i}\boldsymbol{\gamma} + \upsilon_i + \eta_t + \varepsilon_{it}\,, \qquad (1)$$

where Z_i is a dummy variable indicating whether enterprise i is invited to any of the training programs. T_{st} is a time dummy taking one if $t = s$ and 0 otherwise. $\boldsymbol{X_i}$ is a vector of variables capturing the entrepreneurs' characteristics, which are assumed to be time-invariant.[9] υ_i is the enterprise-level unobserved heterogeneity, η_t is the time trend as well as temporal shock that are common to all the enterprises, and ε_{it} is the error term. Under this specification, α_s represents the impact of training programs at period s, which can be first, second, third or fourth follow-up survey.

We apply two econometric strategies to estimate Equation (1). Firstly, we estimate the intention-to-treat (ITT) effect and the treatment on the treated (TOT) effect. In order to estimate TOT, we use instrument variable approach by instrumenting actual participation status with the random invitation status, following the lead of Imbens and Angrist (1994). Concretely speaking, we replace Z_i with P_i in Equation (1), where P_i is a dummy variable taking one if the enterprise i participated in either (or both) of the training programs, using Z_i as an instrument for P_i. By so doing, we estimate the training impact on compliers, who were induced to participate in the training by the random invitation. The estimated TOT effect is expected to be similar but slightly greater than ITT estimates because the compliance rate for both training programs are high as discussed in Section 2.3.

Secondly, to control for the time-invariant enterprise-level heterogeneity, υ_i, we

9 Precisely speaking, age of entrepreneur and years of operation are time-variant but we treat them as time-invariant by fixing to the values at the baseline survey.

apply ANCOVA model. McKenzie (2012) suggests that this model is preferred to the fixed effect model when a dependent variable has low-autocorrelation and data is available for several points. Specifically, we add the (mean of) baseline value of the dependent variable in the right-hand side in addition to other control variables, $\boldsymbol{X_i}$. In order to take advantage of our panel data, we cluster the error term, ε_{it}, at the enterprise level so that the autocorrelation within a respective enterprise can be controlled.

Next, in order to analyze the impact of each component of the two training programs and their combination, we split Z_i into $Z^k{}_i$, where $k=TT$ for Group TT, $k=TC$ for Group TC, and $k=CT$ for Group CT. Our specification is written as;

$$y_{it} = \sum_s \alpha^{TT}{}_s Z^{TT}{}_i T_{st} + \sum_s \alpha^{TC}{}_s Z^{TC}{}_i T_{st} + \sum_s \alpha^{CT}{}_s Z^{CT}{}_i T_{st} + \boldsymbol{X_i}\boldsymbol{\gamma} + \upsilon_i + \eta_t + \varepsilon_{it}\,. \qquad (2)$$

Similar to the interpretation of Equation (1), $\alpha^k{}_s$ is our parameter of interest, representing the training impact of training k at period s. For TOT estimation of Equation (2), we replace $Z^k{}_i$ with $P^k{}_i$, where $P^k{}_i$ is a dummy variable taking one if the enterprise i complied with the assigned treatment status, and use $Z^k{}_i$ as an instrument for P_i.[10]

4.4. Econometric Analyses

Table 4 presents the estimates of Equation (1) for the management score. Using the baseline management score as a control variable, we have four data points for the regression analyses. s = "soon after CR" and s = "soon after OS" indicate that the data

10 Four entrepreneurs in Group TT, who participated only in the on-site training, are treated as non-participants in the instrument regression because we are mainly interested in the impact of the combination of the two training programs by estimating $\alpha^{TT}{}_s$.

was collected in the first and second follow-up survey, respectively. The coefficients, α_s, show the shorter-run training impact. In order to capture longer-run impact of the training, s = "1.5 years after" and s = "3 years later" indicate that the data was collected in the third and fourth follow-up survey, respectively, and α_s show the longer-term training impact. In all estimations, entrepreneurs' age, its squared, male dummy, Chagga dummy, years of education, past business training experience, former textile employee dummy, and years of operation are controlled although their coefficients are not reported. In addition, in order to remove bias in data collection, we control for enumerator fixed effect in all specification of our econometric analyses.

As expected, the coefficients in column 1, which reports ITT estimates, are slightly smaller than TOT estimates in column 2. The first-stage *F*-statistic reported toward the bottom of Table 4 is large, and thus, our instrument variable strategy is valid. The training impacts are shown to be significant both in the short and longer-run. The coefficient of 3.35 in the first row of column 2—our preferred specification—can be interpreted as that treated entrepreneurs adopt, on average, 3.35 more management practices listed in Appendix Table A1 than those in Group CC at the time of the first follow-up survey. The magnitude of the coefficients are greater in the short run than in long run. This supports our argument that the trained enterprises started to adopt many management practices soon after the training interventions mostly because of their excitement while they gradually select only the useful practices. The coefficients for the interaction terms with "3 years after" are all significant, illustrating the longer-run impacts of our training. The significant coefficients for time dummies reported toward the bottom capture, at least partially, changes in management score among Group CC by knowledge spillover.

In order to analyze which components of management practices were adopted after the training, we disaggregate the management score into five components, namely, sales promotion, record keeping, marketing, quality control and *Kaizen*, and present the results of regressions with the each of the disaggregated score as a dependent variable in columns 3 to 7. Since our instrumental strategy is valid, we only report the TOT estimates. Columns 3 to 7 show that the increase in total management score is not sorely explained by any particular component. Instead, the training has impact evenly among the all components. In other words, some enterprises started to adopt some component of management practices while other enterprises adopted others. This supports our argument that the participants select useful part of management practices for their own business, which differ from one enterprise to another.

Table 5 presents the estimates of Equation (1) for value added and sales revenue. Controlling for the mean values in 2008 and 2009 as the baseline value, we use four data points, i.e., 2010, 2011, 2012 and 2013, for the regression analyses. The coefficient of the interaction of training and "2010 dummy", for instance, captures the immediate impact of the training on business performance, while that of training and "2013 dummy" captures the impact observed three years after the training intervention. In columns 1 to 3, we report the impact on value added and columns 4 to 6 report the impact on sales revenue. We mainly report the results with levels of business performance as outcome variables, while columns 3 and 6 use the log of business performance. Similar to Table 4, the first-stage *F*-statistics is large, and thus, our instrument strategy is valid.

Table 5 shows positive and marginally significant coefficient for value added in 2013 while most of the other coefficients are positive but not significant. Among the insignificant coefficients, *t*-statistics are larger for those interacted with 2012 and 2013

dummies. These observations suggest that the training has no impact on business performance in the short-run but the impact gradually emerged in the longer-run. As Table 3 suggests that the combination of the two training had greater impact on business performance, we estimate Equation (2) to separately analyze the impact of the combination of two training programs. Table 6 presents the results. The coefficient for both training is marginally significant for 2012 and strongly significant for 2013. The magnitude of coefficient is large; the estimated impact of the combination of the two programs on value added in 2013 is 9,626 USD (see column 2), which is about two-thirds of the baseline value added reported toward the bottom of Table 6.

In order to check the robustness of our results, Appendix Table A3 reports the results with narrowly defined participation status of the on-site training. As described in Section 2.3., some enterprises did not receive the consultation until the end of the course as recommended by the training instructors. Although all 54 enterprises in Groups TT and CT accepted our initial consultation, only 33 completed the on-site training and we regard only those 33 enterprises as on-site training participants. Appendix Table A3 presents similar results; the training has no impact on business performance in the short-run but the impact, particularly those of the combination of the two programs, emerged in the longer-run. Appendix Table A4 shows the results using fixed effect model, instead of ANCOVA model. Although the size and significant level of the estimated coefficients are different from ANCOVA estimates, the qualitative results remain the same.

Lastly, Table 7 shows the regressions with a dummy variable representing whether an enterprise made any investment at each data period as a dependent variable. As the data coverage is not mutually exclusive, we report the results of Probit regression run

separately for each period. Table 7 shows that the training participants are no more likely to have made investment shortly after the training intervention. The participants, particularly those participated in the classroom training, however, are significantly more likely to do so between January 2012 and December 2013. This suggests that the training changed investment behavior of the training participants, which results in higher business performance in the longer-term.

5. Conclusion

This paper studies the randomized experiment in Tanzania to assess the short and longer term impacts of managerial training on entrepreneurs' management capacity and business performance. Our findings are in line with the stylized finding in existing studies that training improves management capacity of entrepreneurs. While the impact of training on business performance is mixed in the literature, we find that the impact on business performance gradually emerged in a longer term. Hence, our finding suggests the potential usefulness of management training, or *Kaizen* training in particular, in developing industrial sector in the developing economies. In addition, as we illustrate the presence of rampant spillover of managerial knowledge from the training participants to non-participants, the social impacts of our training is most likely to be greater than our estimates in this paper. To the extent that managerial training has greater social returns than private returns due to positive externality, the governmental intervention to provide such training is warranted.

Our findings also suggest that existing studies of management training intervention evaluate the impacts too early. It takes longer time for trained entrepreneurs to substantially change their business to improve the performance, and thus, researchers

should trace the training participants for longer period of time than in most of the existing studies. In so doing, the construction of panel data covering the long span is critically important.

References

Berge, L.I.O., Bjorvatn, K., Juniwaty, K.S., Tungodden, B., 2012. Business training in Tanzania: From research-driven experiment to local implementation. *Journal of African Economies*, 21(5), 808-827.

Berge, L.I.O., Bjorvatn, K., Tungodden, B., 2014. Human and financial capital for microenterprise development: Evidence from a field and lab experiment. *Management Science* 61, 707-722.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., Roberts, J., 2013. Does management matter? Evidence from India. *Quarterly Journal of Economics*, 128(1), 1-51.

Bloom, N., Genakos, C., Sadun, R., van Reenen, J., 2012. Management practices across firms and countries. *Academy of Management Perspective*, 26(1), 12-33.

Bloom, N., Schankerman, M., van Reenen, J., 2013. Identifying technology spillovers and product market rivalry. *Econometrica* 81, 1347-1393.

Bloom, N., Van Reenen, J., 2010. Why do management practices differ across firms and countries? *Journal of Economic Perspective*, 24(1), 203-224.

Blumenschein, K., Blomquist, G.C., Johannesson, M., Horn, N., Freeman, P., 2008. Eliciting willingness to pay without bias: Evidence from a field experiment. *Economic Journal*, 118(525), 114-137.

Bruhn, M., Karlan, D., Schoar, A., 2010. What capital is missing in developing countries? *American Economic Review: Papers and Proceedings*, 100(2), 629-633.

Bruhn, M., Zia, B., 2013. Stimulating managerial capital in emerging markets: The impact of business training for young entrepreneurs. *Journal of Development Effectiveness*, 5(2), 232-266.

De Grip, A., Sauermann, J., 2012. The effects of training on own and co-worker

productivity: Evidence from a field experiment. *Economic Journal*, 122, 376-399.

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

Drexler, A., Fischer, G., Schoar, A., 2014. Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics*, 6(2), 1-31.

Field, E., Jayachandran, S., Pande, R., 2010. Do traditional institutions constrain female entrepreneurship? A field experiment on business training in India. *American Economic Review*, 100 (2), 125-129.

Higuchi, Y., Nam, V.H., and Sonobe T. 2015. Sustained impacts of *Kaizen* training. *Journal of Economic Behavior and Organization*, 120, 189-206.

Imai, M., 2012. Gemba Kaizen: *A commonsense Approach to Continuous Improvement Strategy*, Second Edition. New York: McGraw-hill.

International Labour Organisation, (ILO) 2003. *ILO/SIYB Global Outreach Report*. Geneva: ILO.

Imbens, G., Angrist, J., 1994. Identification and estimation of local average treatment effects. *Econometrica*, 62, 467-476.

Karlan, D., Knight, R., Udry, C., 2012. Hoping to win, expected to lose: Theory and lessons on micro enterprise development. *NBER Working Paper No. 18325*.

Karlan, D., Valdivia, M., 2011. Teaching entrepreneurship: Impact of business training of microfinance clients and institutions. *Review of Economics and Statistics*, 93(2), 510-527.

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

McKenzie, D., Woodruff, C., 2014. What are we learning from business training and entrepreneurship evaluations around the developing world. *World Bank Research Observer*, 29(1), 48-82.

Morgan, J., Liker, J.K., 2006. *The Toyota Product Development System: Integrating People, Process, and Technology*. London: Productivity Press.

Mottaleb, K.A., Sonobe, T., 2011. An inquiry into the rapid growth of the garment industry in Bangladesh. *Economic Development and Cultural Change*, 60, 67-89.

Pack, H. 1987. *Productivity, Technology, and Industrial Development :A case study in textiles*. New York: Oxford University Press.

Rubin, D.B., 1978. Bayesian inference for causal effects: The role of randomization. *Annals of Statistics*, 6(1), 34-58.

Sonobe, T., Otsuka, K., 2011. *Cluster-Based Industrial Development: A Comparative Study of Asia and Africa*. New York: Palgrave Macmillan.

Sonobe, T., Otsuka, K., 2014. *Cluster-Based Industrial Development: KAIZEN Management for MSE Growth in Developing Countries.* New York: Palgrave Macmillan.

Van Biesebroeck, J., 2003. Productivity dynamics with technology choice: An application to automobile assembly. *Review of Economic Studies*, 70, 167-198.

World Bank, 2012. *World Development Report 2013: Jobs.* Washington DC: World Bank.

Table 1: Balance Check

	(1) Group TT mean	(2) Group TC mean	(3) Group CT mean	(4) Group CC mean	(5) All *p*-value
Panel A: Characteristics of entrepreneurs					
Age (as of the baseline survey)	44.5	44.8	45.2	44.9	0.99
	(9.06)	(7.52)	(9.49)	(7.97)	
Male	0.08	0.17	0.14	0.26	0.33
(yes = 1)	(0.27)	(0.38)	(0.36)	(0.44)	
Chagga	0.15	0.25	0.32	0.29	0.53
(yes = 1)	(0.37)	(0.44)	(0.48)	(0.46)	
Years of education	11.3	10.3	10.6	10.7	0.63
	(2.62)	(2.12)	(2.66)	(2.85)	
Past business training experience	0.73	0.67	0.61	0.58	0.67
(yes = 1)	(0.45)	(0.48)	(0.50)	(0.50)	
Former textile employee	0.15	0.25	0.25	0.19	0.80
(yes = 1)	(0.37)	(0.44)	(0.44)	(0.40)	
Years of operation (as of baseline survey)	11.9	11.8	12.0	10.6	0.78
(yes = 1)	(5.45)	(4.85)	(6.34)	(6.19)	
Member of TANCRAFT	0.42	0.46	0.46	0.58	0.66
(yes = 1)	(0.509	(0.51)	(0.51)	(0.50)	
Member of HOT	0.46	0.46	0.54	0.29	0.28
(yes = 1)	(0.51)	(0.51)	(0.51)	(0.46)	
Member of ADAT	0.12	0.08	0.00	0.13	0.29
(yes = 1)	(0.33)	(0.28)	(0.00)	(0.34)	
Panel B: Baseline management score and business performance					
Management score	11.7	10.3	10.2	10.2	0.30
[0-27]	(3.53)	(2.39)	(3.88)	(3.34)	
Real value added	14,473	13,551	12,895	19,441	0.52
[USD]	(10,964)	(12,171)	(13,916)	(28,744)	
Real sales revenue	23,328	23,130	21,235	31,483	0.58
[USD]	(15,784)	(20,723)	(20,352)	(49,171)	
Number of workers	5.5	4.9	5.3	4.2	0.79
	(4.63)	(5.44)	(6.44)	(3.32)	
No. enterprises in the Group	26	24	28	31	109
No. participated in the classroom training	22	23	0	0	45
No. participated in the on-site training	26	0	28	0	54

Notes: Numbers in parenthesis are standard deviations. *P*-values are from the *F*-test concerning the null hypothesis that the mean values are the same for all the four groups. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Baseline real value added, real sales revenue and number of workers are those of the average of 2008 and 2009. Sales revenue and value added are presented in PPP-adjusted USD using "PPP conversion factor, GDP (LCU per international $)", available at World Bank DATABANK.

Table 2: Knowledge Spillover

	(1) Group TT mean	(2) Group TC mean	(3) Group CT mean	(4) Group CC mean	(5) Total mean
# of sample entrepreneurs					
known before the training programs	36.1	40.3	29.4	22.1	31.4
	(21.51)	(13.27)	(21.35)	(13.30)	(18.85)
known soon after the training programs	45.1	48.4	33.5	24.5	37.1
	(19.47)	(13.08)	(21.83)	(14.59)	(19.84)
talked about *Kaizen* with	42.2	46.0	31.6	23.1	35.0
	(17.50)	(13.24)	(22.21)	(13.69)	(19.06)
Willingness to pay (yes = 1)					
before the training programs	0.65	0.75	0.64	0.68	0.67
	(0.49)	(0.44)	(0.49)	(0.48)	(0.47)
soon after the classroom training	0.88	0.96	0.86	0.77	(0.86)
	(0.33)	(0.20)	(0.36)	(0.43)	(0.35)
soon after the on-site training	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)
No. enterprises in the Group	26	24	28	31	109

Notes: Numbers in parenthesis are standard deviations.

Table 3: Changes in Management Score and Business Performance by Group

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Group TT	Group TC	Group CT	Group CC	TT v.s. CC	TC v.s. CC	CT v.s. CC
	Mean	mean	mean	mean	*p*-value	*p*-value	*p*-value
Management score [0-27]							
between the two programs	17.8	16.9	11.3	10.7	0.00***	0.00***	0.60
	[18]	[16.5]	[10.5]	[9]			
soon after the training programs	17.7	16.8	14.5	12.6	0.00***	0.00***	0.14
	[17.5]	[15.5]	[14.5]	[12]			
1.5 years after the programs	20.3	19.7	19.8	17.4	0.00***	0.02**	0.01***
	[22]	[20]	[19]	[18]			
3 years after the programs	16.4	16.9	15.6	13.0	0.00***	0.00***	0.02**
	[17]	[16.5]	[15]	[12]			
Value added [USD]							
in 2010	24,623	20,750	14,789	24,364	0.97	0.57	0.09*
	[20,248]	[15,730]	[9,432]	[18,993]			
in 2011	18,092	23,667	20,909	22,605	0.44	0.88	0.77
	[14,801]	[13,549]	[14,718]	[14,669]			
in 2012	17,380	12,059	16,445	12,574	0.23	0.88	0.51
	[12,544]	[9,106]	[9,069]	[9,087]			
in 2013	18,914	12,460	13,275	12,535	0.08*	0.98	0.84
	[13,953]	[9,390]	[7,230]	[7,316]			
Sales revenue [USD]							
in 2010	34,527	33,337	24,204	39,484	0.62	0.57	0.11
	[29,881]	[25,101]	[17,623]	[25,101]			
in 2011	33,459	35,727	40,579	38,641	0.59	0.79	0.85
	[29,460]	[27,385]	[36,260]	[24,881]			
in 2012	32,379	24,453	31,515	25,132	0.25	0.90	0.53
	[26,390]	[19,353]	[21,112]	[17,074]			
in 2013	35,786	24,528	26,654	25,818	0.15	0.82	0.90
	[28,191]	[21,143]	[18,241]	[19,899]			
Investment (yes = 1)							
Apr. 2010- Apr. 2011	0.42	0.21	0.14	0.23	0.11	0.88	0.42
	(0.50)	(0.41)	(0.36)	(0.43)			
Jan. 2011- Sep. 2012	0.23	0.17	0.27	0.26	0.82	0.47	0.93
	(0.43)	(0.39)	(0.45)	(0.44)			
Jan. 2012- Dec. 2013	0.36	0.33	0.24	0.10	0.02**	0.04**	0.19
	(0.49)	(0.48)	(0.44)	(0.31)			

Notes: Numbers in bracket are medians while numbers in parenthesis are standard deviations. *P*-values are from the *t*-test concerning the null hypothesis that the mean values are the same among the two groups. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Sales revenue and value added are presented in PPP-adjusted USD.

Table 4: Impact on Management Score (ANCOVA)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Total Management Score		Sales Promotion	Record Keeping	Marketing	Quality control	*Kaizen*
	ITT	TOT	TOT	TOT	TOT	TOT	TOT
Any training	3.29***	3.35***	0.18	0.51**	1.05***	0.43**	1.11***
* Soon after CR	(4.21)	(4.43)	(0.76)	(2.35)	(4.56)	(2.32)	(3.14)
Any training	2.79***	2.85***	0.46*	0.26	0.59**	0.28*	1.21***
* Soon after OS	(3.40)	(3.56)	(1.81)	(1.28)	(2.20)	(1.84)	(2.84)
Any training	1.16*	1.18*	0.071	-0.019	0.18	0.11	0.75*
* 1.5 years after	(1.77)	(1.90)	(0.26)	(-0.14)	(1.31)	(0.60)	(1.86)
Any training	2.55***	2.61***	0.46*	0.74***	0.29	0.17	0.93**
* 3 years after	(3.48)	(3.60)	(1.86)	(3.79)	(1.54)	(0.92)	(2.23)
Time dummy	1.34	1.33	-0.21	0.44**	0.77**	0.13	0.18
(Soon after OS)	(1.57)	(1.62)	(-1.10)	(2.09)	(2.53)	(0.92)	(0.41)
Time dummy	5.56***	5.54***	0.66**	1.08***	1.80***	0.29	1.68***
(1.5 years after)	(6.72)	(6.97)	(2.54)	(4.42)	(8.47)	(1.51)	(3.31)
Time dummy	2.98***	2.94***	-0.67**	0.72***	2.34***	-0.069	0.76
(3 years after)	(2.99)	(3.03)	(-2.55)	(3.19)	(9.78)	(-0.28)	(1.25)
No. observations	427	427	427	427	427	427	427
No. enterprises	109	109	109	109	109	109	109
First-stage *F*-statistics		2742.8	2750.4	2722.1	2740.7	2744.4	2734.3
Baseline mean	10.6	10.6	1.1	2.6	1.4	0.9	4.6

Notes: Numbers in parentheses are *t*-statistics based on standard errors clustered at the enterprise level. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. For the intention-to-treat effects, the reported is the coefficients of the dummy variable taking one if the enterprise was assigned to any of the training programs. For the treatment effects on the treated, the reported is the coefficients of the dummy variable taking one if the enterprise participated in any of the training programs. In all estimations, entrepreneurs' age, its squared, male dummy, Chagga dummy, years of education, past training experience, former textile employee dummy, years of operation and association dummies, the baseline values for each dependent variable, and enumerator fixed effect are controlled although the coefficients are not reported.

Table 5: Impact on Business Performance (ANCOVA)

	(1)	(2)	(3)	(4)	(5)	(6)
	Value added			Sales revenue		
	Level		Log	Level		Log
	ITT	TOT	TOT	ITT	TOT	TOT
Any training	-3008.3	-3029.7	-0.063	-5175.0	-5213.7	-0.038
* 2010	(-0.71)	(-0.73)	(-0.40)	(-0.76)	(-0.78)	(-0.24)
Any training	302.0	329.8	0.036	2737.5	2783.7	0.12
* 2011	(0.08)	(0.09)	(0.20)	(0.47)	(0.49)	(0.72)
Any training	4587.1	4675.6	0.23	8178.5	8333.3	0.15
* 2012	(1.44)	(1.49)	(1.42)	(1.51)	(1.56)	(1.07)
Any training	4327.7*	4409.8*	0.21	7239.4	7376.3	0.15
* 2013	(1.74)	(1.79)	(1.18)	(1.55)	(1.60)	(0.97)
Time dummy	196.6	218.1	0.050	787.8	825.4	0.096
(2011)	(0.04)	(0.05)	(0.28)	(0.13)	(0.14)	(0.60)
Time dummy	-5139.5	-5202.3	-0.37*	-8634.3	-8743.8	-0.19
(2012)	(-1.23)	(-1.29)	(-1.85)	(-1.18)	(-1.24)	(-1.09)
Time dummy	-5179.2	-5242.0	-0.29	-7948.6	-8058.1	-0.14
(2013)	(-1.21)	(-1.27)	(-1.30)	(-1.05)	(-1.10)	(-0.74)
No. observations	418	418	418	418	418	418
No. enterprises	109	109	109	109	109	109
First-stage *F*-statistics		2567.9	2613.3		2580.8	2618.2
Baseline mean	15277.5	15277.5	9.07	25066.0	25066.0	9.63

Notes: Numbers in parentheses are *t*-statistics based on standard errors clustered at the enterprise level. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. For the intention-to-treat effects, the reported is the coefficients of the dummy variable taking one if the enterprise was invited to any of the training programs. For the treatment effects on the treated, the reported is the coefficients of the dummy variable taking one if the enterprise participated in any of the training programs. In all estimations, entrepreneurs' age, its squared, male dummy, Chagga dummy, years of education, past training experience, former textile employee dummy, years of operation and association dummies, the baseline values for each dependent variable, and enumerator fixed effect are controlled although the coefficients are not reported.

Table 6: Impact on Business Performance (ANCOVA)

	(1)	(2)	(3)	(4)	(5)	(6)
	Value added			Sales revenue		
	Level		Log	Level		Log
	ITT	TOT	TOT	ITT	TOT	TOT
Both training	1067.3	1191.9	0.14	-1061.9	-1335.1	0.11
* 2010	(0.22)	(0.22)	(0.66)	(-0.15)	(-0.16)	(0.51)
Both training	-1468.5	-1814.9	-0.039	1484.1	1691.3	0.13
* 2011	(-0.35)	(-0.38)	(-0.16)	(0.25)	(0.25)	(0.58)
Both training	6225.3*	7471.4*	0.39*	11059.1*	13288.9*	0.29
* 2012	(1.66)	(1.70)	(1.65)	(1.87)	(1.93)	(1.27)
Both training	8144.5***	9626.3***	0.39	14300.8**	16915.7**	0.31
* 2013	(2.63)	(2.65)	(1.49)	(2.43)	(2.49)	(1.29)
CR training	-1867.9	-1964.7	0.063	-1645.5	-1751.2	0.091
* 2010	(-0.37)	(-0.39)	(0.34)	(-0.21)	(-0.22)	(0.50)
CR training	2885.3	2857.1	0.070	1199.0	1174.2	0.049
* 2011	(0.54)	(0.56)	(0.33)	(0.16)	(0.17)	(0.25)
CR training	1692.8	1745.7	0.16	3622.0	3748.8	0.12
* 2012	(0.55)	(0.57)	(0.87)	(0.70)	(0.73)	(0.80)
CR training	2131.6	2198.6	0.20	3089.1	3184.2	0.11
* 2013	(0.78)	(0.80)	(1.12)	(0.60)	(0.62)	(0.66)
OS training	-7200.1	-7237.4*	-0.33*	-10980.3	-11043.7	-0.25
* 2010	(-1.62)	(-1.71)	(-1.74)	(-1.50)	(-1.58)	(-1.35)
OS training	-157.5	-190.9	0.064	5828.3	5771.2	0.20
* 2011	(-0.03)	(-0.04)	(0.26)	(0.76)	(0.79)	(0.92)
OS training	5802.6	5785.8	0.19	9961.8	9941.6	0.100
* 2012	(1.12)	(1.17)	(0.93)	(1.08)	(1.13)	(0.51)
OS training	2671.9	2655.0	0.072	4415.3	4395.1	0.064
* 2013	(0.80)	(0.83)	(0.34)	(0.72)	(0.75)	(0.32)
Time dummy	-67.4	-28.8	0.037	493.1	462.9	0.083
(2011)	(-0.01)	(-0.01)	(0.20)	(0.08)	(0.08)	(0.51)
Time dummy	-5643.3	-5784.3	-0.41**	-9392.0	-9624.2	-0.22
(2012)	(-1.35)	(-1.44)	(-2.03)	(-1.28)	(-1.36)	(-1.27)
Time dummy	-5683.0	-5824.0	-0.33	-8706.3	-8938.5	-0.17
(2013)	(-1.32)	(-1.41)	(-1.48)	(-1.14)	(-1.22)	(-0.92)
No. observations	418	418	418	418	418	418
No. enterprises	109	109	109	109	109	109
First-stage *F*-statistics		128.9	132.1		129.0	132.2
Baseline mean	15277.5	15277.5	9.07	25066.0	25066.0	9.63

Notes: Numbers in parentheses are *t*-statistics based on standard errors clustered at the enterprise level. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. For the intention-to-treat effects, the reported is the coefficients of the dummy variable taking one if the enterprise was assigned to each of the treatment groups. For the treatment effects on the treated, the reported is the coefficients of the dummy variable taking one if the enterprise complied with the assigned treatment. In all estimations, entrepreneurs' age, its squared, male dummy, Chagga dummy, years of education, past training experience, former textile employee dummy, years of operation and association dummies, the baseline values for each dependent variable, and enumerator fixed effect are controlled although the coefficients are not reported.

Table 7: Impact on Capital Investment (Probit)

Panel A: Impact of training aggregated

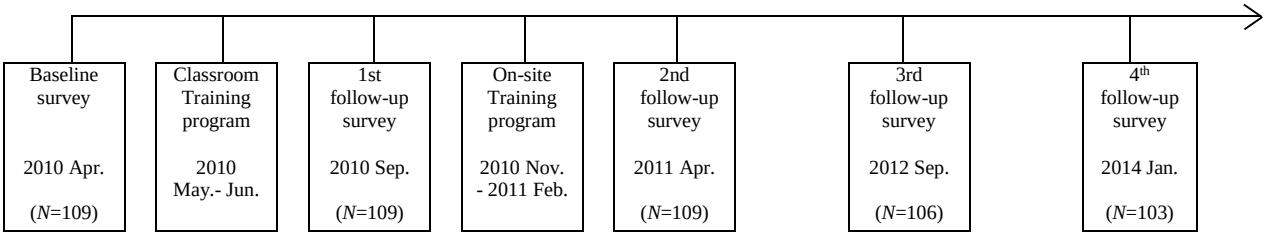
	(1)	(2)	(3)	(4)
	=1 if made any investment			
	Apr. 2010-Sep. 2010 ITT	Oct. 2010-Apr. 2011 ITT	Jan. 2011-Sep. 2012 ITT	Jan. 2012-Dec. 2013 ITT
Any training	0.49	-0.0060	-0.18	0.79*
	(1.13)	(-0.02)	(-0.51)	(1.93)
No. observations	109	106	100	103
No. enterprises	109	109	106	103

Panel B: Impact of each component of training

	(1)	(2)	(3)	(4)
	=1 if made any investment			
	Apr. 2010-Sep. 2010 ITT	Oct. 2010-Apr. 2011 ITT	Jan. 2011-Sep. 2012 ITT	Jan. 2012-Dec. 2013 ITT
Both training	0.54	0.56	-0.018	0.94*
	(1.03)	(1.28)	(-0.04)	(1.92)
CR training	0.52	-0.35	-0.33	0.87*
	(1.07)	(-0.65)	(-0.78)	(1.82)
OS training	0.41	-0.34	-0.16	0.55
	(0.81)	(-0.83)	(-0.39)	(1.20)
No. observations	109	106	100	103
No. enterprises	109	109	106	103

Notes: Numbers in parentheses are *t*-statistics based on standard errors robust to heteroscedasticity. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. In all estimations, entrepreneurs' age, its squared, male dummy, Chagga dummy, years of education, past training experience, former textile employee dummy, years of operation and association dummies, the baseline values for each dependent variable, and enumerator fixed effect are controlled although the coefficients are not reported.

Figure 1: Timeline



Appendix Table A1: Components of Management Score and Baseline Adoption Rates (%)

Sales promotion	
The enterprise had any expenditure for advertisement including newspaper, radio and internet in the last 3 months.	10
The enterprise has any signboards in front of the workshop.	39
The enterprise distributes complimentary cards or calendar?	27
The enterprise issues invoices or receipts with workshop's name or phone number.	36
Record keeping	
The enterprise preserves business documents (e.g., receipts or invoices) when making a purchase.	48
The enterprise separates business and household expenses.	62
The enterprise keeps record of sales.	84
The enterprise keeps record of material purchase.	70
Marketing	
The entrepreneur can clearly describe the characteristics of their customers.	42
The entrepreneur can clearly describe the strength of their enterprise compared with neighboring enterprises.	24
The entrepreneur has clear sales target or profit target in this year.	45
The entrepreneur has clear plan for growth of the enterprise in five years from now.	28
Quality control	
The enterprise assigns any workers to inspect the quality of the products before sales.	10
The enterprise keeps records of quality defects.	22
The enterprise records customers' complaints about the products sold.	45
The enterprise instructs the worker the way of preventing the defect.	9
Kaizen	
The enterprise has a designated place for all tools.	34
The enterprise has labels in the storage of tools so that workers can easily find them.	3
The enterprise has a designated place for raw material storage.	76
The enterprise separately stores raw materials from the scrap.	75
The enterprise has no scrap cloths around the floor.	13
The enterprise daily removes scraps and cleans the floor of the workplace.	83
The enterprise does machine maintenance at least once a week.	29
The enterprise regularly holds a meeting in which all the production workers participate.	28
The enterprise has a designated area for all the production activities within the workshop.	29
The enterprise has Do you have a flowchart indicating the sequence of activities followed in the production process.	8
The entrepreneur completely knows the sequence and duration of each of the production activities.	82

Appendix Table A2: Baseline Correlates of Management Score and Business Performance (OLS)

	(1)	(2)	(3)	(4)	(5)
	Baseline management score	Baseline value added		Baseline Sales revenue	
Age	-0.65	-2859.3	-1805.3	-5894.7	-4282.3
	(-1.22)	(-1.17)	(-0.91)	(-1.51)	(-1.25)
Age squared	0.0059	29.9	20.3	63.9	49.2
	(1.04)	(1.18)	(0.96)	(1.51)	(1.30)
Male	-1.40*	2063.0	4329.9	10444.2	13912.0
(yes = 1)	(-1.69)	(0.29)	(0.66)	(0.74)	(1.01)
Chagga	0.64	4314.6	3280.2	5524.1	3941.9
(yes = 1)	(0.90)	(1.28)	(0.95)	(1.14)	(0.80)
Years of education	0.31***	1053.7**	551.4	1544.1**	775.6
	(2.68)	(2.16)	(1.06)	(2.01)	(0.94)
Past business training experience	1.41*	11127.2**	8844.2*	20052.7**	16560.4**
(yes = 1)	(1.89)	(2.03)	(1.90)	(2.28)	(2.18)
Former textile employee	-0.80	3168.7	4474.6	4655.8	6653.5
(yes = 1)	(-0.83)	(0.54)	(0.77)	(0.56)	(0.82)
Years of operation	0.13**	530.8*	326.9	970.6**	658.7
	(2.24)	(1.82)	(1.03)	(2.02)	(1.33)
Baseline management score			1623.1**		2482.9**
			(2.51)		(2.56)
No. observations	109	109	109	109	109

Notes: Numbers in parentheses are *t*-statistics based on standard errors robust to heteroscedasticity. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. In all estimations, association dummies and enumerator fixed effect are controlled although the coefficients are not reported.

Appendix Table A3: Robustness Check with Narrowly Defined On-site Training Participation

Panel A: Impact of training aggregated

	(1)	(2)	(3)	(4)	(5)	(6)
	Value added			Sales revenue		
	Level		Log	Level		Log
	ITT	TOT	TOT	ITT	TOT	TOT
Any training	-3008.3	-6010.2	-0.079	-5175.0	-10011.4	-0.038
* 2010	(-0.71)	(-0.62)	(-0.21)	(-0.76)	(-0.64)	(-0.10)
Any training	302.0	965.1	0.24	2737.5	19261.9	0.40
* 2011	(0.08)	(0.26)	(1.23)	(0.47)	(1.58)	(1.01)
Any training	4587.1	11892.6	0.60	8178.5	21415.2	0.41
* 2012	(1.44)	(1.50)	(1.48)	(1.51)	(1.57)	(1.11)
Any training	4327.7*	11321.0*	0.55	7239.4	6746.9	0.31*
* 2013	(1.74)	(1.74)	(1.23)	(1.55)	(1.27)	(1.85)
Time dummy	196.6	364.5	0.020	787.8	1185.4	0.10
(2011)	(0.04)	(0.08)	(0.12)	(0.13)	(0.18)	(0.67)
Time dummy	-5139.5	-6252.1	-0.41*	-8634.3	-10644.9	-0.22
(2012)	(-1.23)	(-1.36)	(-1.86)	(-1.18)	(-1.33)	(-1.14)
Time dummy	-5179.2	-6291.8	-0.33	-7948.6	-9959.2	-0.17
(2013)	(-1.21)	(-1.35)	(-1.36)	(-1.05)	(-1.21)	(-0.82)
No. observations	418	418	418	418	418	418
No. enterprises	109	109	109	109	109	109
First-stage *F*-statistics		20.0	20.1		19.8	20.2
Baseline mean	15277.5	15277.5	9.07	25066.0	25066.0	9.63

Notes: Numbers in parentheses are *t*-statistics based on standard errors clustered at the enterprise level. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. For the intention-to-treat effects, the reported is the coefficients of the dummy variable taking one if the enterprise was invited to any of the training programs. For the treatment effects on the treated, the reported is the coefficients of the dummy variable taking one if the enterprise participated in any of the training programs. In all estimations, entrepreneurs' age, its squared, male dummy, Chagga dummy, years of education, past training experience, former textile employee dummy, years of operation and association dummies, the baseline values for each dependent variable, and enumerator fixed effect are controlled although the coefficients are not reported.

Panel B: Impact of each component of training

	(1)	(2)	(3)	(4)	(5)	(6)
	Value added			Sales revenue		
	Level		Log	Level		Log
	ITT	TOT	TOT	ITT	TOT	TOT
Both training	1067.3	2326.2	0.24	-1061.9	-887.7	0.19
* 2010	(0.22)	(0.27)	(0.70)	(-0.15)	(-0.07)	(0.56)
Both training	-1468.5	-2139.2	-0.035	1484.1	3608.4	0.21
* 2011	(-0.35)	(-0.30)	(-0.09)	(0.25)	(0.34)	(0.62)
Both training	6225.3*	12401.6	0.63	11059.1*	22338.8*	0.48
* 2012	(1.66)	(1.64)	(1.64)	(1.87)	(1.83)	(1.27)
Both training	8144.5***	15983.7**	0.65	14300.8**	28415.7**	0.52
* 2013	(2.63)	(2.36)	(1.51)	(2.43)	(2.20)	(1.30)
CR training	-1867.9	-1630.1	0.078	-1645.5	-1112.0	0.10
* 2010	(-0.37)	(-0.33)	(0.42)	(-0.21)	(-0.14)	(0.56)
CR training	2885.3	3138.7	0.083	1199.0	1729.7	0.060
* 2011	(0.54)	(0.62)	(0.39)	(0.16)	(0.24)	(0.31)
CR training	1692.8	2015.9	0.17	3622.0	4290.7	0.13
* 2012	(0.55)	(0.63)	(0.91)	(0.70)	(0.79)	(0.84)
CR training	2131.6	2442.8	0.21	3089.1	3667.4	0.12
* 2013	(0.78)	(0.87)	(1.14)	(0.60)	(0.69)	(0.69)
OS training	-7200.1	-11775.3	-0.53	-10980.3	-17756.2	-0.41
* 2010	(-1.62)	(-1.60)	(-1.64)	(-1.50)	(-1.46)	(-1.25)
OS training	-157.5	-107.7	0.11	5828.3	9789.0	0.33
* 2011	(-0.03)	(-0.01)	(0.27)	(0.76)	(0.84)	(0.95)
OS training	5802.6	9162.4	0.30	9961.8	15946.8	0.16
* 2012	(1.12)	(1.16)	(0.92)	(1.08)	(1.13)	(0.51)
OS training	2671.9	4270.6	0.12	4415.3	7280.4	0.10
* 2013	(0.80)	(0.81)	(0.33)	(0.72)	(0.75)	(0.32)
Time dummy	-67.4	77.8	0.046	493.1	697.5	0.094
(2011)	(-0.01)	(0.02)	(0.26)	(0.08)	(0.11)	(0.58)
Time dummy	-5643.3	-6222.2	-0.42**	-9392.0	-10459.4	-0.23
(2012)	(-1.35)	(-1.49)	(-2.00)	(-1.28)	(-1.42)	(-1.26)
Time dummy	-5683.0	-6262.0	-0.34	-8706.3	-9773.7	-0.18
(2013)	(-1.32)	(-1.47)	(-1.47)	(-1.14)	(-1.28)	(-0.92)
No. observations	418	418	418	418	418	418
No. enterprises	109	109	109	109	109	109
First-stage *F*-statistics		13.8	14.0		13.6	14.0
Baseline mean	15277.5	15277.5	9.07	25066.0	25066.0	9.63

Notes: Numbers in parentheses are *t*-statistics based on standard errors clustered at the enterprise level. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. For the intention-to-treat effects, the reported is the coefficients of the dummy variable taking one if the enterprise was assigned to each of the treatment groups. For the treatment effects on the treated, the reported is the coefficients of the dummy variable taking one if the enterprise complied with the assigned treatment. In all estimations, entrepreneurs' age, its squared, male dummy, Chagga dummy, years of education, past training experience, former textile employee dummy, years of operation and association dummies, the baseline values for each dependent variable, and enumerator fixed effect are controlled although the coefficients are not reported.

Appendix Table A4: Robustness Check with Fixed Effect Model

Panel A: Impact of training aggregated

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Management Score		Value added		Sales revenue	
	ITT	TOT	ITT	TOT	ITT	TOT
Any training	3.45***	3.48***	879.4	885.9	375.4	372.9
* Soon after CR/ 2010	(4.72)	(4.72)	(0.18)	(0.18)	(0.05)	(0.05)
Any training	2.92***	2.96***	3731.0	3775.3	7405.4	7476.2
* Soon after OS/ 2011	(3.85)	(3.85)	(0.85)	(0.86)	(1.12)	(1.13)
Any training	1.30*	1.32*	7210.7	7308.4	11937.7	12100.8
* 1.5 years after/ 2012	(1.77)	(1.79)	(1.39)	(1.39)	(1.37)	(1.37)
Any training	2.70***	2.73***	6973.1	7064.5	11258.5	11407.1
* 3 years after/ 2013	(3.39)	(3.38)	(1.43)	(1.43)	(1.41)	(1.41)
Time dummy	-0.67	-0.62	8236.5*	8378.4**	6582.1	6818.9
(Soon after CR/ 2010)	(-0.98)	(-0.91)	(1.93)	(1.98)	(0.95)	(0.99)
Time dummy	0.70	0.74	8228.7*	8411.7*	6469.7	6780.6
(Soon after OS/ 2011)	(0.92)	(0.98)	(1.75)	(1.80)	(0.87)	(0.91)
Time dummy	4.71***	4.74***	5430.5	5494.6	1873.1	1974.9
(1.5 years after/ 2012)	(4.91)	(4.94)	(1.08)	(1.09)	(0.22)	(0.23)
Time dummy	2.09**	2.13**	5390.8	5454.8	2558.7	2660.6
(3 years after/ 2013)	(2.10)	(2.15)	(1.07)	(1.09)	(0.31)	(0.32)
No. observations	536	536	527	527	527	527
No. enterprises	109	109	109	109	109	109

Panel B: Impact of each component of training

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Management Score		Value added		Sales revenue	
	ITT	TOT	ITT	TOT	ITT	TOT
Both training	5.04***	5.97***	3598.4	4238.7	2940.3	3484.2
* Soon after CR/ 2010	(4.54)	(4.58)	(0.69)	(0.69)	(0.37)	(0.38)
Both training	4.04***	4.76***	301.8	244.7	3685.5	4226.7
* Soon after OS/ 2011	(4.22)	(4.31)	(0.06)	(0.04)	(0.53)	(0.51)
Both training	0.85	0.97	8081.3	9643.6	13915.0	16617.0
* 1.5 years after/ 2012	(0.85)	(0.81)	(1.45)	(1.43)	(1.49)	(1.48)
Both training	2.06**	2.41**	9830.5*	11647.7*	17295.0**	20499.7*
* 3 years after/ 2013	(2.15)	(2.09)	(1.92)	(1.88)	(1.98)	(1.96)
CR training	5.58***	5.83***	805.6	840.4	2312.3	2394.2
* Soon after CR/ 2010	(5.92)	(6.46)	(0.16)	(0.16)	(0.29)	(0.29)
CR training	3.59***	3.76***	5420.8	5449.6	4940.6	5018.2
* Soon after OS 2011	(3.83)	(3.92)	(1.04)	(1.03)	(0.68)	(0.68)
CR training	1.50*	1.63*	3166.9	3280.2	5991.8	6223.6
* 1.5 years after/ 2012	(1.78)	(1.93)	(0.63)	(0.63)	(0.69)	(0.69)
CR training	3.90***	4.05***	4068.8	4214.2	6366.0	6598.8
* 3 years after/ 2013	(3.92)	(3.83)	(0.83)	(0.83)	(0.75)	(0.74)
OS training	0.030	0.016	-1212.0	-1250.7	-3017.9	-3106.4
* Soon after CR/ 2010	(0.04)	(0.02)	(-0.22)	(-0.22)	(-0.36)	(-0.38)
OS training	1.56	1.54	5135.3	5102.1	12788.2	12704.2
* Soon after OS/ 2011	(1.59)	(1.58)	(0.85)	(0.85)	(1.47)	(1.46)
OS training	1.60*	1.61*	9941.1	9937.4	15224.8	15205.5
* 1.5 years after/ 2012	(1.79)	(1.79)	(1.48)	(1.48)	(1.36)	(1.35)
OS training	2.28**	2.27**	6810.3	6806.6	9678.3	9659.0
* 3 years after/ 2013	(2.31)	(2.29)	(1.15)	(1.15)	(1.06)	(1.05)
Time dummy	-1.27*	-1.23*	7720.3*	7704.9*	6046.3	5896.9
(Soon after CR/ 2010)	(-1.87)	(-1.80)	(1.85)	(1.85)	(0.88)	(0.86)
Time dummy	0.054	0.11	7929.6*	8003.3*	6368.1	6250.6
(Soon after OS/ 2011)	(0.07)	(0.15)	(1.75)	(1.74)	(0.87)	(0.85)
Time dummy	4.01***	4.11***	4489.2	4314.1	540.9	137.8
(1.5 years after/ 2012)	(4.29)	(4.32)	(0.91)	(0.87)	(0.06)	(0.02)
Time dummy	1.00	1.00	4449.4	4274.4	1226.5	823.4
(3 years after/ 2013)	(1.04)	(1.02)	(0.91)	(0.86)	(0.15)	(0.10)
No. observations	536	536	527	527	527	527
No. enterprises	109	109	109	109	109	109

Notes: Numbers in parentheses are *t*-statistics based on standard errors clustered at the enterprise level. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. For the intention-to-treat effects, the reported is the coefficients of the dummy variable taking one if the enterprise was assigned to each of the training programs. For the treatment effects on the treated, the reported is the coefficients of the dummy variable taking one if the enterprise complied with the assigned treatment. In all estimations, enumerator fixed effect is controlled although the coefficients are not reported.