Mentors or Teachers?

Microenterprise Training in Kenya*

Wyatt Brooks
University of Notre Dame

Kevin Donovan
University of Notre Dame

Terence R. Johnson
University of Notre Dame

January 2017

Abstract

We use a randomized controlled trial to demonstrate that inexperienced female microenterprise owners in a Kenyan slum benefit from mentorship by an experienced entrepreneur in the same community. Using seven rounds of surveys, we find that mentorship increased profits by 20% relative to control. We conduct a formal business education intervention, which has no effect on profits despite changes in business practice. Our results demonstrate that missing information is a salient barrier to profitability, but the type of information matters: localized, specific information (like finding local suppliers) increases profit while abstract, general information (like how to do bookkeeping) does not.

JEL Classification Codes: O12, L26, M53

*Thanks to conference and seminar participants at Arizona State, Notre Dame, the World Bank, York University, the Chicago Fed M&M Development Workshop, Development Day (Michigan), the Econometric Society NASM (Penn), IGC Growth Week at LSE, NEUDC (MIT), and the World Bank ABCDE for comments and insights, especially Nick Bloom, Paco Buera, Bill Evans, Xavi Giné, Joe Kaboski, Molly Lipscomb, David McKenzie, Ezra Oberfield, and Chris Woodruff. This paper also benefitted substantially from discussions by Paulo Bastos and Jing Cai. We thank the Ford Family Program and the Helen Kellogg Institute for International Studies for financial support, including Bob Dowd and Dennis Haraszko for their help coordinating the project, and Lawrence Itela, Jackline Aridi, and Maurice Sikenyi for their excellent work managing the project in Dandora. *Contact Info:* 434 Flanner Hall, Notre Dame, IN 46556. Brooks: wbrooks@nd.edu; Donovan: kdonova6@nd.edu; Johnson: tjohns20@nd.edu

1 Introduction

In urban areas of developing countries, microenterprises are ubiquitous with a large fraction of the workforce engaged in entrepreneurship.[1] Many of these very small firms generate low profit and have few or no employees. Understanding the reason that these firms have low profits, and why they generate so little employment, is key for designing policy to improve the welfare of the urban poor. One possibility is that microenterprise owners lack what Bloom and Van Reenen (2007) and Bruhn, Karlan and Schoar (2010) refer to as *managerial capital*. That is, they lack the skill or know-how to run a business, which limits their profitability and operational scale.

If this lack of information or managerial capital is a serious impediment to growth, formal classes offer an appealing solution: a well-designed business training curriculum may increase business profit, and may be scalable and easily portable between differing environments. However, as summarized by McKenzie and Woodruff (2014) and Blattman and Ralston (2015), where formal business classes have been offered to entrepreneurs, they have had limited impact. One possible interpretation of these results is that, in fact, a lack of information or business knowledge is not an explanation for the lack of profitability in these small firms.

In this paper, we both rationalize the findings of the previous literature and provide evidence that a lack of information is a salient limitation for the operators of microenterprises. Our insight is that a distinction must be made between the abstract principles commonly taught in formal business classes (such as bookkeeping, marketing, managing costs, and selecting product lines) and market-specific, localized information (such as what location has highest demand, what supplier sells at lowest cost, or what products can be sold for the greatest profit). We show that the local expertise provided by a mentor can significantly benefit young entrepreneurs, even when formal classes fail to do so.

We use a randomized controlled trial to evaluate the effects of increasing both information types. We randomly assign young and inexperienced microenterprise owners into three groups: the first group receives one-to-one matching with a mentor from the same

[1] According to Gollin (2008), developing countries have substantially higher rates of entrepreneurship than rich countries. The proportion of workers that identify as entrepreneurs is 75.3% in Nigeria, 74.5% in Bangladesh, and 50.4% in Guatemala. This compares to 8.2% in the United States and 10.6% in the United Kingdom.

community and that operates in the same industry; the second group receives formal business classes; and the third group receives neither. Mentors are drawn from the set of successful business owners within the same community.[2] Rather than direct the content of mentor-mentee interactions, we allow the mentors to relay whatever information they believe will be beneficial to their mentees. Our finding is that the information that mentors typically provide is of a highly localized nature, such as information on suppliers.[3] The business classes are conducted by instructors from a local university, and cover general business topics such as accounting, marketing, cost control and business planning.

We find that mentorship is effective at increasing profit among young microenterprises. Over the twelve months following treatment, weekly profits are on average 20 percent higher among the mentor treatment than the control. We find changes in business practice from mentorship that are consistent with our interpretation that mentees are gaining local, market-specific information from their mentors. Mentees are more likely to have switched suppliers in the aftermath of the treatment, which generates greater profitability per unit sold. We find that this increase in profitability is driven by lower costs rather than higher quality or a more profitable product mix. We find little evidence of changes in firm quality (customer relations, marketing, production quality, etc.) for the mentor treatment.[4]

By exploiting a discontinuity in our mentor selection procedure we can also measure the effect on mentors themselves. A priori, the expected effect on profit of being a mentor is ambiguous. Mentor profit could increase, if mentor-mentee pairs effectively act as "self-help groups," or decrease, if time or other resources spent mentoring negatively impact business operation. By surveying the mentors and similar non-mentor business owners just below our cutoff used to select mentors, we find no effect on profits for mentors around the discontinuity. This implies one-way flows of skill and knowledge to the mentee. This is consistent with models of knowledge transfer or diffusion that typically assume that the gains from an interaction between two firms accrue solely to the less productive firm

[2] Our mentor selection procedure is based on age and experience thresholds, as well as a profit cutoff. This procedure is detailed in Section 3.

[3] An obvious concern is that mentors may view their mentees as competition and, therefore, withhold valuable information. Our findings strongly conflict with this interpretation. We believe this is due to the extreme density of the community in which we work, which implies that mentors do not perceive their mentees to be direct competitors even if they are located only a few kilometers apart.

[4] Interventions that do find these changes either concurrently increase demand (Atkin, Khandelwal and Osman, 2016; Hardy and McCasland, 2016) or focus on larger firms (Bloom et al., 2013).

(e.g. Jovanovic and Rob, 1989; Lucas, 2009; Lucas and Moll, 2014; Buera and Oberfield, 2015; Perla and Tonetti, 2014).

These benefits from mentorship are obtained even though we find muted results from formal business training. Our results relating to formal business classes are consistent with the large and growing literature studying their effects. Formal business training generates a statistically insignificant 1 percent increase in profit relative to the control, which is in line with studies summarized in McKenzie and Woodruff (2014) and Blattman and Ralston (2015). While we find no evidence that profits change, we do find short run changes in business practices.[5] The lack of profit increase coupled with a short run change in business practices is consistent with previous work on microenterprise training, including Bruhn and Zia (2013) and Giné and Mansuri (2014). These results demonstrate that the mentorship treatment impact is not simply driven by an initial absolute lack of knowledge relative to contexts previously studied, but that mentorship is effective even in circumstances where classroom study has the same effects as found previously.[6]

For further evidence to understand the mechanisms behind these results, we exploit two different types of variation. First, we show that the treatment effect of mentorship is concentrated among business owners who more frequently utilize the market to purchase inventory. Among those who buy inventory at least once a week, we find a substantial benefit from mentorship, but none from the class treatment. On the other hand, among those who purchase inventory from the market less than once a week, we cannot distinguish either treatment from the control or each other. Second, we study the dynamics of the treatment effects by exploiting variation in outcomes between mentees who continue to meet with mentors and those who do not.[7] Despite nearly identical average profits at baseline, those who continue to meet with their mentors twelve months after the official end of the treatment period are on average 55 percent more profitable than those who are no longer meeting. This result could be explained by survivorship bias if matches with future benefits are the only ones that continue. We show that this is not the case by

[5] We interpret this as evidence against the possibility that these classes were of categorically lower quality than classes studied in the previous literature on formal business training.

[6] It is worth emphasizing that our goal is not to make the mentorship and class treatments as similar as possible. Instead, we want to show that (1) interacting with a successful business has a positive effect and (2) that a standard training intervention generates similar results to the literature within the same population.

[7] Meeting of mentors and mentees after the treatment period was not incentivized. We discuss the persistence of mentor-mentee relationships in Section 5.2.

showing that continued meeting cannot be predicted by previous changes in profit. Nearly 70 percent of matches were ended by the mentor as opposed to the mentee, which limits the effect of mentee selection. Our interpretation is that mentees benefit from mentorship and matches typically end only when mentors lose interest. The evidence therefore suggests that our positive treatment effects are driven by the flow of information from mentor to mentee.

It is worth emphasizing that our results are only informative about the barriers faced by the owners of very small businesses, as nearly all firms in our sample have no employees. It is possible that formal business skills could be useful to firms operating at a larger scale. For example, we would not expect that training in human resources would be useful to a firm with no employees. Likewise, it is possible that other formal business skills become effective, and even crucial, when businesses reach a larger scale, as Bloom et al. (2013) finds in India.

1.1 Related Literature

The literature on increasing managerial capital in micro and small firms has overwhelmingly focused on in-class training. McKenzie and Woodruff (2014) provide an excellent and comprehensive review of previous studies. The overriding theme of this research is that business practices do change, but translate into little impact on revenue and profit (for example, Bruhn and Zia, 2013; Giné and Mansuri, 2014), despite billions of dollars spent by governments and international organizations. Our training intervention is consistent with these results. The mentorship intervention, however, overcomes two issues in this literature. First, in-class training requires diagnosing what skills the entrepreneurs are missing, and then designing a curriculum to effectively address those deficiencies. This has proved difficult, in part because of the large number of different constraints facing microenterprises (Bruhn, Karlan and Schoar, 2013; Karlan, Knight and Udry, 2014). Our intervention allows each mentor-mentee pair to diagnose and discuss information or skills they believe to be most useful. Second, building evidence on the constraints facing businesses is expensive. Classroom training usually can cost over 100 U.S. dollars per student (Blattman and Ralston, 2015), while more personalized consultants can cost as much as 11,000 U.S. dollars, as in the case of Bruhn, Karlan and Schoar (2013). Our mentors were paid 9.83 U.S. dollars for one month of meetings, yet despite this, 45 percent were still meeting 17 months after the official treatment period. Our training class cost approximately 40 U.S. dollars per student.

Our solution to these issues – leveraging the skill of other individuals facing the same local economic conditions – is closely related to work on the transmission of knowledge through networks. Foster and Rosenzweig (1995), Munshi (2004), Bandiera and Rasul (2006), and Conley and Udry (2010) all document social learning in various contexts, while BenYishay and Mobarak (2015) and Beaman et al. (2015) experimentally vary the entry point of information and track its impact through existing networks. While these papers all show the importance of existing networks for information and technology flow, much less is known about the impact of (potentially) learning from an exogenously selected partner. More closely related to our exercise are Atkin, Khandelwal and Osman (2016), who study this in the context of supplier-customer relationships by randomly allocating new foreign rug orders to small Egyptian rug makers, and Cai and Szeidl (2016), who study the creation of randomly formed business groups in China. Interestingly, despite the fact that they focus on much larger firms, we find many of the same channels. However, they find more persistent effects, highlighting potentially important differences among firms of different scale.

Last, our results highlight a particular micro-foundation for distortions to managerial capital accumulation, albeit in a relatively specialized setting. Bhattacharya, Guner and Ventura (2013) and Da-Rocha, Mendes Tavares and Restuccia (2014) show that frictions limiting managerial capital accumulation can have an important aggregate effect in the context of policy distortion models developed in Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). Our results suggest a link between these distortion models and recent work by Lucas (2009), Lucas and Moll (2014), and Perla and Tonetti (2014) in which growth is generated through spillovers from other firms. Our empirical results support the mechanisms of these models. First, firms are capable of increasing profit through interaction with other firms, even in the case of exogenous matching. Second, interaction with a more profitable firm has a larger effect. Third, these benefits only accrue to the less profitable member of the match.

2 Microenterprise Characteristics in Dandora, Kenya

Dandora is a dense, urban slum to the northeast of Nairobi. It is approximately four square kilometers, and as of the 2009 census, contained 151,046 residents. To assess the business characteristics in the area, we conducted a street-level survey of 3,290 randomly selected business.[8] Table 1 provides summary statistics for business. Column three also includes the same information for "young" firms with owners under 40 years old and less than 5 years of experience, as we eventually draw our sample from this group. These businesses make up 43 percent of all businesses surveyed. The rest of the businesses — those with owners older than 40 or more than 5 years of experience — we classify as "experienced."

The average business in our survey has profit of 16,899 Kenyan shillings (Ksh), or 167 U.S. dollars, in the previous month. This is approximately 72 percent above GDP per capita in Kenya. However, while the average young owner earns 14,266 Ksh, the average experienced owner earns nearly 42 percent more profit per month or 20,168 Ksh. Figure 1 plots the distribution of the natural logarithm of profit for young and experienced enterprises.

This result is driven by selection. However, despite the substantial difference in profit, there is not much difference in observable formal business practices. They are equally likely to offer credit to customers, have a bank account, have taken a loan at some point in the past, or engage in formal accounting or advertising. Moreover, they are roughly equally educated.

We focus on female microenterprise owners, as they make up 71 percent of inexperienced owners. As Figure 2a shows, they are less profitable than their male counterparts at every business experience level. Interestingly, this percentage difference in profits is roughly constant over the first fifteen years of the firms' operating lives (Figure 2b).

[8]Our sampling procedure was intended to reduce convenience sampling, and worked as follows. We generated 200 points randomly throughout Dandora, and then gave each enumerator a list of randomly selected numbers. Starting from a randomly selected point, they were instructed to count businesses until they reached a number on their list, and survey the business owner of that establishment.

Figure 1: Log profit distribution for young and experienced enterprises

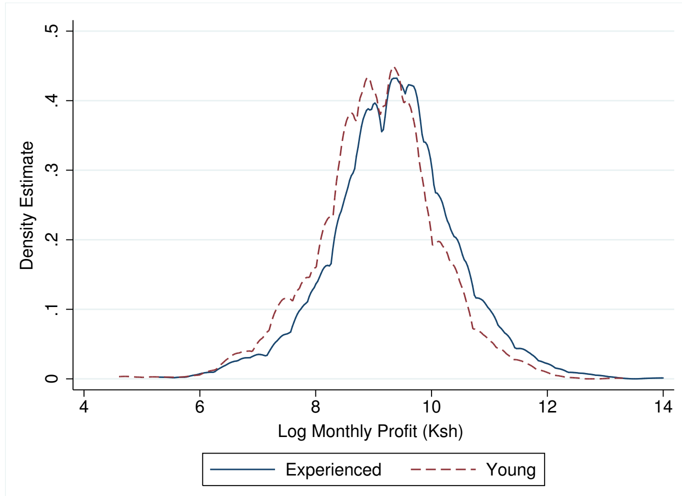
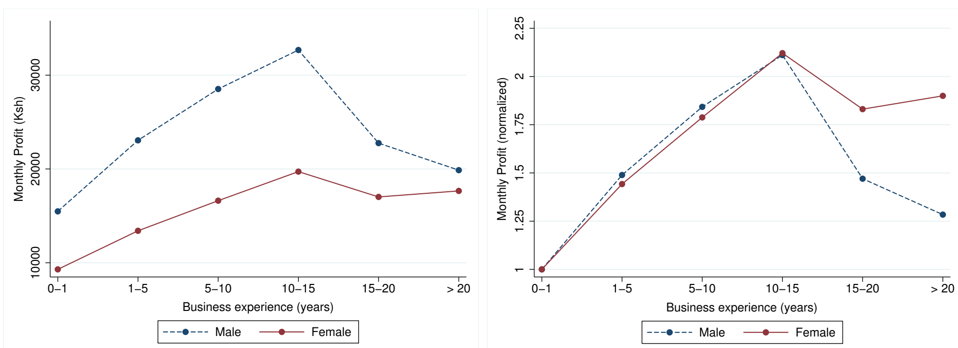


Figure 2: Gender differences over the lifecycle

(a) Average lifecycle profit (b) Average lifecycle profit growth



2.1 Self-Taught vs. Business Owners who Learn from Others

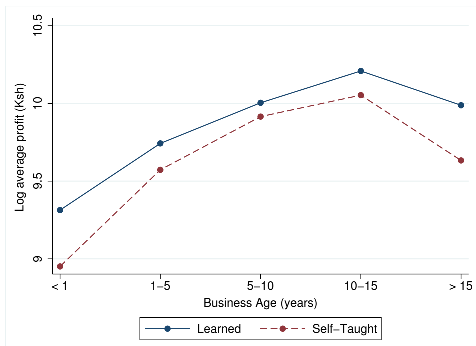
To motivate our study, we assess how profitability varies in the cross-section based on self-reported learning methods. Fifty-five percent of all firms claimed they were self-taught, while the rest claimed to learn either from another business operator, in school, or through an apprenticeship. We then consider various measures of business success across these two groups, and find that those who are self-taught make substantially less profit and operate at a smaller scale. Figure 3 plots three measures of business scale over the lifecycle.[9] First, Figure 3a shows that the self-taught earn less profit at any point over the lifecycle. The average self-taught firm has profit that is 82 percent of firms that learn from others.

[9]Total employment looks quite similar to Figure 3b given there are so few firms that have more than one worker.

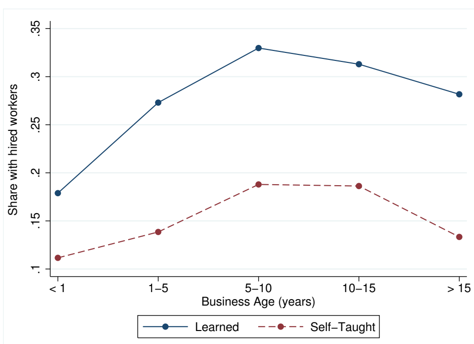
Other measures show similar patterns of self-taught operating at a lower scale than those who learned from others. Figure 3b shows that the self-taught are less likely to have employees and pay a smaller total wage bill.

Figure 3: Business scale differences over the lifeycle

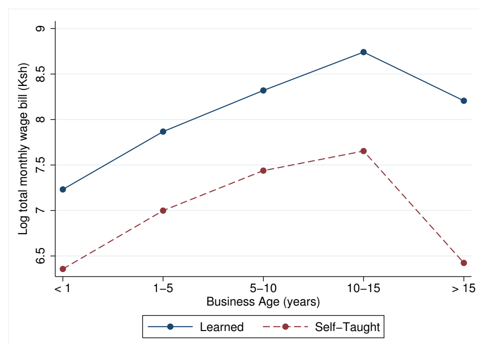
(a) Average profit



(b) Share with employees



(c) Total wage bill



The results provide suggestive evidence that learning from others plays an important role in the profitability of business, though not all businesses have access to it. However, confounding factors (e.g. selection) limit our ability to causal interences about the relationship between profitability and the type or source of information. Understanding this link is the purpose of our experiment.

3 Experimental Design

We use the baseline survey discussed in Section 2 to construct our experimental sample. We restricted our sample to business owners who are under 40 years old and have been running a business for less than 5 years. This includes 1,094 business owners, 787 of whom were operated by females. Out of these 787 female entrepreneurs, we contacted 723 to participate in the study after dropping some with a high fraction of missing baseline data or extreme outliers in the baseline. Of these, 538 (68%) accepted our invitation to participate in the program. We set up relatively strict participation requirements due to the numerous follow-up surveys expected, and in particular required attendance of an in-person orientation. Of the 538 individuals, 372 attended orientation (69% of 538, or 51% of the original 723). Randomization took place among these 538 individuals, and no one was given any indication of their assigned group until arrival at orientation. The control group received a cash payment of 4,800 Ksh (48 U.S. dollars) to encourage participation, which is equal to approximately two weeks of average profit. The classroom treatment received an identical cash payment along with one month of business classes. The mentor group received the cash payment in addition to a mentor drawn from local successful business owners. Of the original 372 individuals at orientation, 369 business owners answered at least one post-treatment survey.

The business classes were conducted by faculty from Strathmore University, a leading university in Kenya that is located in Nairobi. The classes have been used as part of a small and medium size business outreach program by the Strathmore University School of Management and Commerce. The curriculum was therefore based on what they believed to be the best available topics and information to cover. The treatment consisted of four two hour classes that broadly covered marketing, accounting, cost structure and inventory management, and the creation and development of business plans. These topics are similar to programs used in other studies.[10] Classes were offered at a local hall in Dandora, and were offered six times over the course of the week to accommodate individual schedules. While each of the four class topics had a separate instructor, the same instructor conducted all sections of each class topic. We refer to participants in this arm of the experiment as the classroom treatment.

Individuals assigned to the mentor treatment were matched with a mentor drawn from a set of successful local business owners (mentor selection is detailed in the next

[10] Anticipating the results somewhat, we find similar results to previous formal training research using other training programs, suggesting that there is nothing specifically different about our class design that generates our results.

subsection). Once the pool of mentors was chosen, mentees were matched based on narrowly defined business sectors. For example, we match perishable food sellers with perishable food sellers, tailors with other other tailors, and so on. Conditional on matching business sectors, mentors were randomly assigned. Mentees were required to meet with the mentor each week at the mentor's business, which was designed to minimize the cost to the mentors. The meetings, however, were relatively unstructured. We put no constraint on minimum meeting time nor the topics that must be discussed. To facilitate discussion, participants were given optional prompts, including "What were some of the challenges the mentee faced this week?" and "What was a time the mentor faced a similar challenge, and how did she respond to it?"

The treatment was completed at the end of November 2014. To understand the dynamics of the response across different treatment arms, we conducted six follow up surveys in the middle of December 2014 (preceding the Christmas holiday) and then in the last week of January, February, March, June, and November of 2015. The last two surveys contained a longer set of followups with more detailed business practice questions. Throughout the rest of the paper these surveys will be numbered by months since treatment, so that the surveys will be numbered $t \in \{1, 2, 3, 4, 7, 12\}$ will reference December, January, February, March, June, and November.[11]

3.1 Mentor Selection

The pool of mentors was selected from our baseline survey. We first constrained our sample to female business owners who were over 40 years old and had been operating the same business for at least 5 years. This left 366 individuals. We then ran a simple regression to control for age and sector-specific differences:

$$\log(\pi_i) = \alpha + \delta_i + \log(age)_i \gamma + \varepsilon_i \tag{3.1}$$

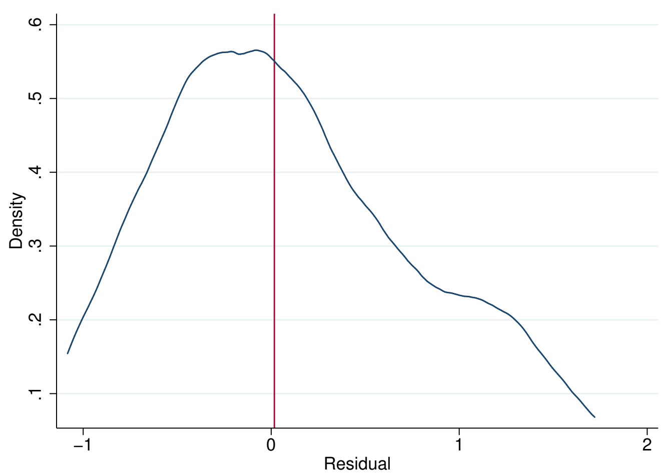
where π_i is baseline profit for individual i, δ_i is a sector fixed effect (manufacturing, retail, restaurant, other services) and age_i is age in years. Our mentors are chosen based on

[11] We conducted a much shorter survey 17 months after the treatment to assess the seasonality of the treatment effect (see Section 6). It is not included in the main analysis because it covered few variables of interest, but its inclusion does not change any of the results presented below.

having the highest residual profit $\widehat{\varepsilon}_i$. Given this, we simply sorted potential mentors by residual profit, then, starting with the most profitable, we recruited mentors until we have enough to link each mentee to a mentor that is in the same tightly defined business sector.[12]

Mentors were offered a token 1,000 shillings (9.83 U.S. dollars) to participate, which is equivalent to approximately two days of profit for the average firm in this group. We also told both mentors and mentees that meetings would take place at the mentor's business to minimize inconvenience to the mentor and maintain some similarity with the travel requirement of the class treatment. These incentives were sufficient to generate high take up, as 95 percent accepted our invitation to take part in the program. Figure 4 plots the distribution of $\widehat{\varepsilon}$ along with the cut-off.

Figure 4: Distribution of $\widehat{\varepsilon}$ and cut-off



As expected, mentors run substantially more successful businesses than those not chosen by all metrics. Profit is about 4 times higher among mentors (5,967 to 20,205 Ksh). Moreover, they are nearly three times as likely to have employees.

3.2 Sample Size and Balance across Survey Waves

Of the 372 individuals who attended orientation, 369 (99%) answered at least one followup. The response rates by wave were 352 (95% of 372), 318 (85%), 319 (86%), 323 (87%),

[12]The sector-specific estimates were small and statistically insignificant, as the correlation between the natural logarithm of baseline profit and $\widehat{\varepsilon}_i$ is 0.98. Using residual profit rather than raw profit did not change the set of potential mentors.

and 325 (87%). In total, 4 individuals completed exactly one followup, 6 completed two, 21 completed three, 41 completed four, and 108 completed five, and 194 completed all six. In Appendix B we provide survey round-specific balance tests. There is no evidence that attrition generates any observable differences across the groups. We further provide the correlation coefficients of baseline observables with number of surveys answered in Table 19 of Appendix B. A few observables are correlated with answering surveys at the 5 percent level, though none at 1 percent. However, the differences are small and we find little difference in estimation results with or without controlling for baseline factors.[13]

3.3 Take Up of Treatments

Attendance at the business class was recorded, and the average person in the classroom treatment attended 3.1 out of 4 classes. One person attended no classes, 11 percent attended one of four classes, 17 percent attended two, 32 percent attended three, and 40 percent attended all four. This is broadly in line with attendance in other studies (McKenzie and Woodruff, 2014). The mentorship treatment was used by all individuals at least once during the intended treatment period. In the last week of the official treatment, 85 percent had met with their mentor within the past week.

4 Empirical Results: Profitability

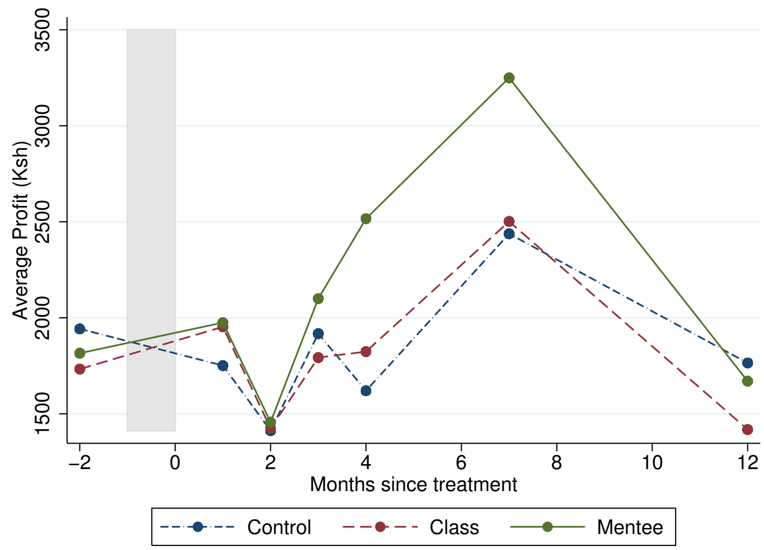
We begin by considering the impact of our various treatments on business profitability and scale in Section 4.1. We find that mentorship increased average profit relative to control, while in-class training had no statistical effect. In Section 4.2 we use a regression discontinuity design to show that there is no change in profit for mentors. Taken together, the results imply that our treatment generates match surplus rather than simply reallocation of profitability across firms.

[13]To give some examples, manufacturing business owners answer 5.8 surveys on average, compared to 5.2 for the rest. Restaurants answer 5.0 surveys, compared to 5.3 surveys for non-restaurants. The other two observables correlated with answering are firm age and owner age, which are naturally highly correlated. A business owner in the bottom 25 percent of the age distribution answers 5.1 surveys on average, compared to 5.3 in the top 25 percent of the age distribution.

4.1 Treatment Impact

We begin by looking at weekly profit for each treatment group. Figure 5 plots the time series of average weekly profit by treatment arm.

Figure 5: Profit time series



The classroom treatment mimics the control group closely throughout the study period, and actually lower than control in month 12.[14] The mentor treatment, however, sees a substantial growth in profit relative to both the control and class that we first pick up in month four and lasts through month seven. This effect fades out by our twelve month followup as mentor-mentee matches dissolve, which is discussed in detail in the next section. To make statistical inferences about these effects, we employ a series of regressions. First, following McKenzie (2012), we pool the data and run the ANCOVA regression to measure the average treatment effect:

$$y_{it} = \alpha + M_i\beta + C_i\gamma + y_{i0}\delta + X_i\eta + \theta_t + \epsilon_{it}. \tag{4.1}$$

Here, y_{it} is the outcome for individual i at time $t \in \{1, 2, 3, 4, 7, 12\}$ months since the treatment, while y_{i0} is the baseline outcome value. $M_i = 1$ if i is in the mentor group,

[14]There is an obvious decline in profit from December to January ($t = 1$ to $t = 2$) across all groups. This is the seasonal effect of a slow down in sales after December holidays, which we confirmed with numerous business owners in the study.

and $C_i = 1$ if i is in the classroom treatment. X_i is a vector of baseline controls including secondary education, log age, and business sector fixed effects, and an indicator for whether the firm employs any workers, while θ_t is a time fixed effect. All pooled regressions have standard errors clustered at the individual level. To understand the dynamics of the response, we supplement the pooled regression with wave-by-wave regressions:

$$y_{it} = \alpha_t + M_i\beta_t + C_i\gamma_t + y_{i0}\delta_t + X_i\eta_t + \varepsilon_{it} \quad \text{for} \quad t \geq 1 \tag{4.2}$$

Table 3 reports the impact on business profit. On average during this time period, mentee profit is 414 Ksh (4.08 USD) higher than the control group ($p = 0.002$), which is nearly 25 percent of baseline mean. The classroom treatment generates no statistically significant effect, and the point estimate of 75 Ksh ($p = 0.597$) is also much smaller. The average pooled effect of mentorship is 5.5 times higher than the average class effect. Comparing the two more formally, we can distinguish them at the five percent level, with a t-test p-value of 0.027. The time series of profit across the three groups shows that the average results are driven by a large increase that begins 4 months post-treatment. Looking back on Figure 5, this follows the general drop in demand following Christmas. In March 2015 (4 months post-treatment), profit of the mentees is 995 Ksh ($p = 0.000$) more than control compared to 313 Ksh more ($p = 0.257$) in the class treatment. This result remains into July 2015 (7 months post-treatment), as profit is 916 Ksh higher ($p = 0.007$) among mentees. Again, at both $t = 4$ and $t = 7$, we can statistically distinguish the two treatment effects at the five percent level, as the p-value of the t-tests are 0.014 and 0.018 respectively. Overall, the mentorship program generates a large average increase in profit relative to the control, while the in-class training program delivers almost no change in profitability. However, the effect fades over time, and we return to this theme in Section 5.

If mentors utilize their own skill or ability to increase mentee profit, better mentors should generate a larger treatment effect. We test this with the following estimating equation:

$$y_{it} = \alpha + M_{1it}\beta_1 + M_{2it}\beta_2 + M_{3it}\beta_1 + C_{it}\gamma + X_i\eta + \theta_t + \varepsilon_{it}. \tag{4.3}$$

where $M_{1it} = 1$ if i has a mentor from the bottom 25 percent of the baseline mentor profit distribution, $M_{2it} = 1$ if i's mentor is in the 25th to 75th percentile, and $M_{3it} = 1$ if i's mentor is in the top 25 percent. The results are in Table 4. On average, the better the mentor, the higher the effect on the mentee. The coefficient increases from 365 among the least profitable mentors to 515 Ksh among the most profitable. However, we cannot statistically distinguish these effects from each other with t-tests. The result is therefore broadly consistent with better mentors providing larger effects.

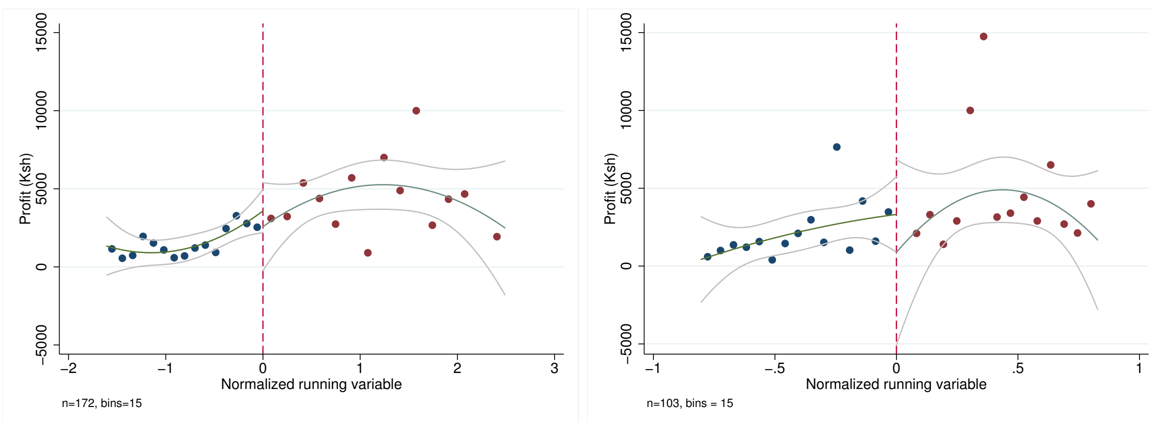
4.2 Impact on Mentors

Since mentee profit increases on average, we also wish to determine if there is any impact on the mentor. One possibility is that the mentor is an experienced agent who can impart knowledge on a less experienced business without receiving any return benefit from the interaction. This assumption is made in a large body of recent theoretical and quantitative work (e.g. Jovanovic and Rob, 1989; Lucas, 2009; Lucas and Moll, 2014; Buera and Oberfield, 2015). Alternatively, the mentor-mentee relationship may be better described as a collaboration or business group, where both sides gain from interacting with the other (e.g. Cai and Szeidl, 2016). We therefore ask whether the interaction imparts any gain in profit to the mentor. To do this, we employ a regression discontinuity design that exploits our mentor selection procedure.

As discussed above, mentors were paid 1000 Ksh for their participation. Therefore, when selecting mentors at the beginning of the program, we also contacted the 150 female business owners closest to the cutoff to participate as well. Ninety five agreed to participate, and they were also paid 1000 Ksh at the start of the program, so that we can compare the two groups with the effect of the cash transfer as a potential confounding factor. Four months after the treatment – a period with a significant profit increase among mentees – we surveyed all mentors, along with these 95 other non-mentor business owners. We then assess the impact of being chosen as a mentor on profit. For preliminary evidence that mentorship has no impact on the mentors, Figure 6 plots profit along with a fitted quadratic and its 95 percent confidence interval. Figure 6a uses the entire sample, while Figure 6b uses the Imbens and Kalyanaraman (2012) procedure to choose the optimal bandwidth. Both use 15 bins on either side of the cutoff.

Figure 6: Profit for mentors and non-mentors

(a) Full sample (b) IK Optimal Bandwidth



While Figure 6 suggests no discontinuity around the cutoff, we next assess this more formally. In particular, letting $\overline{\varepsilon}$ be the cut-off value for mentors derived from regression (3.1), we run the regression

$$\pi_i = \alpha + \tau D_i + f(N_i) + \nu_i \tag{4.4}$$

where π_i is profit, $D_i = 1$ if individual i was chosen as a mentor ($\widehat{\varepsilon}_i \geq \overline{\varepsilon}$ in regression 3.1), $f(N_i)$ is a flexible function of the normalized running variable $N_i = (\widehat{\varepsilon}_i - \overline{\varepsilon})/\sigma_\varepsilon$, and ν_i is the error term. The parameter τ captures the causal impact of being chosen as a mentor. We use local linear regressions to estimate the treatment effects on profit and inventory, along with business practices of record keeping and marketing. The results are in Table 5, and there is no evidence that mentors benefit from being mentors. Moreover, there is no change in marketing or record keeping practices. We do see some evidence that inventory spending decreases, but it cannot be statistically distinguished from zero. Overall, we find little evidence that mentorship changes either business scale or business practices for the mentors. When combined with the effect on mentees, this implies that the total observable gains to the match are driven by less productive member (the mentee), and moreover, mentee gains do not come at a cost to mentors. In Appendix C, we use the

MSE-optimal bandwidth procedure proposed in Calonico, Cattaneo and Titiunik (2014), along with varying the order of the polynomial for $f(\cdot)$. Neither changes the results.

5 Understanding the Treatment Impact: Channels and Dynamics

Section 4 demonstrates two key results. First, mentorship generates a substantial increase in profit while the class does not. Second, the treatment effect has non-trivial dynamics. We find that the changes among the mentorship group primarily relate to market-specific information, not the more general business skills covered in the class. We use regressions (4.1) and (4.2) to study the impact of treatments on supplier choice and inventory. We then study the dynamics of the treatment effects by considering the dissolution of matches over time. Finally, in Section 5.3, we return to business skills covered in classroom training.

5.1 Supplier Churning

We first consider the effect of each treatment on product sourcing by studying the propensity of firms to switch to new suppliers. The treatment effects are presented in Table 6 where the outcome variable is an indicator for whether or not the firm switched suppliers between the first and seventh survey wave. First, we note that supplier churning is very high in this environment. In the control group, 62 percent of firms switched suppliers, which suggests that the ability to acquire information about suppliers is salient for these firms. Second, we find a strong effect on the likelihood of switching suppliers in response to the mentorship treatment. Mentees are 19 percentage points ($p = 0.00$) more likely to switch suppliers than the control group. The class treatment, on the other hand, generates both a statistically and economically insignificant effect of -0.00 ($p = 0.98$). We reject the null hypothesis that these two effects are the same using t-test ($p = 0.002$).

The results suggest that mentorship has the effect of connecting mentees with better suppliers, since their greater propensity to switch is likely associated with greater profit from doing so. To demonstrate that the suppliers are better in the sense that they increase average profit (rather than just increasing the scale of firms), we construct a Lerner index as a measure of average profitability. Because most firms in the sample sell many products, we measure a multi-product Lerner index as the difference between total revenue and total cost divided by total revenue.[15] Notice that this Lerner index is independent of scale, so if firms simply increased in absolute size as a result of the treatment we would not expect any change in this measure. We find that mentorship has a positive and significant effect on this Lerner index of 8.6 percentage points ($p = 0.000$), while the classroom treatment has an insignificant 2.5 percentage point effect ($p = 0.376$). These effects on the multi-product Lerner index could come through a number of channels. First, since these are multi-product firms, mentees may shift their product mix toward higher profit items. Second, firms may increase the quality of their product and higher quality products may have higher markups. Third, perhaps the mentees switch to suppliers that charge lower prices.

To distinguish between these possible channels, we use information collected in waves 7 and 12 on each firm's main product and compute the Lerner index for that single product.[16] In Table 8 we consider the treatment effects on this single-product Lerner index. The Lerner index increases by 0.056 ($p = 0.113$) among mentees, which is an 18 percent increase over the control. While the mentee p-value is slightly above the 10 percent significance level, we can distinguish the two treatment effects from each other at the 10 percent level with a t-test ($p = 0.069$). We also report the effects on sale price of the main product and find no effect on either treatment group. Because we have only two waves with this data, the significance of the results for a single product is weaker than for the entire product line. However, the magnitudes are similar and this provides suggestive evidence about channels. First, these results are not, by definition, driven by changes in product mix. Second, the fact that the price did not increase means that there is no evidence that firms are selling products of higher quality. Therefore, we interpret this as evidence the result is driven by lower costs. Taken together with the evidence on supplier churning, this suggests that an important part of the effect on firm profit from mentorship is that mentees are able to switch to better suppliers that charge lower cost.

Finally, we run an additional test related to the importance of supplier and markets.

[15] The Lerner index takes values between zero and one, and represents and the percentage of the price that is realized as profit.

[16] A firm's main product is the project that accounts for the largest fraction of its profit.

We decompose the treatment effects based on how often owners visited the wholesale market at baseline.[17] We modify regression (4.1) and (4.2) with indicators for whether or not the owner makes at least one weekly trip to the market for supplies. If suppliers and cost are important margins, we should see larger effects among those who more frequently utilize the market. The results are displayed in Table 9. In the pooled regression, we find that mentees who make more frequent trips to the market increase profit by 564 Ksh ($p = 0.000$), compared to a statistically insignificant but negative effect of -377 Ksh ($p = 0.281$) for mentees who make less frequent trips. The p-value for the t-test testing equality of these two effects is $p = 0.013$, so we can statistically distinguish them at the five percent level. Two other interesting results emerge that are consistent with our results. First, among those business owners who use the market less than once per week, at no point can we statistically distinguish the class treatment from the mentorship treatment. Moreover, neither can be statistically distinguished from the control. Second, among those business owners who do utilize the market at least once a week, t-tests reject the equality of the two treatment effects in both the pooled regression ($p = 0.007$) and at months four ($p = 0.017$) and seven ($p = 0.008$), the two periods in which we see an average profit effect of mentorship.

All of these results confirm the importance of suppliers and cost as a critical margin changed by interacting with high quality business owners. In Appendix D, we include other measures of business scale (employment, wages, hours of operation, etc.) and show that these do not change. The last part of this paper therefore focuses then on why the dynamics of the treatment effects. We utilize the separation of mentee-mentor pairs, along with the substantial changes in suppliers that occur among the control group, to show that the profit increase is driven by the flow of information, not by permanent changes in practices.

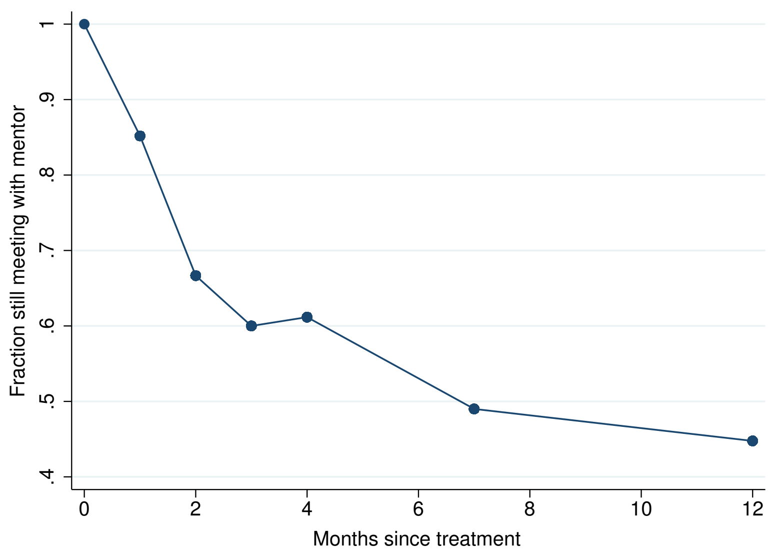
5.2 Match Persistence

We have shown that having a mentor increases profit, and primarily works through the ability to find lower cost suppliers. We now assess the dynamics of these effects. We

[17] We make this classification based on baseline characteristics because all behavior related to supplier relationships is clearly endogenous to the treatment.

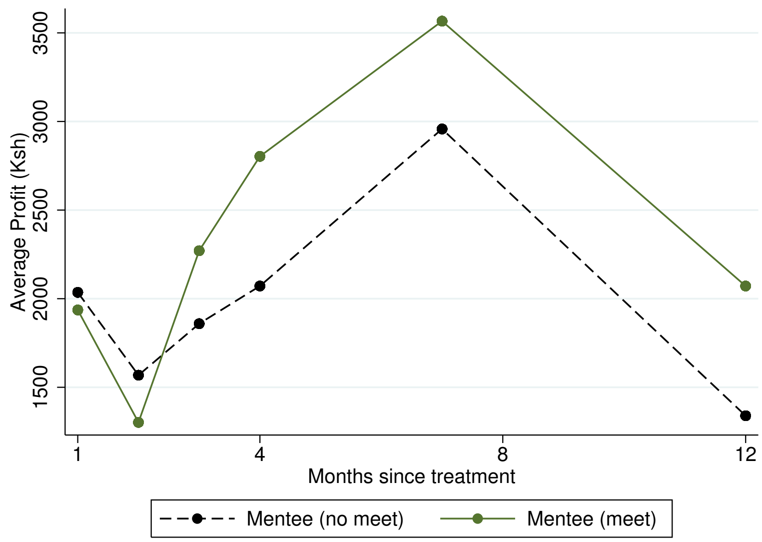
find that the lack of treatment effects from mentorship after one year is due to the fact that matches dissolve over time, coupled with the substantial churn in suppliers. That is, mentees lose their access to the information provided by mentors, and that information is only valuable for a limited period. Figure 7 plots the fraction of mentees still meeting with their mentor over the course of the study. All mentees met with their mentor in the official treatment month. This fraction declines over time with 45 percent still meeting after twelve months. No incentives were provided to continue the relationship beyond the treatment period.

Figure 7: Fraction of mentees still meeting with mentor



As highlighted in the previous section, nearly two thirds of the control group switch suppliers following the treatment, suggesting the value of being able to find the proper supplier is short lived. To the extent that the mentor is the link to this information, mentees that continue meeting should see higher profit. The data is consistent with this. Twelve months after the treatment, average profit for those still meeting with their mentor is 2071.38 compared to 1339.47 among those not meeting - a difference of 55 percent (and statistically significant at 0.05). This result holds despite the fact that the two groups on average have nearly identical baseline profit levels. Moreover, this result is not specific to the final wave, which can be seen in Figure 8, though it is largest in that wave. Four months after the treatment, profit is 35 percent higher ($p = 0.16$) for those still meeting with their mentor, and is 22 percent higher seven months after ($p = 0.23$), though it is worth emphasizing that the results are not precisely estimated enough to statistically distinguish the difference from each other in the relatively small sample.

Figure 8: Average Profit for Mentees



However, Figure 8 is not sufficient to imply that mentors continually deliver profitable information to mentees. That result requires identifying a positive counterfactual effect of continuing meetings after the meetings have already ceased in reality, which is of course an endogenous outcome. Alternatively, if Figure 8 is driven by selection, the counterfactual effect is close to zero. That is, mentees end the relationship with their mentors when meetings are no longer profitable. This would generate a time series like Figure 8, but imply no continued benefits if those mentees (counterfactually) continued to meet with their mentors. To test this idea, we first asked mentees directly why they no longer met with their mentors. Nearly 70 percent claimed it was due to the mentor ending the relationship, suggesting selection is not the key issue.

We also run an additional test not subject to the biases that potentially entered into mentee answers (though we were careful to explain that the program had ended, and any answer would have no positive or negative effects on anything in the future). The idea of this test is the following. If a mentee ends her match when all benefits have expired, we should see profit decreasing over time leading up to the match dissolution. That is, if a mentee has high profit at t_1 and low profit at t_2, then the selection explanation implies the match should end. Therefore, the likelihood of meeting with her mentor at t_3 should be positively related to the change in profits $\pi_{t_2} - \pi_{t_1}$: those who continue to generate profit growth through meetings continue to meet, while those who do not are less likely to meet. Alternatively, if matches end for reasons unrelated to mentee profitability (i.e. the mentor tires of time away from her own business, mentor moves to different city, etc.) then there should be no relationship between meeting and profit changes. We therefore ask whether changes in profitability affect the likelihood of meeting with a mentor in the future with the regressions

$$\begin{aligned} Meet_{it} &= \alpha + \beta \Delta\pi_{i,t-1} + \varepsilon \\ \Delta Meet_{it} &= \alpha + \beta \Delta\pi_{i,t-1} + \varepsilon \end{aligned}$$

run on just the mentees, with standard errors clustered at the individual level. The variable $Meet_t = 1$ if the mentee is still meeting with her mentor and $\Delta Meet_t = Meet_t - Meet_{t-1}$. We define $\Delta\pi_{i,t-1}$ in three different ways, to allow mentees to potentially react to different profit fluctuations. The first is $\Delta\pi_{i,t-1} := \log(\pi_{i,t-1}) - \log(\pi_{i,t-2})$, so that we test whether lagged profit changes impact meeting likelihood. The second is $\Delta\pi_{i,t-1} := \log(\pi_{i,t-1}) - \log(\pi_{i,0})$ which test whether changes from baseline profit predict meetings. Lastly, $\Delta\pi_{i,t-1} := \log(\pi_{i,t-1}) - \log(\overline{\pi}^{control}_{i,t-1})$. This regression considers that mentees may react to how well they are doing relative to the control group. If mentees are responding to changes in profit as is required for the selection explanation, we would expect to see $\widehat{\beta} > 0$ for at least one of these regressions. The results are presented in Table 10, and we find no evidence that meeting likelihood is responding to mentee profit realizations. In all six specifications, we have no statistically significant effects. Moreover, the point estimates are small. Combined with the previous evidence, this suggests that the cause of the decline in average effect is driven by the dissolution of matches by mentors, not necessarily by a decrease in the impact of continued mentorship. Therefore, the results in Figure 8 are not simply a result of selection, but show that the *flow* of information is critical to continued success of the mentee.

5.3 Skills Taught in Class

Lastly, we turn back to studying the business practices covered in the training classes. If the classes simply transferred no skill or knowledge, then it is certainly possible that standard training would have been beneficial had it been successful in changing practices. We refute that argument here, and show that our results are consistent with previous work on training classes. In particular, we see changes in underlying business practices covered in class, but they do not translate into profit changes.

In every survey, we asked about accounting and advertising practices, and Table 11 provides the time series of estimates using regressions (4.1) and (4.2). Marketing practices do not change relative to the control for either treatment. Accounting practices do change across treatments, and in fact we find a significantly larger impact among the class than the mentees. On average, 74 percent of the control does some sort of record keeping, compared 86 percent of those who receive in-class training (19 percent increase) and 77 percent of the mentees (7 percent increase). However, this effect is only present in the first four months following the treatment for the class treatment. This is consistent with short run changes in behavior found in other studies as well (e.g. Karlan, Knight and Udry, 2014), and implies that the in-class training does in fact change behavior without changing business outcomes.

For further evidence that classes change business practices, in our $t = 7$ and $t = 12$ surveys we asked a much longer battery of business practice questions. The questions are primarily drawn from the survey instrument first used in de Mel, McKenzie and Woodruff (2014) and McKenzie and Woodruff (2016) show that these questions are positively correlated with profit in a number of countries. Table 12 provides four aggregate measures of business practices. The *Aggregate Score* variable is the sum of *Marketing score*, *Stock score*, and *Record keeping score*. Each is presented as a standardized z-score to facilitate comparability, but we present the raw numbers in Appendix D for disaggregated categories. We combine the two waves of data and run the regression

$$y_{it} = \alpha + M_i\beta + C_i\gamma + X_i\eta + \theta_t + \epsilon_{it} \qquad \text{for } t = 7, 12$$

We do find changes associated with the stock score, which includes the likelihood of haggling with suppliers, whether or not the business owner compare suppliers, and how often she runs out of stock. This is consistent with the importance of suppliers and cost for generating changes in profit. Interestingly, however, we find similar effects between the class and mentorship group. The difference is that this information does not translate into profitability among the training class. Taken together, the results highlight the fact that the class was indeed successful in generating changes in behavior, but that they did not translate into increased profit. Moreover, they demonstrate that the mentorship treatment impact is not simply driven by an initial absolute lack of knowledge relative to places previously studied, but that mentorship is effective even in circumstances where classroom study has the same effects as found previously.

6 Ruling Out Alternative Explanations

6.1 Seasonality

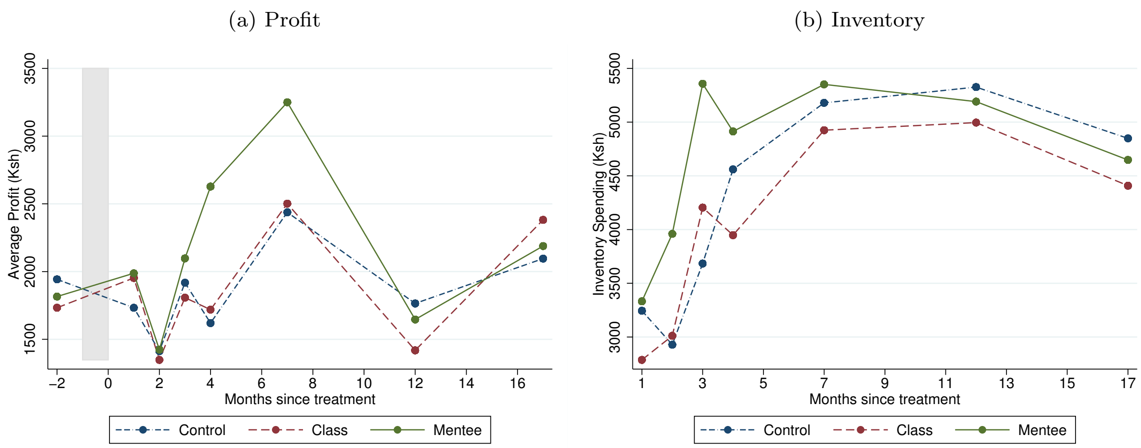
An alternative explanation is that the effect is seasonal. For example, perhaps mentors instructed mentees to do major inventory purchases in the beginning of the year. Then a potential long-term benefit only appears temporary within the context of one year. To answer this question we conducted a short survey 17 months after the treatment month asking only about profit and inventory. If the effect is seasonal, we should see an increase in profit for mentees. Figure 9 plots the profit and inventory time series including the $t = 17$ data, and shows that our results are not driven by seasonality or cyclicality.[18] While there is a cyclical component to the economy, it does not differentially affect the treatment groups. It is not the case, therefore, that aggregate shocks affecting all owners are driving the result.

6.2 Transfers between businesses

Another possibility is that mentors provided monetary help to the mentees. While interesting in its own right, it suggests constraints other than local information driving the

[18] We do not include these results in the broader analysis here because of the short nature of the survey. Their inclusion does not affect the interpretation of the results for either profit or inventory. All estimates are still significant, and there are only small changes in the pooled point estimates. They are available upon request.

Figure 9: Profit and Inventory (including $t = 17$)



results. Qualitatively, when asked about the most important benefit of having a mentor, only 6 percent mentioned anything related to financial help or transfers. More objectively, only two percent reported receiving any financial help from their mentor. Financial transfers do not seem to be critical in increasing mentee profit.

7 Conclusion

The prevalence of small, unprofitable businesses in developing countries presents a persistent and difficult problem. To investigate whether local expertise and business experience can benefit these entrepreneurs relative to a class, we conduct a randomized controlled trial in which all subjects receive a cash grant, some subjects are enrolled in a business course, some receive an experienced mentor, and some receive neither. Our results show that interacting with a successful, local business owner generates a 20 percent increase in profit, but that the effect fades over time, while the class generates no change in profit. We use the dynamics of the response to show that the result is driven by the dissolution of matches over time, as those that still meet earn higher profit. We argue that the reason for the treatment effect and its amelioration is that the value of information and supplier relationships are short-lived, and access to more experienced mentors who have a better understanding of the local market has substantial advantages. This also implies a rationale for the lack of success of formal training classes (at least in terms of higher profit). Most current training programs are designed to be easily portable and replicable, and therefore do not focus on the local information we have shown to be important.

In addition, the program is cost-effective relative to class training. Per dollar spent, the mentorship treatment increased profit by 1.73 USD, while the class increased profit by 0.25 USD. This is driven by both the higher benefits and 19 percent lower cost compared to the class treatment. These calculations are discussed further in Brooks, Donovan and Johnson (2017).

This paper joins a recent literature that shows specific and individualized interaction with peer groups, consultants, and mentors can be harnessed to improve business outcomes. Still unknown is how the the scale of the business determines which type of interaction is most effective. While Bloom et al. (2013) finds that professional consultants have positive effects on large businesses in India, Karlan, Knight and Udry (2014) find no increase in profit from a similar intervention targeting microenterprises in Ghana. We use mentors from the same community and find positive effects on microenterprises in Kenya. Cai and Szeidl (2016) show that peer groups have a positive effect on medium-scale businesses in China. Further research should focus on how the effectiveness of these different types of arrangements interacts with firm scale.

References

Atkin, David, Amit K. Khandelwal, and Adam Osman. 2016. "Exporting and Firm Performance: Evidence from a Randomized Trial." MIT Working Paper.

Bandiera, Oriana, and Imran Rasul. 2006. "Social Networks and Technology Adoption in Northern Mozambique." *Economic Journal*, 116(514): 869–902.

Beaman, Lori, Ariel BenYishay, Jeremy Magruder, and A. Mushfiq Mobarak. 2015. "Can Network Theory-based Targeting Increase Technology Adoption?" Northwestern University Working Paper.

BenYishay, Ariel, and A. Mushfiq Mobarak. 2015. "Social Learning and Incetives for Experimentation and Communication." *Review of Economic Studies*. Forthcoming.

Bhattacharya, Dhritiman, Nezih Guner, and Gustavo Ventura. 2013. "Distortions, Endogenous Managerial Skills and Productivity Differences." *Review of Economic Dynamics*, 16(1): 11–25.

Blattman, Christopher, and Laura Ralston. 2015. "Generating employment in poor and fragile states: Evidence from labor market and entrepreneurship programs." Columbia University Working Paper.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaning Management Practices Across Firms and Countries." *Quarterly Journal of Economics*, 122(4): 1351–1408.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India." *Quarterly Journal of Economics*, 128(1): 1–51.

Brooks, Wyatt, Kevin Donovan, and Terence R. Johnson. 2017. "Is There a Cost-Effective Means of Training Microenterprises?" University of Notre Dame Working Paper.

Bruhn, Miriam, and Bilall Zia. 2013. "Stimulating managerial capital in emerging markets: the impact of business training for young entrepreneurs." *Journal of Development Effectiveness*, 5(2): 232–266.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar. 2010. "What Capital Is Missing in Developing Countries?" *American Economic Review Papers and Proceedings*, 100(2): 629–633.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar. 2013. "The Impact of Consulting Services and Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico." World Bank Policy Research Working Paper 6508.

Buera, Francisco J., and Ezra Oberfield. 2015. "The Global Diffusion of Ideas." Chicago Fed Working Paper.

Cai, Jing, and Adam Szeidl. 2016. "Interfirm Relationships and Business Performance." University of Michigan Working Paper.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, 82(6): 2295–2326.

Conley, Timothy G., and Christopher R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review*, 100(1): 35–69.

Da-Rocha, Jose Maria, Marina Mendes Tavares, and Diego Restuccia. 2014. "Policy Distortions and Aggregate Productivity with Endogenous Establishment-Level Productivity." University of Toronto Working Paper.

de Mel, Suresh, David McKenzie, and Christopher Woodruff. 2014. "Business training and female enterprise start-up, growth, and dynamics: Experimental Evidence from Sri Lanka." *Journal of Development Economics*, 106(1): 199–210.

Foster, Andrew D., and Mark R. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy*, 103(6): 1176–1209.

Giné, Xavier, and Ghazala Mansuri. 2014. "Money or Ideas? A Field Experiment on Constraints to Entreprenuership in Rural Pakistan." World Bank Policy Research Working Paper 6959.

Gollin, Douglas. 2008. "Nobody's Business But My Own: Self Employment and Small Entreprise in Economic Development." *Journal of Monetary Economics*, 55(2): 219–233.

Hardy, Morgan, and Jamie McCasland. 2016. "It Takes Two: Experimental Evidence on the Determinants of Technology Diffusion." University of British Colombia Working Paper.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124(4): 1403–1448.

Imbens, Guido, and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies*, 79(3): 933–959.

Jovanovic, Boyan, and Rafael Rob. 1989. "The Growth and Diffusion of Knowledge." *Review of Economic Studies*, 56(4): 569–582.

Karlan, Dean, Ryan Knight, and Christopher Udry. 2014. "Consulting and Capital Experiments with Microenterprise Tailors in Ghana." *Journal of Economic Behavior and Organization*. Forthcoming.

Lucas, Robert E. 2009. "Ideas and Growth." *Economica*, 76(301): 1–19.

Lucas, Robert E., and Benjamin Moll. 2014. "Knowledge Growth and the Allocation of Time." *Journal of Political Economy*, 122(1): 1–51.

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

McKenzie, David, and Christopher Woodruff. 2014. "What are we Learning from Business Training and Entrepreneurship Evaluations around the Developing World?" *World Bank Research Observer*, 29(1): 48–82.

McKenzie, David, and Christopher Woodruff. 2016. "Business Practices in Small Firms in Developing Countries." *Management Science*. Forthcoming.

Munshi, Kaivan. 2004. "Social learning in a heterogenous population: technology diffusion in the Indian Green Revolution." *Journal of Development Economics*, 73(1): 185–213.

Perla, Jesse, and Christopher Tonetti. 2014. "Equilibrium Imitation and Growth." *Journal of Political Economy*, 122(1): 52–76.

Restuccia, Diego, and Richard Rogerson. 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." *Review of Economic Dynamics*, 11(4): 707–720.

Appendices

A Main Tables

Table 1: Baseline Characteristics

	Overall (3290)	Young Firms (1405)
Firm Scale:		
Profit (last month)	16,899	14,226
Firm Age	5.6	2.1
Has Employees?	0.21	0.18
Number of Emp (if $n > 0$)	1.8	1.5
Business Practices:		
Offer credit	0.67	0.69
Have bank account	0.36	0.30
Taken loan	0.21	0.15
Practice accounting	0.11	0.12
Advertise	0.10	0.09
Owner:		
Age	34.0	28.9
Female	0.65	0.71
Secondary Education	0.58	0.58

Table notes: Trimmed profit drops the top and bottom 1 percent of answers. 3171 establishments answered about profit.

Table 2: Balancing Test at Baseline

	Control Mean (1)	Class - Control (2)	Mentor - Control (3)
Firm Scale:			
Profit (last month)	10,054	-360.95	-975.25
		(1175.44)	(1186.76)
Firm Age	2.39	0.19	-0.05
		(0.23)	(0.23)
Has Employees?	0.21	-0.06	-0.02
		(0.05)	(0.05)
Number of Emp.	0.21	-0.05	0.02
		(0.06)	(0.06)
Business Practices:			
Offer credit	0.74	0.00	-0.02
		(0.06)	(0.06)
Have bank account	0.30	-0.03	-0.03
		(0.06)	(0.06)
Taken loan	0.14	-0.03	-0.05
		(0.04)	(0.04)
Practice accounting	0.11	-0.07	0.00
		(0.04)	(0.04)
Advertise	0.07	-0.02	0.04
		(0.03)	(0.03)
Sector:			
Manufacturing	0.04	0.00	-0.03
		(0.02)	(0.02)
Retail	0.69	-0.12	0.03
		(0.06)**	(0.06)
Restaurant	0.14	0.06	-0.02
		(0.05)	(0.05)
Other services	0.17	0.06	0.06
		(0.05)	(0.05)
Owner Characteristics:			
Age	29.1	0.87	-0.25
		(0.65)	(0.64)
Secondary Education	0.51	-0.04	-0.00
		(0.06)	(0.06)
Observations	119	129	124

Table Notes: Columns 1-3 are the coefficient estimates from the regression $y_i = \alpha + \gamma C_i + \beta M_i + \varepsilon_i$, where C_i and M_i are indicators for the class and mentorship treatments. Column 1 is $\widehat{\alpha}$. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***.

Table 3: OLS Estimates of Profit Equation at Different Time Periods

		Months since treatment					
	Pooled	(1)	(2)	(3)	(4)	(7)	(12)
Mentee	414.46	313.17	50.22	321.06	994.68	915.59	-0.40
	(133.07)***	(200.26)	(208.94)	(272.72)	(279.15)***	(336.22)***	(211.37)
Class	75.25	255.82	17.05	34.84	312.58	141.01	-172.06
	(147.34)	(196.04)	(203.32)	(264.57)	(275.37)	(330.48)	(211.54)
Control mean	1803.48	1733.11	1412.43	1903.30	1620.28	2437.84	1764.84
p-value, $H_0 : M = C$	0.027	0.719	0.872	0.285	0.014	0.018	0.419
Obs.	1902	345	311	312	316	302	318
R^2	0.091	0.056	0.028	0.057	0.081	0.076	0.100
Controls	Y	Y	Y	Y	Y	Y	Y

Table notes: Robust standard errors are in parentheses. Standard errors for pooled regressions are clustered at individual level and include wave fixed effects. Controls include secondary education, age of owner, sector fixed effects, and an indicator for any employees. The top and bottom one percent of dependent variables are trimmed. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***. The p-value provided is the p-value for testing equality of treatment effects across the two treatment arms.

Table 4: Heterogeneous Mentor Effects from OLS Profit Regression

	Pooled	Months since treatment (1)	(2)	(3)	(4)	(7)	(12)
Mentee: mentor in $(0, 25)$ pctile	356.03	235.92	-43.99	232.89	747.57	1000.83	-37.31
	$(180.17)^{**}$	(256.74)	(259.35)	(336.44)	$(360.20)^{**}$	$(412.41)^{**}$	(266.36)
Mentee: mentor in $(25, 75)$ pctile	449.11	429.69	93.80	596.22	1151.34	629.19	-133.49
	$(172.01)^{***}$	(270.31)	(278.65)	(372.14)	$(370.33)^{***}$	(461.45)	(279.95)
Mentee: mentor in $(75, 100)$ pctile	514.54	-88.66	285.35	-117.46	1317.05	1338.68	522.33
	$(270.54)^{*}$	(406.72)	(450.86)	(372.14)	$(567.77)^{**}$	$(666.64)^{**}$	(433.43)
Class	74.03	192.74	14.91	32.28	304.47	140.60	-173.59
	(142.49)	(201.10)	(203.85)	(264.82)	(275.82)	(331.04)	(211.56)
Control mean	1803.48	1733.11	1412.43	1903.30	1620.28	2437.84	1764.84
p-value, $H_0 : M_H = M_L$	0.601	0.457	0.490	0.557	0.355	0.634	0.228
p-value, $H_0 : M_L = C$	0.136	0.865	0.817	0.543	0.211	0.033	0.608
Obs.	1902	345	315	312	316	302	316
R^2	0.090	0.059	0.003	0.095	0.085	0.079	0.106
Controls	Y	Y	Y	Y	Y	Y	Y

Table notes: Robust standard errors are in parentheses. Standard errors for pooled regressions are clustered at individual level and include wave fixed effects. Controls include secondary education, age of owner, sector fixed effects, and an indicator for any employees. The top and bottom one percent of dependent variables are trimmed, though results are robust to other (or no) trimming procedures. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***. The first p-value is for test of equality of "high" (percentile 75-100) mentor and "low" (percentile 0-25) mentor effects, while the second is the p-value for equality of means of "low" mentor and class.

Table 5: Regression discontinuity results for mentor treatment effect

% of IK optimal bandwidth	Scale: Profit	Scale: Inventory	Practices: Marketing	Practices: Record keeping
100	-482.61	-1526.83	0.01	0.02
	(1325.07)	(2296.83)	(0.11)	(0.18)
150	313.67	-943.97	0.01	0.07
	(1408.75)	(2028.38)	(0.09)	(0.14)
200	329.92	-148.09	0.01	0.10
	(1324.69)	(1734.28)	(0.07)	(0.13)
Treatment Average	4387.34	8501.58	0.08	0.85
Control Average	1791.94	4005.06	0.13	0.63

Table notes: Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***. Profit and inventory are both trimmed at 1 percent, but results are robust to other (or no) procedures.

Table 6: Likelihood of Switching Suppliers

	Switch suppliers?
Mentee	0.187
	(0.065)***
Class	-0.002
	(0.069)
Control mean	0.617
p-value, $H_0 : M = C$	0.002
Obs.	304
R^2	0.070
Controls	Y

Table notes: Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***. Controls include secondary education, age of owner, sector fixed effects, and an indicator for any employees. p-values given are for the t-tests comparing equality of the different treatment effects.

Table 7: OLS Treatment Effect on Multi-Product Lerner Index

		Months Since Treatment			
	Pooled	(3)	(4)	(7)	(12)
Mentee	0.086	0.203	0.051	0.127	-0.037
	(0.023)***	(0.054)***	(0.049)	(0.044)***	(0.037)
Class	0.025	0.012	0.063	0.065	-0.040
	(0.022)	(0.043)	(0.045)	(0.042)	(0.037)
Control mean	0.369	0.388	0.385	0.382	0.280
p-value, $H_0 : M = C$	0.006	0.000	0.817	0.145	0.937
Obs.	1184	295	299	299	291
R^2	0.087	0.112	0.086	0.101	0.095
Controls	Y	Y	Y	Y	Y

Table notes: Standard errors are in parentheses. Standard errors for pooled regressions are clustered at individual level and include wave fixed effects. Controls include secondary education, age of owner, sector fixed effects, and an indicator for any employees. The top and bottom one percent of dependent variables are trimmed. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***.

Table 8: OLS Treatment Effect on Price and Lerner Index in Different Time Periods

	Sales Price				Lerner Index			
	Pooled	(1)	(7)	(12)	Pooled	(1)	(7)	(12)
Mentee	52.38	38.60	61.26	56.51	0.079	0.071	0.110	0.058
	(49.86)	(59.63)	(70.88)	(59.38)	$(0.217)^{***}$	$(0.036)^{**}$	(0.040)$***$	(0.038
Class	19.08	-11.46	-22.75	94.06	0.019	0.026	0.010	0.017
	(46.58)	(58.39)	(67.72)	(58.81)	(0.021)	(0.035)	(0.038)	(0.037)
Control mean	238.32	203.89	264.45	205.29	0.269	0.197	0.268	0.273
p-value, $H_0 : M = C$	0.482	0.395	0.214	0.526	0.005	0.204	0.009	0.282
Obs.	921	351	276	294	914	351	280	283
R^2	0.087	0.084	0.129	0.101	0.063	0.046	0.090	0.038
Controls	Y	Y	Y	Y	Y	Y	Y	Y

Table notes: Standard errors are in parentheses. Standard errors for pooled regressions are clustered at individual level and include wave fixed effects. Controls include secondary education, age of owner, sector fixed effects, and an indicator for any employees. The top and bottom one percent of dependent variables are trimmed. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***.

Table 9: Profit Effects Based on Frequency of Market Visits

		Months since treatment					
	Pooled	(1)	(2)	(3)	(4)	(7)	(12)
Mentee: $\geq$ 1 trip/week	564.31	484.95	216.30	414.453	1024.51	1169.53	111.63
	(144.54)***	(219.22)**	(229.30)	(296.12)	(304.71)***	(359.65)***	(232.74)
Mentee: $<$ 1 trip/week	-376.66	-1076.37	-864.59	-344.60	802.79	-10.74	-544.66
	(348.89)	(558.90)*	(558.27)	(775.16)	(757.66)	(1010.31)	(546.89)
Class: $\geq$ 1 trip/week	115.80	277.15	20.98	85.01	281.86	201.37	-104.49
	(142.03)	(214.23)	(217.75)	(283.11)	(297.63)	(348.88)	(228.21)
Class: $<$ 1 trip/week	-328.27	-622.40	-294.16	-470.32	401.37	-259.02	-623.19
	(460.19)	(566.30)	(573.29)	(781.81)	(759.80)	(331.04)	(574.71)
Control mean	1803.48	1733.11	1412.43	1903.30	1620.28	2437.84	1764.84
p-value, $H_0 : M_H = M_L$	0.013	0.010	0.075	0.361	0.787	0.272	0.272
p-value, $H_0 : C_H = C_L$	0.349	0.138	0.607	0.504	0.883	0.671	0.402
p-value, $H_0 : M_H = C_H$	0.007	0.350	0.397	0.268	0.017	0.008	0.362
p-value, $H_0 : M_L = C_L$	0.900	0.331	0.212	0.840	0.534	0.736	0.873
Obs.	1902	345	315	312	316	302	316
R^2	0.095	0.075	0.042	0.093	0.083	0.099	0.105
Controls	Y	Y	Y	Y	Y	Y	Y

Table notes: The treatment effects are for the treatments interacted with frequency of market visits in baseline survey. Robust standard errors are in parentheses. Standard errors for pooled regressions are clustered at individual level and include wave fixed effects. Controls include secondary education, age of owner, sector fixed effects, and an indicator for any employees. The top and bottom one percent of dependent variables are trimmed. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***. M_H is the the treatment effect for mentees who take at least one trip to the market per week (e.g. "high") and M_L are those that do not (e.g. "low"). C_H and C_L are defined identically for the class group. p-values given are for the t-tests comparing equality of the different treatment effects.

Table 10: Relationship between Meeting with Mentor and Previous Profit Realizations

Panel A: Meet$_t$						
	Meet$_t$	Meet$_t$	Meet$_t$	Meet$_t$	Meet$_t$	Meet$_t$
$\log \pi_{t-1} - \log \pi_{t-2}$	0.002	-0.001	–	–	–	–
	(0.017)	(0.018)				
$\log \pi_{t-1} - \log \pi_0$	–	–	0.030	0.033	–	–
			(0.019)	(0.021)		
$\log \pi_{t-1} - \log \overline{\pi}_{t-1}^{control}$	–	–	–	–	0.026	0.034
					(0.022)	(0.022)
Constant	0.563	0.650	0.558	0.643	0.610	0.868
	(0.026)***	(0.063)***	(0.026)***	(0.063)***	(0.023)***	(0.042)***
Obs.	460	460	469	383	551	551
R^2	0.000	0.012	0.005	0.018	0.002	0.060
Wave F.E.	N	Y	N	Y	N	Y

Panel B: Meet$_t$ - Meet$_{t-1}$						
	Meet$_t$ - Meet$_{t-1}$	Meet$_t$ - Meet$_{t-1}$	Meet$_t$ - Meet$_{t-1}$	Meet$_t$ - Meet$_{t-1}$	Meet$_t$ - Meet$_{t-1}$	Meet$_t$ - Meet$_{t-1}$
$\log \pi_{t-1} - \log \pi_{t-2}$	-0.015	-0.011	–	–	–	–
	(0.026)	(0.027)				
$\log \pi_{t-1} - \log \pi_0$	–	–	-0.001	0.000	–	–
			(0.019)	(0.018)		
$\log \pi_{t-1} - \log \overline{\pi}_{t-1}^{control}$	–	–	–	–	-0.004	0.005
					(0.025)	(0.025)
Constant	-0.036	-0.202	-0.042	-0.205	-0.043	-0.205
	(0.019)*	(0.085)**	(0.024)**	(0.084)**	(0.019)**	(0.083)**
Obs.	414	414	423	423	423	423
R^2	0.001	0.015	0.000	0.014	0.000	0.015
Wave F.E.	N	Y	N	Y	N	Y

Table notes: Robust standard errors are in parentheses, and are clustered at individual level. Controls include secondary education, age of owner, sector fixed effects, and an indicator for any employees. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***. The variable $Meet_t = 1$ if an individual has met with their mentor in period t.

Table 11: OLS Treatment Effects for Business Practices at Different Time Periods

Panel A: Record Keeping		Months since treatment					
	Pooled	(1)	(2)	(3)	(4)	(7)	(12)
Mentee	0.072	-0.008	0.119	0.099	0.143	-0.022	0.105
	(0.027)***	(0.054)	(0.057)**	(0.059)	(0.061)**	(0.069)	(0.063)
Class	0.146	0.198	0.185	0.122	0.308	-0.047	0.098
	(0.027)***	(0.052)***	(0.057)***	(0.059)**	(0.060)***	(0.067)	(0.063)
Control Mean	0.654	0.719	0.683	0.689	0.561	0.629	0.636
p-value, $H_0 : M_L = C$	0.003	0.000	0.238	0.688	0.007	0.710	0.909
Obs.	1941	351	317	318	322	308	325
R^2	0.052	0.090	0.099	0.057	0.090	0.062	0.049
Controls	Y	Y	Y	Y	Y	Y	Y

Panel B: Marketing/Advertising		Months since treatment					
	Pooled	(1)	(2)	(3)	(4)	(7)	(12)
Mentee	-0.004	0.003	-0.022	0.077	-0.014	-0.015	-0.058
	(0.019)	(0.052)	(0.048)	(0.044)*	(0.036)	(0.055)	(0.046)
Class	-0.014	0.030	-0.050	-0.008	0.007	0.010	-0.086
	(0.020)	(0.052)	(0.047)	(0.043)	(0.035)	(0.053)	(0.045)*
Control Mean	0.201	0.202	0.163	0.097	0.075	0.186	0.182
p-value, $H_0 : M_L = C$	0.643	0.611	0.541	0.050	0.551	0.634	0.554
Obs.	1941	351	317	318	322	308	325
R^2	0.028	0.050	0.040	0.052	0.052	0.056	0.040
Controls	Y	Y	Y	Y	Y	Y	Y

Table notes: Robust standard errors are in parentheses. Pooled regression are clustered at individual level and include wave fixed effects. Controls include secondary education, age of owner, sector fixed effects, and an indicator for any employees. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***.

Table 12: OLS Treatment Effects on Aggregated Business Practice Measures used in McKenzie and Woodruff (2016)

		Score Components		
	Aggregate z-score	Marketing z-score	Stock z-score	Record keeping z-score
Mentee	0.140	0.006	0.248	0.088
	(0.098)	(0.096)	(0.099)**	(0.095)
Class	0.118	-0.022	0.248	0.063
	(0.093)	(0.096)	(0.096)**	(0.095)
p-value, $H_0 : M = C$	0.810	0.779	1.000	0.794
Obs.	633	633	633	633
R^2	0.076	0.059	0.061	0.062
Controls	Y	Y	Y	Y

Table notes: Robust standard errors are in parentheses. Controls include secondary education, age of owner, sector fixed effects, and an indicator for any employees. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***.

B Further Balance Tests and Attrition [for online appendix]

Table 13: Wave 1 Balance Test

	Control (114)	Class (125)	Mentor (113)
Firm Scale:			
Profit (last month)	10252	9783	9268
Firm Age	2.4	2.6	2.4
Has Employees?	0.09	0.10	0.13
Number of Emp (if $n > 0$)	1.3	1.3	1.3
Business Practices:			
Offer credit	0.75	0.75	0.75
Have bank account	0.30	0.28	0.27
Taken loan	0.15	0.10	0.09
Practice accounting	0.01	0.01	0.01
Advertise	0.06	0.05	0.11
Sector:			
Manufacturing	0.04	0.05	0.01
Retail	0.69	0.57	0.65
Restaurant	0.14	0.19	0.12
Other services	0.16	0.23	0.24
Owner Characteristics:			
Age	29.3	29.8	28.9
Secondary Education	0.52	0.48	0.51

Table 14: Wave 2 Balance Test

	Control (104)	Class (113)	Mentor (101)
Firm Scale:			
Profit (last month)	9675	9355	9161
Firm Age	2.49	2.59	2.38
Has Employees?	0.09	0.08	0.12
Number of Emp (if $n > 0$)	1.00	1.44	1.33
Business Practices:			
Offer credit	0.74	0.77	0.72
Have bank account	0.32	0.27	0.28
Taken loan	0.14	0.11	0.08
Practice accounting	0.01	0.01	0.00
Advertise	0.05	0.05	0.11
Sector:			
Manufacturing	0.05	0.04	0.01
Retail	0.67	0.57	0.69
Restaurant	0.15	0.19	0.09
Other services	0.15	0.22	0.22
Owner Characteristics:			
Age	29.2	29.4	28.9
Secondary Education	0.54	0.49	0.51

Table 15: Wave 3 Balance Test

	Control (103)	Class (115)	Mentor (101)
Firm Scale:			
Profit (last month)	9942	9802	9547
Firm Age	2.40	2.63	2.31
Has Employees?	0.11	0.10	0.12
Number of Emp (if $n > 0$)	1.27	1.36	1.5
Business Practices:			
Offer credit	0.73	0.76	0.72
Have bank account	0.29	0.28	0.29
Taken loan	0.15	0.10	0.08
Practice accounting	0.01	0.01	0.01
Advertise	0.07	0.03	0.09
Sector:			
Manufacturing	0.05	0.05	0.01
Retail	0.70	0.57	0.66
Restaurant	0.14	0.19	0.11
Other services	0.16	0.22	0.24
Owner Characteristics:			
Age	29.1	29.6	28.7
Secondary Education	0.51	0.45	0.53

Table 16: Wave 4 Balance Test

	Control (107)	Class (113)	Mentor (103)
Firm Scale:			
Profit (last month)	10380	9452	9371
Firm Age	2.38	2.67	2.37
Has Employees?	0.09	0.10	0.15
Number of Emp (if $n > 0$)	1.30	1.36	1.40
Business Practices:			
Offer credit	0.75	0.75	0.69
Have bank account	0.30	0.28	0.27
Taken loan	0.15	0.11	0.09
Practice accounting	0.01	0.01	0.01
Advertise	0.07	0.05	0.09
Sector:			
Manufacturing	0.05	0.05	0.01
Retail	0.69	0.54	0.66
Restaurant	0.14	0.20	0.13
Other services	0.17	0.23	0.22
Owner Characteristics:			
Age	29.7	29.7	29.2
Secondary Education	0.53	0.49	0.50

Table 17: Wave 5 Balance Test

	Control (101)	Class (110)	Mentor (104)
Firm Scale:			
Profit (last month)	10198	8986	9195
Firm Age	2.45	2.60	2.26
Has Employees?	0.09	0.09	0.15
Number of Emp (if $n > 0$)	1.33	1.40	1.40
Business Practices:			
Offer credit	0.74	0.75	0.71
Have bank account	0.31	0.26	0.25
Taken loan	0.15	0.10	0.07
Practice accounting	0.01	0.00	0.01
Advertise	0.05	0.05	0.12
Sector:			
Manufacturing	0.05	0.05	0.01
Retail	0.69	0.54	0.66
Restaurant	0.14	0.20	0.13
Other services	0.17	0.23	0.22
Owner Characteristics:			
Age	29.6	29.6	29.4
Secondary Education	0.50	0.49	0.51

Table 18: Wave 6 Balance Test

	Control (110)	Class (109)	Mentor (104)
Firm Scale:			
Profit (last month)	10293	8986	9167
Firm Age	2.48	2.60	2.31
Has Employees?	0.21	0.16	0.21
Number of Emp (if $n > 0$)	1.33	1.03	1.27
Business Practices:			
Offer credit	0.75	0.75	0.70
Have bank account	0.31	0.26	0.26
Taken loan	0.15	0.10	0.07
Practice accounting	0.01	0.00	0.01
Advertise	0.05	0.05	0.11
Sector:			
Manufacturing	0.04	0.05	0.01
Retail	0.70	0.54	0.66
Restaurant	0.15	0.17	0.12
Other services	0.14	0.23	0.25
Owner Characteristics:			
Age	29.6	29.6	29.3
Secondary Education	0.52	0.48	0.51

Table 19: Correlation of baseline observables with number of surveys completed

Variable	Correlation coefficient
Firm Scale:	
Profit (last month)	0.031
Firm Age	0.121**
Has Employees?	-0.051
Number of Emp (if $n > 0$)	0.041
Business Practices:	
Offer credit	0.081
Have bank account	0.067
Taken loan	-0.047
Practice accounting	-0.020
Advertise	-0.053
Sector:	
Manufacturing	0.101**
Retail	-0.006
Restaurant	-0.089*
Other services	0.003
Owner Characteristics:	
Age	0.077**
Secondary Education	0.073

Statistical significance at 0.10, 0.05, and 0.01 are denoted *, **, and ***.

C Robustness of Regression Discontinuity Assumptions [for online appendix]

In this Appendix, we use the MSE-optimal bandwidth procedure in Calonico, Cattaneo and Titiunik (2014) instead of the Imbens and Kalyanaraman (2012) procedure used in the text. We also allow for higher-order polynomials. In no specification do we find any statistically significant effects.

Table 20: Regression discontinuity results using Calonico, Cattaneo and Titiunik (2014)

	Scale		Practices	
Polynomial order	Profit	Inventory	Marketing	Record keeping
0	-224.28	-92.11	0.050	0.136
	(900.02)	(1859.80)	(0.089)	(0.196)
1	-670.03	-1203.4	0.086	0.224
	(2291.80)	(2772.1)	(0.125)	(0.362)
2	-4048.20	-2137.7	0.232	0.345
	(4267.00)	(5236.4)	(0.142)	(0.453)
Treatment Average	4387.34	8501.58	0.08	0.85
Control Average	1791.94	4005.06	0.13	0.63

Table notes: Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***. Profit and inventory are both trimmed at 1 percent, but results are robust to other (or no) procedures.

D Further Results on Business Scale and Skills [for online appendix]

D.1 More Measures of Business Scale

At $t = 7$ and $t = 12$, we included more detailed measures of business scale, including inventory stock, employment, and hours of business operation. In this Appendix section, we test whether these margins change in response to either of the treatments. We find that they do not, again highlighting the importance of the supplier and cost channels emphasized in the main text. To test them, we run a variation of the pooled regression (4.1) with the caveat that the variables are only available for two periods. The results are in Table 21.

Table 21: OLS Treatment Effects on Business Scale Measures

	Stock of inventory (Ksh)	Any employees?	Number of employees	Total wage bill (Ksh)	Hours open (last week)
Mentee	1304.43	-0.032	0.001	244.98	0.42
	(2104.53)	(0.022)	(0.034)	(212.08)	(2.29)
Class	-398.74	-0.027	-0.029	112.56	2.21
	(1881.65)	(0.022)	(0.034)	(165.48)	(2.08)
Control mean	11030.83	0.092	0.099	360.00	49.42
p-value, $H_0 : M = C$	0.364	0.806	0.334	0.579	0.423
Obs.	633	633	633	633	633
R^2	0.067	0.240	0.206	0.162	0.048
Controls	Y	Y	Y	Y	Y

Table notes: Standard errors are in parentheses. Results are presented for $t = 7$ and $t = 12$. Controls include a time fixed effect, secondary education, age of owner, sector fixed effects, and an indicator for any employees at baseline The top and bottom one percent of dependent variables are trimmed for all dependent variables except the 0-1 employee indicator, though results are robust to other (or no) trimming procedures. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***.

D.2 Decomposition of Business Scores

Table 22: Marketing Practices Decomposed

Panel A: $t = 7$		Marketing Score Components				
	Marketing Score	Check competitor price	Check competitor products	Have sales	Upsell	Advertise
Mentee	0.185	0.036	0.096	0.064	0.072	-0.083
	(0.195)	(0.061)	(0.060)	(0.067)	(0.073)	(0.061)
Class	-0.263	-0.058	-0.031	-0.017	-0.064	-0.092
	(0.161)	(0.054)	(0.053)	(0.061)	(0.071)	(0.058)
Control mean	1.505	0.206	0.186	0.289	0.536	0.289
p-value, $H_0 : M = C$	0.015	0.082	0.020	0.200	0.048	0.867
Obs.	308	308	308	308	308	308
R^2	0.104	0.036	0.060	0.072	0.048	0.094
Controls	Y	Y	Y	Y	Y	Y

Panel B: $t = 12$		Marketing Score Components				
	Marketing Score	Check competitor price	Check competitor products	Have sales	Upsell	Advertise
Mentee	-0.203	-0.111	-0.072	-0.038	0.041	-0.023
	(0.185)	(0.063)*	(0.065)	(0.056)	(0.053)	(0.055)
Class	0.248	0.106	0.108	0.020	0.132	-0.117
	(0.202)	(0.067)	(0.068)	(0.058)	(0.058)**	(0.050)**
Control mean	1.400	0.373	0.391	0.227	0.191	0.218
p-value, $H_0 : M = C$	0.022	0.001	0.008	0.292	0.127	0.061
Obs.	325	325	325	325	325	325
R^2	0.081	0.064	0.047	0.046	0.090	0.037
Controls	Y	Y	Y	Y	Y	Y

Table notes: Robust standard errors are in parentheses. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***. Marketing score is computed by summing all its components.

Table 23: Stock Practices Decomposed

Panel A: $t = 7$		Stock Score Components		
	Stock Score	Haggle with suppliers	Compare suppliers	Run out of stock
Mentee	0.521	0.141	0.158	-0.221
	(0.118)***	(0.061)**	(0.067)**	(0.052)***
Class	0.445	0.107	0.119	-0.220
	(0.116)***	(0.061)*	(0.066)*	(0.048)***
Control mean	1.021	0.691	0.598	0.268
p-value, $H_0 : M = C$	0.482	0.517	0.534	0.952
Obs.	308	308	308	308
R^2	0.099	0.041	0.048	0.125
Controls	Y	Y	Y	Y

Panel B: $t = 12$		Stock Score Components		
	Stock Score	Haggle with suppliers	Compare suppliers	Run out of stock
Mentee	-0.12	-0.03	-0.05	0.02
	(0.13)	(0.07)	(0.07)	(0.06)
Class	-0.05	-0.04	0.02	0.01
	(0.13)	(0.07)	(0.07)	(0.06)
Control mean	1.51	0.21	0.19	0.29
p-value, $H_0 : M = C$	0.004	0.042	0.021	0.067
Obs.	306	306	306	306
R^2	0.025	0.010	0.015	0.008
Controls	Y	Y	Y	Y

Table notes: Robust standard errors are in parentheses. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***. Aggregate stock score is computed as *Haggle + Compare - Run out of stock*.

Table 24: Record Keeping Practices Decomposed

Panel A: $t = 7$		Record Keeping Score Components		
	Record Keeping Score	Record every sale	Consult records	Budget costs
Mentee	0.239	0.053	0.033	0.153
	(0.187)	(0.069)	(0.070)	(0.068)**
Class	0.070	-0.026	0.042	0.054
	(0.179)	(0.066)	(0.068)	(0.069)
Control mean	1.711	0.598	0.557	0.557
p-value, $H_0 : M = C$	0.346	0.243	0.901	0.131
Obs.	308	308	308	308
R^2	0.088	0.091	0.070	0.073
Controls	Y	Y	Y	Y

Panel B: $t = 12$		Record Keeping Score Components		
	Record Keeping Score	Record every sale	Consult records	Budget costs
Mentee	-0.001	0.090	0.033	-0.125
	(0.140)	(0.060)	(0.065)	(0.060)**
Class	0.094	0.021	0.032	0.401
	(0.151)	(0.062)	(0.066)	(0.066)
Control mean	1.391	0.700	0.345	0.345
p-value, $H_0 : M = C$	0.509	0.254	0.983	0.007
Obs.	325	325	325	325
R^2	0.071	0.052	0.054	0.065
Controls	Y	Y	Y	Y

Table notes: Standard errors are in parentheses. Statistical significance at 0.10, 0.05, and 0.01 is denoted by *, **, and, ***. Record keeping score is computed by summing all its components.