SIER Working Paper Series No. 102

Limited Impact of Business Development Programs on Entrepreneurs' Profitability in the Presence of Ambiguity Aversion

By

Dmitry Shapiro

February, 2019

Institute of Economic Research
Seoul National University

Limited Impact of Business Development Programs on Entrepreneurs' Profitability in the Presence of Ambiguity Aversion

Dmitry Shapiro*

February 25, 2019

Abstract

This paper develops a theoretical framework to explain a limited effect of business development programs (BDPs) on entrepreneurs' profit. We argue BDPs limited effect is due to mismatch between a BDPs' narrow focus on business-promoting strategies and a wider context in which microentrepreneurs operate. Entrepreneurs are ambiguity-averse and have multiple sources of income, e.g. business and wage incomes, that are correlated with each other. We show that for a sufficiently ambiguity-averse entrepreneur with multiple income sources efficient training can result in profit decline. We, further, show that both the ambiguity aversion and the multiplicity of income sources are crucial for our results. Only when the wider context (multiple income sources, ambiguity-aversion) is considered, the business-training impact is limited and can result in post-training profit decline. The limited impact is caused by the diversifying role that the business income plays in households' finances.

Keywords: Business development programs, ambiguity aversion, microentrepreneurship

JEL Classification Codes: O12, O16, D1

*Department of Economics, College of Social Sciences, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, South Korea. Phone: 82-2-880-2287, Email: dmitry.shapiro@snu.ac.kr. I am grateful to Chris Ahlin, Suresh de Mel, David Dicks, Kevin Donovan, Dean Karlan, Xavier Gine, Antoinette Schoar, Tavneet Suri, Hebe Verrest, Chris Udry as well as to attendants of the economist seminar at the University of Mannheim and Seoul National University for their feedback and comments. Financial support from the Center for National Competitiveness in the Institute of Economic Research of Seoul National University is gratefully acknowledged.

1 Introduction

Muhammad Yunis, in his *"Banker to the Poor"*, argued that teaching microentrepreneurs is a waste (Yunis, 1999). One cannot improve loan use since borrowers already use it efficiently. After all, the fact that the poor are alive despite all the adversity they face, is the best proof of their innate ability. Recent research, however, questions the scope of the "poor but rational" view. Karlan and Valdivia (2011) test whether microentrepreneurs maximize their profit given constraints and find that "*... [many microentrepreneurs'] activities prove to be generating an economic loss*" (p. 510). de Mel, McKenzie and Woodruff (2008) find that real returns on capital vary with borrowers' entrepreneurial ability, indicating that not everyone has the innate ability to do the best with what one has. Finally, there is no *a priori* reason why the "poor but rational" view would be true, as the poor lack the human capital and the connections that help build successful business (Banerjee, 2013).

One well-recognized way to make loan use more efficient is using business-training programs to improve microentrepreneurs' business knowledge (Prediger and Gut, 2014). Yet, the effect of business-training programs is mixed. Meta studies have shown that while entrepreneurship programs do have positive impact on business knowledge and practice, they have no impact on business expansion or income (Cho and Honorati, 2014). To make matters worse, some studies have documented negative effect of business trainings on profits. Karlan and Valdivia (2011) report that training of female entrepreneurs in Peru led to a noticeable improvement in "bad months", and less noticeable or even a decline in good months. Karlan, Knight and Udry (2012) study the effect of training on a group of tailors in Ghana. Business literacy of tailors in their sample increased, but profits declined. Bruhn and Zia (2011) conduct training of 445 clients in Bosnia and Herzegovina. They find that while basic financial knowledge improved, there was no improvement in the survival rate of business start-ups. Additionally, they find that profit declines, though insignificantly. de Mel, Mckenzie and Woodruff (2014) conduct a study using a sample of 1252 women in Sri-Lanka. They find that, for women with existing businesses, the training had no impact on profit. The training, however, had a positive impact on new owners' profitability. Finally, Drexler et al. (2014) report that only simplistic training — which consists mostly of basic rules of thumb — improves profit while the complex one does not.

An immediate explanation, which is that training programs are too complicated for microentrepreneurs to comprehend, is not supported by the evidence. Most papers report noticeable increase in business literacy after training. Giné and Mansuri (2014) specifically note that "*business training did lead to an increase in business knowledge, so lack of understanding is not the issue*" (p. 19). Limited impact of training does not appear to be due to improved accounting either. Drexler et al. (2014), for example, find that although there is a reduction in mistakes and more consistency across measures of how people calculate profits or sales, it does not affect main results. de Mel et al. (2014) compares self-reported profits to revenue and cost figures and control for detailed measures of accounting practices as a further robustness check. They also find no significant evidence that training changed reporting.

One explanation suggested in empirical literature is that a weakness of business development programs (BDPs), and the business-trainings they provide, is due to their narrow focus on business-promoting strategies that ignores a wide economic context in which microfinance clients operate. First, many microfinance clients are neither interested nor "*...particularly good at growing [their] businesses*" (Banerjee, 2013, p. 512). In a survey conducted in India, 80% of parents hoped their children would get government jobs, while 0% hoped their children would build successful business (Banerjee and Duflo, 2011). Verrest (2013) argued that BDPs "*are relevant to only a minority of entrepreneurs*" due to variations in household vulnerability and entrepreneurial ambition (p. 58). Second, microentrepreneurs do not view their business activities solely as a way to bring in more money. Instead, they consider it either as a valuable diversification tool in order to deal with irregularity in income sources (Krishna, 2004); or to reduce household's vulnerability to negative shocks such as job-loss or illness (Ellis, 2000); or for consumption and income smoothing (Bateman and Chang, 2009; Banerjee and Duflo, 2011).

The goal of this paper is to develop a theoretical model that shows how a mismatch between BDPs' narrow focus on business-promoting goals and a complex reality in which microentrepreneurs run their businesses can be responsible for a limited, or even negative, impact of business-trainings on microentrepreneurs' profits.

To capture wider context in which microentrepreneurs operate, we introduce two assumptions. First, we assume the microentrepreneurs' objectives are not limited to profit-maximization. In our model, the microentrepreneur has two ambitions: business-oriented ambition and livelihoods-oriented ambition. We model the business-oriented ambition as maximizing expected profit, and the livelihoods-oriented ambition as maximizing the "*rainy day*" profit, or more formally, the worst-case profit.[1] The microentrepreneur's utility is assumed to be a weighted average of the two ambitions: business-oriented ambition (expected profit) and livelihoods-oriented ambition (rainy day, or worst-case profit). As we will show, assuming such utility function is mathematically equivalent to a microentrepreneur being ambiguity-averse.

Second, we assume the microentrepreneur has multiple sources of income. One source of income in from the business activity. Other income sources can include income from farming, wage employment, or incomes from informal risk-sharing arrangements. Having multiple sources of income is quite common among poor households, yet this assumption is rarely used in the theoretical microfinance literature. A common approach is to focus on the business part of entrepreneurial income by, for example, normalizing all other incomes to zero.[2] As I will show in this paper, assuming away multiple sources of income comes with loss of generality. The training's impact will differ depending on whether the microentrepreneur has one or multiple sources of income available. For

[1] Broadly speaking, the livelihoods approach applies a holistic view on the poverty that includes economic, social, infrastructural and environmental factors. For example, Verrest (2013) classifies households with a livelihoods-motivation as those whose goal is to secure their livelihoods, be it by ensuring that they have "an apple for a rainy day", or by increasing consumption, or by having a hobby.

[2] The focus on the business part of the household's income is a common assumption starting from classical papers such as Besley and Coate (1995), Ghosh and Ray (2001) to more recent papers including Chowdhury (2005), Ahlin and Waters (2014), de Quidt et al. (2015) and Shapiro (2015).

example, only in multiple income setting can the post-training profit decline.

The microentrepreneur's income is determined by the amount of capital invested, and a state of nature. States of nature is the only source of uncertainty, and we assume that higher states make business activity more profitable. The microentrepreneur takes a business training course, which introduces a new business practice. The new practice is superior to the old one in that in every state it makes the capital more productive and, consequently, the business activity more profitable. Furthermore, we assume that the training's impact is stronger in states where capital is more productive, i.e., in higher states. The training does not affect profitability of non-business activities. Given the superiority of the new practice, the training (weakly) improves microentrepreneur's utility and, therefore, business training has always a positive effect from the welfare point of view.

The training has two effects on the expected profit. The first effect is *profit improvement effect*. It measures how the expected profit changes if the microentrepreneur invests the pre-training level of capital. It does not depend on microentrepreneurs' preferences or other sources of income, and due to superiority of the new practice it is always positive. The second effect is *capital adjustment effect*. After adopting the new practice, the post-training optimal level of capital changes. The capital adjustment effect measures how the expected profit changes due to change in the post-training investment. The capital adjustment effect depends on preferences and on non-business sources of income. It can be either positive or negative. Negativity of capital adjustment effect undermines the efficiency of the business-training and limits its impact. Whenever the capital adjustment effect is negative it means the microentrepreneur does not fully reap the benefits of the training. Instead of taking advantage of improved profitability and expanding her enterprise by investing more, the microentrepreneur finds it safer to invest less than before thereby invalidating the training's effect.

Whether the capital adjustment effect is positive or negative depends on the strength of the business-oriented ambition and on how much diversification the business income provides when combined with other incomes. If either the microentrepreneur's business ambition is sufficiently strong, or if the business income is not sufficiently diversifying then the training results in a positive capital adjustment effect and, therefore, in a higher post-training profit. In the former case, the post-training profit goes up because microentrepreneurs' goals and environment are sufficiently aligned with BDPs' focus on business promoting strategies. In the latter case, the post-training profit goes up because the diversification role of business income is weak and does not affect the microentrepreneur's investment decisions in a way that would hurt the post-training expected profit.

If, however, the microentrepreneur has sufficiently strong livelihoods-ambition *and* business income is sufficiently diversifying, then it is possible for the capital adjustment effect to be negative thereby limiting the efficiency of business-training. Moreover, it is possible for the negative capital adjustment effect to dominate the profit improvement effect so that the post-training profit declines. The intuition is as follows. Consider a livelihood-oriented entrepreneur who wants to maximize her rainy day profit. The training's focus on business part of microentrepreneur's income makes states with higher capital profitability to become even more profitable and, therefore, less of a concern for a livelihoods-oriented microentrepreneur. In such states, business activities generate enough funds to cushion against negative shocks to other sources of income. It is states with lower capital profitability that are more likely to be rainy-day-states and, therefore, to have a stronger effect on the utility of a livelihoods-oriented entrepreneur. Since optimal capital investment is lower in states with lower capital profitability, the microentrepreneur — instead of taking advantage of the improved profitability by investing more — invests less which results in the negative capital adjustment effect and, possibly, in profit decline.

Overall, the contribution of the paper is as follows. First, to the best of our knowledge this is the first theoretical paper providing an explanation to a limited, including negative, effect of training on the profit. McKenzie and Woodruff (2014) offered an empirical explanation, arguing that such issues as sample size and sample heterogeneity made it harder to detect the effect of training on profitability. In a follow-up paper, however, de Mel, McKenzie and Woodruff (2014) addressed those issues by using a large and homogeneous sample of female entrepreneurs. The authors found little impact on the profitability and concluded that "*the lack of impacts in most of the existing literature … may not be just due to power issues*" (p. 200). They also conjectured that business training programs might be less effective than previously thought. Second, differently from earlier theoretical literature, our model provides a holistic view of households by explicitly taking into account multiple sources of incomes, diversification needs, and a variation in entrepreneurial ambitions. Third, we show that the holistic modeling of the microentrepreneurial decision is crucial to understanding how efficient training can have mixed to negative impact. Only when diversification *and* lack of entrepreneurial ambition are introduced, can the training have negative impact on profit.

2 Basic Setup.

Consider a microentrepreneur who has access to multiple sources of income that include profit from business activities. We assume that profit from business activities, $\pi(s, K)$, depends on the amount of capital invested, K, and a realized state of nature, s. States are the only source of uncertainty and are labeled by integer numbers from 1 to n, $s \in \{1, \ldots, n\}$. The probability of state s is p_s. For any s, the state-profit function, $\pi(s, K)$, is a non-negative, single-peaked, concave differentiable function of K with $\pi'_K(s, K) < \infty$ for $K > 0$.

States are ordered in such a way that higher states are more favorable for a microentrepreneurial profit both in terms of profitability and marginal profitability of capital. Specifically, let $K^*(s)$ denote the optimal capital level in state s, $K^*(s) = arg\max_K \pi(s, K)$. We assume that $\pi(1, K^*(1)) < \cdots < \pi(n, K^*(n))$, which we will refer to as the *better*-assumption. We also assume that states and capital are complements, i.e. $\pi''_{sK} > 0$, which we will refer to as the *complementarity* assumption. The *better*-assumption implies that higher states have a higher upside potential. The complementarity assumption implies that capital's marginal profitability is higher in higher states. It follows from the complementarity assumption that $K^*(1) < \cdots < K^*(n)$. Finally, we assume that capital is necessary for the business-income so that $\pi(1, 0) = \cdots = \pi(n, 0) = 0$. That is without capital investment the profit is zero.

In addition to profits from business activities, the microentrepreneur also receives incomes from other sources. Let $h(s)$ be incomes from all sources other than business activities. As the notation indicates, we assume that income from non-business activities does not depend on K but can depend on the state of the nature. Given $\pi(s, K)$ and $h(s)$ the total incomes of the microentrepreneur is $I(s, K) = \pi(s, K) + h(s)$. It is immediate to verity that $I(s, K)$ is a concave single-peaked function of K such that $I''_{sK} > 0$. $h(s)$ is assumed to be such that $I(s, K)$ satisfies the *better*-assumption.

That households commonly rely on diversified income portfolios with multiple income sources — including subsistence and farming activities, wage employment, small-scale enterprises, informal risk-sharing arrangements, and temporary or permanent migration — is well-documented. For example, for rural households in South Asia, 60 percent of household income comes from non-farming sources; in sub-Saharan the number is in a range between 30 to 50 percent; and in southern Africa it may attain 80-90 percent (Ellis, 2000, p. 233).[3] At the same time, in theoretical literature it is a common assumption to normalize non-business income to zero which, as we show in the paper, is not without loss of generality.

The microentrepreneur is not solely focused on maximizing her expected profit. Instead, her utility is assumed to be a combination of two ambitions: *business-oriented* ambition and *livelihoods-oriented* ambition. We model business-oriented ambition as expected-profit maximization. We model livelihoods-oriented ambition as maximizing the "rainy day" profit, where we define the rainy day profit as the profit in the worst-case state, $\min_s \pi(s, K)$.[4] We will use terms "*rainy day profit*" and "*worst-case profit*" interchangeably throughout this paper.

Formally, the microentrepreneur's objective function is the weighed average of business-oriented and livelihoods-oriented ambitions,

$$U(K) = \left\{ (1-\eta) \sum_s I_s \pi(s, K) + \eta \min_s I(s, K) \right\}, \tag{2.1}$$

which she maximizes with respect to K. Parameter η is the weight that the microentrepreneur puts on her livelihood ambition, as captured by the rainy-day profit. When $\eta = 0$, the microentrepreneur has only business-oriented ambition, and her objective is the expected profit. When $\eta = 1$, the microentrepreneur has only livelihood-oriented ambition and her objective is the rainy-day profit.

The utility function above can be rationalized using the ambiguity-aversion framework.[5] Con-

[3]To provide a more specific numbers: a survey of households in Masaka district, Uganda, showed that for an average household, 64% of its income came from farm income, 20% from business profits and 10.6% from wages (Table 3.1, Ellis 2000). Survey of households in Mamone, a poor semi-arid community in South Africa, showed that the primary income source was remittances and other transfers (63.4%), wages accounted for 9.1%, business activities for 6.3% and farming activities for 12.8% (Table 3.2, Ellis, 2001). In Botswana, wage employment accounted for 21.5% of household income portfolio, crop and livestock farming for 45.8%, other activities (beer brewing, basket weaving, carpentry) for 18.5% (Valentine, 1993).

[4]Entrepreneurs are categorized on being either business- or livelihoods-oriented, based on their answers to the in-depth interviews. People with business-oriented ambition are those who perceived entrepreneurship as the way out of poverty and whose dream is to have their own well-organized business. People with livelihoods-oriented ambition are those who view their entrepreneurship as a secondary income to secure their livelihoods, to create savings ("*an apple for a rainy day*") as well as to increase consumption, or to have a hobby (Verrest 2013, p. 63, 64).

[5]The literature has documented the role of ambiguity aversion on willingness to switch towards new technologies

sider an ambiguity-averse microentrepreneur who does not know the underlying state-distribution and instead assumes that it belongs to a set of priors $\mathcal{Q} = \{q : q_s \geq \eta_s,\ q_s \geq 0,\ \sum_s q_s = 1\}$, where $\eta_s \geq 0$ (LeRoy and Werner, 2001, p. 82). We assume that $(\eta_1, \ldots, \eta_n)$ is proportional to correct probabilities, $\{p_s\}$, that is $\eta_s = (1-\eta) \cdot p_s$, where $1-\eta = \eta_1 + \ldots + \eta_n$. The microentrepreneur has maxmin preferences and chooses K to maximize the expected profit under the *worst* prior in $\mathcal{Q}$:

$$\max_K \min_{q \in \mathcal{Q}} \sum_s q_s I(s, K).$$

It is immediate to verify that the microentrepreneur's objective function can be re-written as:

$$U(K) = \min_{q \in \mathcal{Q}} \sum_s q_s I(s, K) = (1-\eta) \sum_s p_s I(s, K) + \eta \min_s I(s, K),^6 \tag{2.2}$$

which coincides with (2.1). Here $\{q_s\}$ is a distribution from $\mathcal{Q}$ and $\{p_s\}$ is the objective distribution.

In general, parameter η has two mathematically equivalent interpretations. The first one, used to derive (2.2), is that η measures the ambiguity of outcomes' distribution with higher η corresponding to higher ambiguity. When $\eta = 0$, the only element in $\mathcal{Q}$ is the objective distribution; when $\eta = 1$, $\mathcal{Q}$ contains all the possible priors. The second interpretation is that η measures the degree of microentrepreneur's ambiguity aversion with a higher η corresponding to higher ambiguity aversion. When $\eta = 0$, the microentrepreneur is risk-neutral; when $\eta = 1$, the microentrepreneur is the worst-case profit maximizer. Both interpretations are mathematically equivalent. Economic difference between the two is that, in the former case, the microentrepreneur's preferences are maxmin and do not depend on η; in the latter case, they do.[7]

From the concavity of the state-profit functions immediately follows that $U(K)$ is a concave and single-peaked function for any $\eta \in [0, 1]$. We will denote the optimal capital level as K^*_η. It will be convenient to denote K^*_0 as simply K^*, and K^*_1 as K_w where w stands for the word "*worst*". By definition, K^* maximizes the expected profit, and K_w maximizes the worst-case profit. Characterization of K_w will be of particular importance to us and is given by Proposition 1, which is proved in the Appendix. The proof is a straightforward application of complementarity and the *better*-assumption.

---

and practices, and the effect is distinct from risk-aversion. Engle-Warnick et al. (2007) document that farmers in Peru use a traditional variety of potato with low expected yield that, nonetheless, generates enough potatoes to feed a farmer's family. This is despite the availability of new varieties of potatoes such as the *Papa Caprio* that provide substantial yield improvement. They show that it is ambiguity-aversion and not risk-aversion that was responsible for the crop adoption decision. Similarly, Barham et al. (2014) examined adoption of genetically modified corn and soya beans among 191 Midwestern US grain farmers. Risk preference, measured using a coefficient of relative risk aversion, had no significant impact on adoption. Ambiguity aversion did have a significant effect and expedited adoption of the less ambiguous genetically-modified corn seeds.

[6] Let s_w denote the worst-case state (any worst state if there are more than one) given K: $I(s_w, K) \leq I(t, K)$ for any $t \neq s_w$. For a given K, the worst prior assigns the smallest probability to all states but the worst: $q_s^{worst}(K) = (1-\eta)p_s$ for $s \neq s_w$; and the worst state gets the remaining probability, $q_{s_w}^{worst}(K) = (1-\eta)p_{s_w} + \eta$.

[7] There are models with ambiguity-averse preferences that allow for separating the two interpretations. Objective function (2.2) is a special case of biseparable preferences with neo-additive weighting as introduced in equation (3) in Baillon et al. (2017) where $a_t = \eta$ and $b_t = -\eta$.

Proposition 1 *The optimal worst-case capital, K_w, is such that either*

i) $K_w = K^*(1)$*; or*

ii) there exist two states, $s < s'$*, such that* $\pi(s, K_w) = \pi(s', K_w) \leq \pi(t, K_w)$ *for any* t*, and* $\pi'_K(s, K_w) < 0 < \pi'_K(s', K_w)$*.*

The next Proposition shows how ambiguity-aversion affects choice of capital, as well as entrepreneurial's utility and expected profit. Having a higher degree of ambiguity-aversion distorts the capital's choice away from the level that maximizes expected-profit. It negatively affects the expected profit, but positively affects the rainy-day profit.

Proposition 2 *If* $K_w < K^*$ $(K_w > K^*)$*, then* K^*_η *is a decreasing (increasing) function of* η*. Microentrepreneur's utility,* $(1 - \eta)E_s\pi(s, K^*_\eta) + \eta\pi_w(K^*_\eta)$*, and expected profit,* $E_s\pi(s, K^*_\eta)$*, are decreasing functions of* η*. The worst-case profit,* $\pi_w(K^*_\eta)$*, is an increasing function of* η*.*

3 Business Training

We will apply the framework developed in the previous section to study the impact of a microentrepreneurs' training offered by business development programs (BDPs). The scope and level of training vary between different BDPs. In Karlan, Knight and Udry (2012), the training was on a small scale and involved targeted lessons such as keeping time and transaction records, separating business and personal money, etc. de Mel, McKenzie and Woodruff (2014), on the other hand, used the global Start-and-Improve Your Business (SIYB) training program. The SIYB is a program with an outreach of more than 4.5 million people in more than 95 countries. It involves 3 to 5 day training courses and covers topics such as organization of staff, record keeping and stock control, marketing and financial planning.

3.1 Effect of Business Training. Simple case

We introduce the training into the framework of Section 2 as follows. We assume that the microentrepreneur takes a training to learn about a new business practice, or a technology. There is no cost associated with taking the training and no cost associated with implementing the new practice. Thus it is always (weakly) optimal for the microentrepreneur to undertake the training. In what follows, we will use superscript *new* to refer to variables and functions related to the new practice. For example, $\pi^{new}(s, K)$ is the state-profit function under the new practice; K^{*new}_η is the capital level that maximizes the microentrepreneur's post-training objective function.

We assume that the business-training only affects the income from business activities and does not affect incomes from other activities, $h(s)$. Within this subsection, we will assume that the training has the following effect on state-profit functions $\pi^{new}(s, K) = \lambda_s\pi(s, K)$, where $\lambda_s \geq 1$. This assumption greatly simplifies the proofs of all of the propositions but is not necessary for the main message of the paper. In the next subsection, we will show what happens when the business-training is modeled in a more general fashion.

Given that profits are assumed to be non-negative, $\pi^{new}(s, K) > \pi(s, K)$ for every s and every K. We will refer to this property of the training as the *uniform profit improvement*. The uniform profit improvement assumption is very generous. It assumes that the training is so efficient that for any given state of the nature and any given capital investment decision the post-training profit is higher. This assumption is intentionally generous. By assuming away the profit decline due to training inefficiency, it allows us to focus on other factors that can be responsible for the post-training profit decline.

To capture BDP's focus on promoting business-oriented strategies, we assume that the effect of training is stronger in higher states, which are the states where capital is more productive. In other words, business training programs are relatively more efficient at teaching microentrepreneurs how to take full benefits of business-favorable states. Specifically, we assume that $1 \leq \lambda_1 < \cdots < \lambda_n$. For example, if trainees learn how to find cheaper suppliers or become more efficient at inventory management that will have stronger effect during good states when the demand is high.[8]

First, we look at the impact of training on microentrepreneurial utility. Given that profit functions are improved in every state and for every capital level, it is straightforward to show that the training has positive effect on microentrepreneur's welfare. Indeed,

$$U(K^*_\eta) < U^{new}(K^*_\eta) \leq U^{new}(K^{*new}_\eta). \tag{3.1}$$

Here, the first inequality is due to the fact that $\pi^{new}(s, K) > \pi(s, K)$ which implies $I^{new}(s, K) > I(s, K)$. The second inequality is due to the fact that K^{*new}_η is the optimal capital level for the post-training utility function. From welfare perspective, it indicates that the business training is valuable as it has a positive effect on microentrepreneurs' well-being regardless of its effect on expected profit. We summarize the reasoning above in Proposition 3.

Proposition 3 *The training strictly increases a microentrepreneurial utility.*

A consequence of Proposition 3 is that a microentrepreneur will always prefer the new practice and will always adopt it. That microentrepreneurs tend to follow, at least in the short-run, the practices they learn during the training course is well-documented in the literature. In Karlan, Knight and Udry (2012) the authors document that "*the consultants' recommendations were adopted for a time*" (p. 5). Similarly, in Karlan and Valdiva (2011) many people responded in a follow-up survey that they switched to the new practice. Table 8 in the McKenzie and Woodruff (2014)'s survey summarizes the effect of training on business practice adoption with the conclusion that "*almost all studies find a positive effect of business training on business practices*" (p. 67).

[8]Brooks et al. (2018) mention Prudence who was a participant of one the treatment (the mentor treatment) and who used to purchase inventory from suppliers at the entrance of a market area. After training she started to purchase at stalls deep into the market and only after comparing prices. Her cost dropped from 250 Ksh to 100 Ksh as a result, while she kept her sale price exactly the same. Clearly, a reduction in marginal cost has stronger effect during states that are favorable to the business activity.

Next, we study the impact of training on the expected profit. Since income from other sources, $h(s)$, does not depend on K, the effect of the training on expected income, $E_s I^{new}(s, K)$, is exactly the same as on the expected profit, $E_s \pi^{new}(s, K)$, as the difference between the two is a constant, $E_s h(s)$. In particular, if one increases (decreases) then so does another. The training has two effects on the expected profit:

$$E_s \pi^{new}(s, K_\eta^{*new}) - E_s \pi(s, K_\eta^*) = \underbrace{[E_s \pi^{new}(s, K_\eta^*) - E_s \pi(s, K_\eta^*)]}_{profit\ improvement} + \underbrace{[E_s \pi^{new}(s, K_\eta^{*new}) - E_s \pi^{new}(s, K_\eta^*)]}_{capital\ adjustment}. \tag{3.2}$$

The first effect is the *profit improvement effect*, which is $E_s \pi^{new}(s, K_\eta^{*new}) - E_s \pi(s, K_\eta^*)$. The uniform profit improvement ensures that it is always positive. If the microentrepreneur does not change capital investment, her expected profit goes up. The second effect is the *capital adjustment effect*, which is $E_s \pi^{new}(s, K_\eta^{*new}) - E_s \pi^{new}(s, K_\eta^*)$. Since K_η^* is no longer optimal, the microentrepreneur will change her capital investment to K_η^{*new}, which will affect the profit. Unlike the profit improvement effect, the capital adjustment effect can be either positive or negative. Having the negative capital adjustment effect limits the efficiency of the business-training. Whenever it is negative it means that the microentrepreneur does not take advantage of improved profitability but instead adjusts her investment in such a way that it hurts her expected profit.

There are two important benchmarks when the capital adjustment effect is guaranteed to be positive. The first case is when the microentrepreneur has only business-oriented ambition. In this case, the focus of business-training programs and entrepreneurs' preferences are aligned which is why the entrepreneurial choice of capital is guaranteed to increase expected profit. The proof is trivial since, by definition, K_0^{*new} maximizes $E_s \pi^{new}(s, K)$ and, therefore, the capital adjustment effect is positive.

Proposition 4 *If the microentrepreneur has only business-oriented ambitions, $\eta = 0$, the capital adjustment effect is non-negative.*

When $\eta \neq 0$, the BDPs' goal to promote business-oriented strategies does not align with the livelihood-oriented objective of the microentrepreneur. In general, this mismatch makes it possible for the capital adjustment effect to be negative since the microentrepreneur's objective is no longer to maximize the expected profits. However, one instance when the capital adjustment effect is guaranteed to be positive is when the business-income is the only income source for the microentrepreneur.

Proposition 5 *Assume that $\eta < 1$. When the microentrepreneur's only source of income is the business income the capital adjustment effect is positive and the total effect of the business-training are positive. When $\eta = 1$ the capital adjustment effect is non-negative and the total effect is positive.*

Results of Proposition 4 and Proposition 5 are very intuitive. BDPs are designed to promote business-oriented strategies such as business growth or production strengthening (Verrest, 2013). In our model, the training improves profitability of business activities ignoring their diversifying role with other sources of income as well as a possible lack of entrepreneurial ambition. When $\eta = 0$, the microentrepreneur's only objective is business oriented. The training's focus is aligned with microentrepreneur's objectives and profit goes up. When the microentrepreneur's enterprise is the only source of income then business income has no diversification role. The absence of the diversification aspect from the decision-making allows for microentrepreneur's objective and BDPs' focus to remain sufficiently aligned, even for positive η, so that the post-training profit goes up.

Consider now the case of multiple income sources. It turns out that as long as non-business income sources are sufficiently diversifying to provide a cushion against business-income shocks then the capital adjustment effect will be negative. Notably, it is well-documented that small and micro-entrepreneurs do not view their business activities in isolation as a way to bring in more money. Instead, they consider it either as a valuable diversification tool in order to deal with irregularity in income sources (Krishna, 2004); or to reduce household's vulnerability to negative shocks such as job-loss or illness (Ellis, 2000); or for consumption and income smoothing (Bateman and Chang, 2009; Banerjee and Duflo, 2011). To see what 'sufficiently diversifying' means, consider the case of $\eta = 1$. With business income being the only income source, state 1 is the worst state for every K. This is because $\pi(s, 0) = 0$ for every s and state 1 has the lowest marginal profitability of capital. Thus, the optimal worst-case capital $K_w = K^*(1)$. Investing more than $K^*(1)$ is suboptimal because it means lower profit in the worst-case state, which is state 1. Having non-business incomes allows for a possibility that $K_w > K^*(1)$. In this case, non-business incomes are sufficiently diversifying as they provide enough cushion even for the most pessimistic worst-case maximizing microentrepreneur to be willing to take more risk and invest more.

However, as the next Proposition shows if $\eta = 1$, $K_w < K^*$ and non-business incomes are sufficiently diversifying then the capital adjustment effect is negative. The intuition is as follows. When $K_w > K^*(1)$ state 1 is not necessarily the worst-case state. Due to diversifying effect of non-business income sources, for some values of K it is higher states that are worst-case states. Since the training has stronger impact on higher states, it effectively un-does the diversifying role of the non-business income sources. For every given K, the post-training worst-case state is lower than it was before the training. Lower states require lower investments, since the capital is less profitable, and, therefore, the livelihoods-oriented microentrepreneur invests less, $K_w^{new} < K_w$.

The microentrepreneur, instead of taking advantage of improved profitability by investing more, chooses to invest less. Whenever $K_w < K^*$ — i.e. the the microentrepreneur was already under-investing prior to the training — it results in the negative capital adjustment effect, thereby limiting the effectiveness of the the business-training and reducing its impact on the microentrepreneur's profit.

Proposition 6 *Assume that $\eta = 1$ and $K^*(1) < K_w \leq K^*$. Then the capital adjustment effect is negative.*

In order for $K_w \neq K^*(1)$ there must exist state s such that $\pi(1, K^*(1)) + h(1) > \pi(s, K^*(1)) + h(s)$. That is, non-business income should be large enough so that state 1 is no longer a uniformly worst-case state. In particular, it cannot be zero if conditions of Proposition 6 are to be satisfied. This is an important observation since it is common in literature to normalize the non-business income to zero, however, as this result shows ignoring the diversifying aspect of non-business incomes is not an innocuous assumption.

More generally, not only $\{h(s)\}$ cannot be zero but also it cannot be a monotonely increasing function of s either. Instead, it has to be eitehr a non-monotone or a decreasing function of s. In terms of empirical evidence, on the one hand, it is well-documented that many risks affecting income sources available to poor households, e.g. own-farm production and agricultural wage labor, exhibit a high correlation (Ellis, 2000, p. 60). Disastrous events, such as droughts, can adversely affect all income streams simultaneously. On the other hand, it is also well-documented that in much of the developing world, informal risk-sharing arrangements, which help coping with severe income fluctuations due to adverse factors, are widespread. (Ambrus et al., 2014). Risk-sharing is routinely mentioned as the most common way for household to deal with negative shocks including death, sickness, crime and court cases, shocks in income generating activities, etc. (De Weerdt and Dercon, 2006; Mazzucato, 2009). Moreover, they are substantial enough to be successful at smoothing households' consumption. While household income in developing countries varies greatly, consumption is remarkably smooth (Fafchamps and Lund, 2003). Finally, in addition to informal risk-sharing arrangements, governmental programs can provide some income boost in the disastrous state. For example, in Botswana the government drought relief program during 1985-86 created wage employment opportunities substituting for decreased share of livestock in crops in income portfolios (Valentine, 1993).

As an example of a non-monotone or a decreasing $\{h(s)\}$, assume that the only non-business income sources are income from informal risk-sharing arrangement with distant family members or other villagers, and employment income. State 1 is a bad state for the microentrepreneur so that if state 1 is realized she receives some help through her risk-sharing arrangement, $h(1) > 0$. State 2 is an intermediate state where the business is profitable enough, so that no help from risk-sharing is needed, but there are no employment opportunities, $h(2) = 0$, and state 3 is a good state where the microentrepreneur receives positive labor income, $h(3) > 0$. Another example would be where if the only non-business income source is the governmental subsidy via disaster relief programs, where state 1 is the disaster state. Then one would have $h(1) > 0$ and $h(2) = \cdots = h(n) = 0$. Finally, it can be that the microentrepreneur receives net payments from a risk-sharing arrangement in her bad states, so that $h(s) > 0$ for low states, and contributes net payments in good states, so that $h(s) < 0$ in high states. This example is based on the empirical finding that risk-sharing mechanisms are very efficient at smoothing households consumption. Furthermore, as in the first example, one can add employment income as an additional income source. As long as it is not available or is very low in low states, the income from non-business sources will be a non-monotone function of s.

We conclude this section by looking at how the BDP's effect changes with η. Given that

Proposition 6 was proved for $\eta = 1$, and when $\eta = 0$ the capital adjustment effect is always positive, a natural conjecture would be that the capital adjustment effect is lower when η is higher. However, in general, this is not the case. When $\eta > 0$ the livelihoods-oriented ambition pushes the microentrepreneur's capital choice away from profit-maximizing level. The training can undo this effect which will result in it having a larger impact for intermediate values of η. To see that consider the limiting case when $\lambda_1 = \cdots = \lambda_{n-1} = 1$ and $\lambda_n = \infty$. Then the post-training choice of capital will not be sensitive to changes in η so that $K_\eta^* = K^*(n)$ and the post-training profit will not change. The pre-training profit, however, is a decreasing function of η. Then the total effect of the training, $E_s\pi^{new}(s, K_\eta^{*new}) - E_s\pi(s, K_\eta^*)$, is an increasing function of η.

At the same time, the conjecture that the training's impact should be lower for higher η is not entirely incorrect. As long as λ_n/λ_1 is not too large, so that the extreme example above is not applicable, then the total effect is a decreasing function of η. Furthermore, for any given training, the total effect becomes a decreasing function when η is large enough. That is, eventually as ambiguity-aversion becomes higher, the effect of training on profit becomes lower.

Proposition 7 *The total effect of training on expected profit is*

i) not necessarily decreasing function of η;

ii) for any $\{\lambda_s\}$, there exists $\eta^0 < 1$ such that it is a decreasing function of η for any $\eta > \eta^0$;

iii) there exists $\Lambda^0 > 1$ such that if $\lambda_n/\lambda_1 < \Lambda^0$ the total effect is a decreasing function of η.

3.2 Business Training. A General Case

In the previous subsection we assumed specific functional form of profit-improvement, $\pi^{new}(s, K) = \lambda_s \pi(s, K)$. The main advantage of this assumption was its simplicity and the fact that $K^{*new}(s) = K^*(s)$, which simplified most of the proofs. A natural question is how robust the results from the previous subsection are to alternative ways of modeling the business-training. We address this question in this subsection.

As before, we assume that the business-training only affects the income from business activities and does not affect incomes from other activities, $h(s)$. We also assume that post-training state-profit functions are single-peaked, concave functions of K such that $\pi^{new}(s, K) > \pi(s, K)$ for every s and every K. In the previous section we referred to the last assumption as the *uniform profit improvement assumption.* Finally, we will continue to assume that BDPs' focus on promoting business-oriented strategies has stronger effect in higher states, i.e. those states where capital is more productive, and for higher levels of capital.

We will consider two types of profit improvements: an *additive* improvement and a *multiplicative* improvement. Under the additive improvement, the post-training state-profit functions satisfy concavity, single-peakedness the uniform profit improvement, and are such that

$$(\pi^{new}(t, K) - \pi(t, K))' \geq (\pi^{new}(s, K) - \pi(s, K))' \geq 0 \text{ when } t > s, \tag{3.3}$$

and $\pi^{new}(t,0) - \pi(t,0) \geq \pi^{new}(s,0) - \pi(s,0)$ when $t > s$. Notice that for reasons discussed earlier, the improvement is stronger in higher states (the first inequality in (3.3)) and for higher levels of K (the second inequality in (3.3)).

One example of the additive improvement is $\pi^{new}(s,K) = \lambda_s + \pi(s,K)$ where $\lambda_n \geq \cdots \geq \lambda_1 \geq 0$. Another example of the additive improvement is for the case when the pre-training state-profit function is $\pi(s,K) = F(s,K) - RK$, where $F(s,K)$ is a standard production function such that $F''_{sK} > 0$, so that capital is more productive in higher states. The post-training state-profit function is $\pi^{new}(s,K) = \lambda_s F(s,K) - RK$. That is, training increases capital's productivity. In this case, condition (3.3) becomes

$$(\lambda_t - 1)F'(t,K) \geq (\lambda_s - 1)F'(s,K) \geq 0.$$

It is satisfied if $\lambda_n \geq \cdots \geq \lambda_1 \geq 1$.

Under the multiplicative improvement, the new state-profit functions satisfy concavity, single-peakedness, the uniform profit improvement, and are such that

$$\frac{\pi^{new}(s,K)}{\pi(s,K)} = g(s)\frac{\pi^{new}(1,K)}{\pi(1,K)}, \tag{3.4}$$

where $g(s)$ is a weakly increasing function of s, and $\dfrac{\pi^{new}(s,K)}{\pi(s,K)}$ is a weakly increasing function of capital for every s. Again, by design the business-training is weakly more efficient in higher states and for higher levels of K. One example of the multiplicative improvement is $\pi^{new}(s,K) = \lambda_s \pi(s,K)$, where $\lambda_s \geq 1$. Requirement that $g(s)$ is a weakly increasing function of s is satisfied if $\lambda_n \geq \cdots \geq \lambda_1$.

It turns out that when the effect of training is modeled in a more general fashion it is still the case than when the microentrepreneur is risk-neutral or has only source of income the training will have positive effect on profit. The two Propositions are exact analogues of Propositions 4 and 5.

Proposition 8 *If the microentrepreneur has only business-oriented ambitions, i.e.* $\eta = 0$*, then capital adjustment effect is non-negative, and post-training expected profit strictly increases.*

Proposition 9 *Assume that* $\eta < 1$*. When the microentrepreneur's only source of income is the business income the capital adjustment effect is non-negative and the total effect of the business-training are positive. When* $\eta = 1$ *the capital adjustment effect is non-negative and the total effect is positive.*

While the proof of Proposition 9 is more technical, it follows the exact same steps as that of Proposition 5. The intuition behind both Propositions is the same as before. When $\eta = 0$, the focus of business-training programs and entrepreneurs' preferences are aligned which is why the training is guaranteed to have a positive effect on profit. When the microentrepreneur's enterprise is the only source of income then business income has no diversification role, and microentrepreneur's objective and BDPs focus are sufficiently aligned for the post-training profit to go up.

Next, we look at conditions under which the capital adjustment is negative. Just like in a simpler case of Proposition 6, it is necessary that $K^*(1) < K_w$. Non-business incomes sources must be substantial enough and must provide enough diversification to the business-income so that the livelihoods-oriented entrepreneur is willing to invest above $K^*(1)$. However, in a more general setting of this section this is not enough. Since $K^{*new}(s)$ is no longer equal to $K^*(s)$ it is possible to have a training so efficient at improving capital's marginal profitability that it results in $K^{*new}(s) > K_w$ for every s so that $K_w^{new} > K_w$. Then the capital adjustment effect is positive. Thus, we need to impose an additional restriction on how much the training can improve capital's marginal profitability. The last condition of Proposition 10 is an analogous to the assumption $\lambda_1 < \cdots < \lambda_n$ from the previous subsection.

Proposition 10 *Let $\eta = 1$ and $K^*(1) < K_w \leq K^*$. Let s_w be the lowest pre-training worst-case state given K_w, and let $K^{*new}(s_w) < K_w$. If either all inequalities in (3.3) are strict, or if $g(s)$ is strictly increasing function, then the capital adjustment effect is negative.*

3.3 An Example of Post-Training Profit Decline

Negativity of capital adjustment effect undermines the efficiency of the business-training. Whenever the capital adjustment effect is negative it means the microentrepreneur does not fully reap the benefits of the training. Instead of taking advantage of improved profitability and expanding her enterprise by investing more, the microentrepreneur finds it safer to invest less than before thereby invalidating the training's effect. A natural question then is how big this effect is. It turns out that it is, in fact, possible for the negative capital adjustment effect to be large enough to outweigh the profit improvement effect and make the total effect negative. In other words, as documented in empirical literature, it is possible to observe post-training decline in the expected profit even though the training improves microentrepreneur's profit functions for every s and every K.

Consider the following example. There are six states, each of which is equally likely, $p_s = 1/6$. The microentrepreneur has three income sources: business income, labor income, and income due to informal risk-sharing or insurance arrangements. The microentrepreneur has endowment of labor normalized to 1. Labor can be used to generate business and employment incomes. The microentrepreneur can also borrow a capital at rate R. The capital can be only used for business activities. If the microentrepreneur splits the labor between employment and business as $(1-L, L)$ and invests capital K, then her business income in state s is given by $sF(K, L) - RK = s(\sqrt{K} + \sqrt{L}) - RK$, and her employment income in state s is $w(1-L)$. We assumed that $w_s = w$ so that wages are neither positively nor negatively correlated with the state of nature.[9] Finally,

[9]Depending on circumstances, employment and business incomes can be positively or negatively correlated. For example, own-farm production and agricultural wage labor will exhibit a high correlation. At the same time, Verrest (2013) documents that many households use business activities as a diversification tool against possible negative

the microentrepreneur receives transfers from informal risk-sharing arrangements, A_s. Transfers depend on realized state of nature s. We assume that the net informal insurance payment is zero, $E_s A_s = 0$, and that $A_s = A > 0$ when $s \leq 3$, and $A_s = -A < 0$ when $s \geq 4$.[10]

The timing is as follows. First, the microentrepreneur chooses K, then the state of nature is realized, then the microentrepreneur optimally allocates labor between the employment and her enterprise. Thus, we assume that the choice of labor is flexible and can be adjusted given the realized state of nature. Capital investment, on the other hand, is inherently risky, as it is made before the uncertainty is realized.

Thus, given state realization s, and the microentrepreneur's choice of L and K the microentrepreneur's income is

$$I^{ex-post}(s, L, K) = w(1 - L) + s(\sqrt{K} + \sqrt{L}) - RK + A_s.$$

We then define $I(s, K)$ as $\max_L I^{ex-post}(s, L, K)$. The post-training income is defined as

$$I^{ex-post}(s, L, K) = w(1 - L) + s\lambda_s(\sqrt{K} + \sqrt{L}) - RK + A_s.$$

Consider the following numerical example. Let $R = 0.7$, $w = 1$, $A = 3$, and the values of λ's are such that $\lambda_1 = 1$, $\lambda_2 = 1.02$, $\lambda_3 = 1.05$, $\lambda_4 = 1.15$, $\lambda_5 = 1.16$ and $\lambda_6 = 1.17$. As one can verify, the training, despite its uniform profit improvement, will result in lower expected post-training profit.

labor shocks: *"They [home-based economic activities] may provide savings in the form of cash or kind as "an apple for a rainy day" when other incomes disappear because jobs are lost or people fall ill."* (p. 64). Thus, in the example section we do not take either side and simply assume that there is no correlation between labor and business incomes. By continuity, our example will continue to hold for small values of positive and negative correlation between wages and business incomes.

[10] An implicit assumption here is that informal insurance payments cannot be used for investment, which is generally consistent with empirical evidence that shows most common reason for accepting such payments is to meet immediate consumption needs rather than investment purposes. Only 3.8% of all gifts and 18.4% of informal loans are used for investment purposes, mostly schooling (Fafchamps and Lund, 2003).

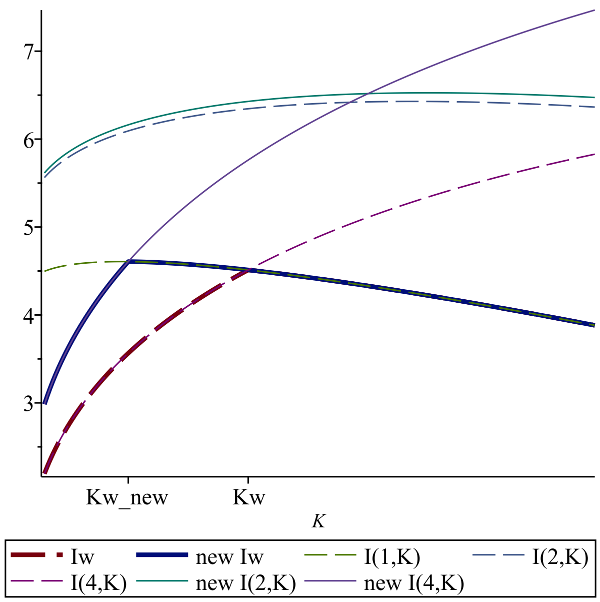

Figure 1: $I(s, K)$ and $I_w(K)$ before and after the training. Profit in state 1 did not change; profits in all other states went up. The capital adjustment effect is negative $K_w^{new} < K_w < K^*$.

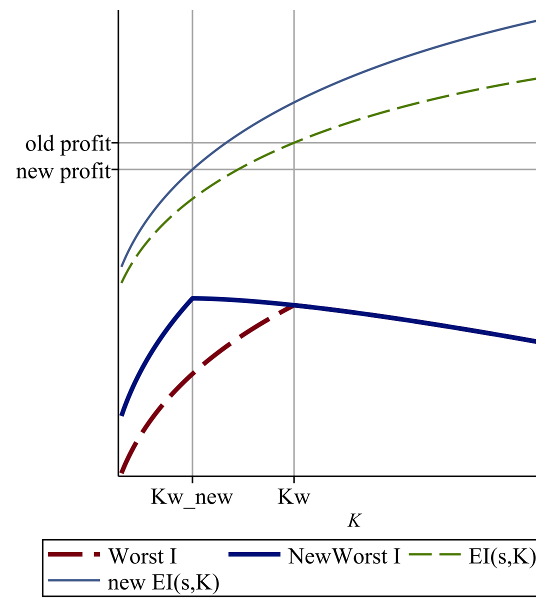

Figure 2: $I_w(K)$ and $EI(s, K)$ before and after the training: $EI^{new}(s, K_w^{new}) \approx 6.49$ is less than $EI(s, K_w) \approx 6.51$. Expected profts are $E\pi(s, K_w^{new}) \approx 6.37$ and $E\pi(s, K_w) \approx 6.39$.

Figures 1 and 2 visualize the example. $I(s, K)$ for states 3, 5 and 6 are not plotted to keep Figure 1 simple. For the microentrepreneur with the livelihoods-ambition, the two relevant states are states 1 and 4. For states with positive safety-net payments, states $s = 2$ and 3 dominate state 1; for states with negative safety-net payments, states $s = 5$ and 6 dominate state 4. The only worst-case states in this example are states 1 and 4. Training makes state 4 more profitable so that $K_w^{new} < K_w$ and, therefore, the capital adjustment effect is negative. The exact values are $K_w \approx 1.17$ and $K_w^{new} \approx 0.54$. In this example, the negative capital adjustment effect dominates the profit improvement effect. Post-training expected income $E_s I^{new}(s, K_w^{new}) \approx 6.49 < E_s I(s, K_w) \approx 6.51$. By design, expected income from business activity declines as well: $E_s \pi^{new}(s, K_w^{new}) \approx 6.37 < E_s \pi(s, K_w) \approx 6.39$.

We can modify the example above to see for which parameter values the capital adjustment effect is negative and for which parameter values it is so high that the total effect on the profit is negative. We see how things change as we vary η, p and λ. As there are six states in our example, we have to define p and λ in a one-dimensional way. We do it as follows. p is probability of state 1. States 2 through 6 have probabilities $(1-p)/5$. When $p = 1/6$ all states are equally likely. As for the training, its effect in state s given λ is defined as $1 + (\lambda_s - 1)(\lambda - 1)$, where $\{\lambda_s\}$ are parameters used to build Figure 1 and 2. When $\lambda = 1$ then the training has no effect. When $\lambda = 2$ then the effect of the training in state s is exactly λ_s.

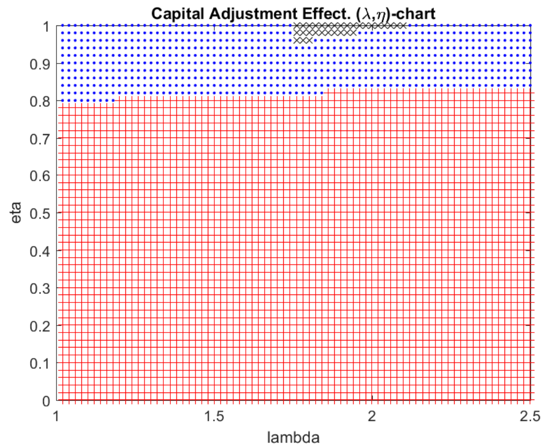

Figure 3: The figure is built for $p_s = 1/6$. Black circles correspond to area where the post-training profit declines. Blue dots correspond to the area where the capital adjustment effect is negative but the total effect is positive. Red '+' signs correspond to the area where the capital adjustment is positive.

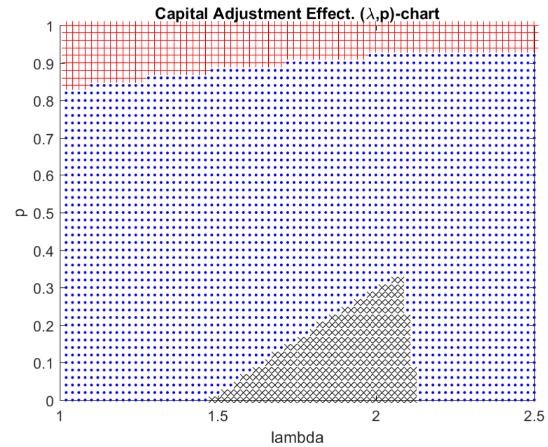

Figure 4: The figure is built for $\eta = 1$. Black circles correspond to area where the post-training profit declines. Blue dots correspond to the area where the capital adjustment effect is negative but the total effect is positive. Red '+' signs correspond to the area where the capital adjustment is positive.

From Figure 3 one can see that a high-degree of ambiguity-aversion is needed in order to have the negative capital adjustment effect. To have the total effect negative, one need to have a very high η and an intermediate range of λ's. When η is set equal to 1, as on Figure 4 then the capital adjustment effect is negative for almost all parameter values. The total effect is negative for low values of p, and intermediate values of λ's.

4 Concluding Remarks

Most experts agree that there is a need to improve business knowledge among small- and micro-entrepreneurs as entrepreneurs in developing countries are often unaware of even the basic business practices such as keeping personal and business finances separate, or keeping records of their transactions and inventory. The business training programs aimed at microentrepreneurs around the world date back as far as the seventies. Yet, the effect of these training programs is not as strong as one would hope. Dar and Tzannatos (1999), as well as an updated review by Betcherman et al. (2004), find that the impact of training programs has wide variation with some programs demonstrating a positive effect, while others having no effect or even a negative effect. In a more recent work, McKenzie and Woodruff (2014) reach a similar conclusion, though they argue that a lack of impact on profit could be caused by methodological issues such as sample size or heterogeneity.

This paper provides a theoretical framework to understand a mixed impact of business-training. We rely on a holistic view of a microentrepreneur as someone whose livelihood and ambitions are more complex than just being an entrepreneur. First, the microentrepreneur has several sources of income in addition to income from business activities. Second, the microentrepreneur has other goals in addition to maximizing total income. The impact of business training, however, is narrow in a sense that it only impacts income from business activities.

We show the training effect is heterogeneous and depends on microentrepreneur's ambition. This is consistent with the observation that "*BDPs have been more successful for some entrepreneurs than the others*" (Verrest, 2013, p. 58). For microentrepreneurs with strong profit-maximizing ambition, the training effect is positive. For other microentrepreneurs, however, it is possible that adoption of new business practices will have much lower, if any, impact on their profits. We further show that the reason behind the limited effect of business trainings is the BDPs' focus on growing microentrepreneur's business that does not take into account a wider context in which the microentrepreneur operates. When the microentrepreneur has other goals beyond profit-maximization, and other income sources beyond business income, the training impact can be limited and even negative.

We also develop a behavioral explanation of how training can have negative impact on the expected profit. We use the perceived ignorance hypothesis to suggest another source that can undermine the training's efficiency: a higher ambiguity of the new business practice. If the new practice is viewed as more ambiguous than the old one then it can either prevent the switch to the new practice or can lead to lower expected profit after the switch. We use this to explain why a simplistic rule-of-thumb training in Drexler et al. (2014) worked better than a more complex one.

Overall, our paper highlights the importance of having a holistic view on microfinance clients. While not a novel insight for empirical literature, to the best of our knowledge ours is one of the first theoretical papers that shows the importance of the holistic view and how using it affects the model's outcomes.

5 Appendix. Proofs

Proof of Proposition 1: If $K_w = K^*(1)$, we are done. Assume now that $K_w \neq K^*(1)$. Let $S_w(K)$ denote the set of all worst states given K, $S_w(K) = \{s : I(s,K) \leq I(t,K) \text{ for all } t \neq s\}$. One can show that if $K_w \neq K^*(1)$ then $S_w(K_w)$ has at least two elements in it. Proof by contradiction. Assume to the contrary that $S_w(K_w)$ has exactly one element, s'. Then, on one hand, $K_w \neq K^*(s')$. This is because

$$I(s', K^*(s')) > I(1, K^*(1)) > I(1, K^*(s')) \geq \min_s I(s, K^*(s')). \tag{5.1}$$

The first inequality is by the *better*-assumption. The second inequality follows from $K^*(s') \neq K^*(1)$, and single-peakedness of $I(1, \cdot)$. On the other hand, $K_w = K^*(s')$. Indeed, by continuity, s' is the unique worst-case states in the neighborhood of K_w. Then, $I(s', K_w)$ can be neither strictly increasing nor strictly decreasing at K_w. Otherwise, K just above (or just below) K_w would deliver higher worst-case profit. Thus, $K_w = K^*(s')$. We reached a contradiction. Therefore, s' is not unique state given K_w.

Let s be the lowest and s' be the highest states in $S_w(K_w)$. One can apply the same reasoning as in (5.1) to show that $I'_K(s, K_w) \neq 0$ and $I'_K(s', K_w) \neq 0$. Furthermore, it cannot be the case that both state-profit functions are increasing (decreasing). Indeed, if, for example, $I'_K(s, K_w) > 0$ and $I'_K(s', K_w) > 0$ then $K_w < K^*(s) < K^*(s')$. State s is the smallest worst-case state given K_w. Therefore, by the complementarity for every $s'' \in S_w(K_w)$, the corresponding state-profit function is increasing at K_w: $I'_K(s'', K_w) > 0$. By continuity, in a sufficiently small neighborhood of K_w, only states from S_w can be the worst states.[11] But then, K slightly above K_w will result in a higher worst-case profit. Similarly, it cannot be the case that $I'_K(s, K_w) < 0$ and $I'_K(s', K_w) < 0$.

Thus, $I'_K(s, K_w)$ and $I'_K(s', K_w)$ have different signs and neither is equal to zero. By the complementarity assumption, it has to be the case that $I'_K(s, K_w) < 0 < I'_K(s', K_w)$, which completes the proof. ■

Proof of Proposition 2: We will prove the proposition for the case $K_w < K^*$ only. The case $K_w > K^*$ is similar.

Since $U(K)$ is a concave function of K, it has left and right derivatives and the left derivative is greater or equal than the right derivative. For a given η, the necessary and sufficient condition for K^*_η to maximize $U(K)$ is

$$(1-\eta)E_s\pi'_K(s, K^*_\eta) + \eta \cdot (\pi_w(K^*_\eta-))'_K \geq 0 \geq (1-\eta)E_s\pi'_K(s, K^*_\eta) + \eta \cdot (\pi_w(K^*_\eta+))'_K. \tag{5.2}$$

If the utility function is differentiable at K^*_η, then (5.2) becomes

$$(1-\eta)E_s\pi'_K(s, K^*_\eta) + \eta \cdot (\pi_w(K^*_\eta))'_K = 0. \tag{5.3}$$

By concavity, when $K^* > K_w$ then $K^* \geq K^*_\eta \geq K_w$ for every η. This is because when $K > K^*$ ($K < K_w$), both the expected profit and the worst-case profit have negative (positive) left and right derivatives.

We can now prove that K^*_η is a decreasing function of η.[12] In the proof, we will use the fact that $(\pi_w)'_K(K^*_\eta+) \leq (\pi_w)'_K(K^*_\eta-) < 0$ when $K^*_\eta > K_w$, and $E_s\pi'_K(s, K^*_\eta) > 0$ when $K^*_\eta < K^*$. The proof is by contradiction. Assume not. Then there exist $\eta_1 < \eta_2$ such that $K^*_{\eta_1} < K^*_{\eta_2}$. That means

[11] For every $t \notin S_w(K_w)$ and every $t' \in S_w(K_w)$, it is the case that $I(t, K_w) > I(t', K_w)$. Then, for any K sufficiently close to K_w, it is also the case that $I(t, K) > I(t', K)$. Therefore, $t \notin S_w(K_w)$ cannot be the worst-case state for K that are sufficiently close to K_w.

[12] It is not a strictly decreasing function of η. When the worst-case state is unique, as in (5.3), it is a strictly increasing function of η. When it is not unique, as in (5.2), it is weakly decreasing. That is, there is a range of η's that would correspond to the same optimal K^*_η.

that

$$
\begin{aligned}
0 &\geq (1-\eta_1)E_s\pi'_K(s,K^*_{\eta_1}) + \eta_1\cdot(\pi_w(K^*_{\eta_1}+))'_K \\
&> (1-\eta_1)E_s\pi'_K(s,K^*_{\eta_2}) + \eta_1\cdot(\pi_w(K^*_{\eta_2}-))'_K \\
&> (1-\eta_2)E_s\pi'_K(s,K^*_{\eta_2}) + \eta_2\cdot(\pi_w(K^*_{\eta_2}-))'_K,
\end{aligned}
$$

which, given (5.2), means that $K^*_{\eta_2}$ cannot be optimal given η_2. Here, the first inequality follows from the fact that $K^*_{\eta_1}$ is optimal given η_1, see (5.2); the second inequality follows from the fact that the utility function is concave and $K^*_{\eta_1} < K^*_{\eta_2}$; the last inequality follows from the fact that $\eta_2 > \eta_1$ and that the derivative of π_w is negative (because $K^*_{\eta_2} > K^*_{\eta_1} \geq K_w$), while the derivative of the expected profit is positive.

The rest of the proposition is straightforward. Given that the worst-case profit $\pi_w(\cdot)$ is a concave function that reached its maximum at K_w, given that $K^*_\eta \geq K_w$ and that K^*_η is a decreasing function of η, we can conclude that the worst-case profit, $\pi_w(\cdot)$, is an increasing function of η. Similarly, given that the expected profit, $E_s\pi(s,\cdot)$, is a concave function with the maximum at K^*, given that $K^* \geq K^*_\eta$, and that K^*_η is a decreasing function of η, we can conclude that the expected profit is a decreasing function of η.

Finally, the utility funciton is a weakly decreasing function of η. Consider $\eta_1 < \eta_2$ and let $K^*_{\eta_1}$ and $K^*_{\eta_2}$ be corresponding optimal capital levels. Then

$$
\begin{aligned}
(1-\eta_1)E_s\pi(s,K^*_{\eta_1}) + \eta_1\pi_w(K^*_{\eta_1}) &\geq (1-\eta_1)E_s\pi(s,K^*_{\eta_2}) + \eta_1\pi_w(K^*_{\eta_2}) \\
&> (1-\eta_2)E_s\pi(s,K^*_{\eta_2}) + \eta_2\pi_w(K^*_{\eta_2}).
\end{aligned}
$$

The first inequality follows from the fact that $K^*_{\eta_1}$ is optimal given η_1. The second inequality follows from the fact that the expected profit is greater than the worst-case profit, $E_s\pi(s,K) > \pi_w(K)$, and that $\eta_1 < \eta_2$. This completes the proof of the proposition. ■

Proof of Proposition 5: We prove it for the case of $\eta < 1$. The case $\eta = 1$ follows from continuity and the fact that the profit improvement effect is always strictly positive. When the microentrepreneur has a single income source, her income is equal to her profit. Since $\pi(1,0) = \cdots = \pi(n,0) = 0$ the complementarity implies that before training state 1 is the worst-case state for every K. Since $\lambda_1 < \cdots < \lambda_n$ state 1 remains to be the worst-case state after training for every K. The microentrepreneur's utility therefore is $(1-\eta)E_s\pi(s,K) + \eta\pi(1,K)$.

We prove that $K^*_\eta < K^{*new}_\eta \leq K^{*new}$ which then would guarantee that the capital adjustment effect is positive. First, $K^*_\eta < K^*$ for every $\eta > 0$. The FOC is $(1-\eta)E_s\pi'(s,K^*_\eta) + \eta\pi'(1,K^*_\eta) = 0$. By complementarity $K^*_\eta \geq K^*(1)$. This is because $\pi'(n,K) > \ldots \pi'(1,K) \geq 0$ when $K < K^*(1)$. Thus $K < K^*(1)$ would not satisfy the FOC. If $\eta = 1$ then $K^*_\eta = K^*(1)$ which is clearly less than K^*. When $\eta < 1$ then $\pi'(1,K^*_\eta) < 0$ which implies that $E_s\pi'(s,K^*_\eta) > 0$ and, therefore, $K^*_\eta < K^*$. Similarly, one can show that $K^{*new}_\eta < K^{*new}$.

To establish that $K^*_\eta < K^{*new}_\eta$ notice that derivative of the post-training utility function at

K^*_η and notice that it is positive. Let τ be the lowest state such that $\pi'(s, K^{*new}_\eta) > 0$. By complementarity all states higher than τ have positive profit derivative at K^{*new}_η and all states lower than τ have non-positive profit derivative. Then,

$$\begin{aligned}(1-\eta)E_s\lambda_s\pi'(s,K^*_\eta)+\eta\lambda_1\pi'(1,K^*_\eta) &= (1-\eta)E_s(\lambda_s-\lambda_1)\pi'(s,K^*_\eta)\geq \\ &\geq (1-\eta)(\lambda_\tau-\lambda_1)E_s\pi'(s,K^*_\eta)>0.\end{aligned}$$

The last inequality follows from $\lambda_1 < \cdots < \lambda_n$ and that $K^*_\eta < K^*$. This completes the proof since $K^*_\eta < K^{*new}_\eta \leq K^{*new}$ implies that the capital adjustment effect is positive. ■

Proof of Proposition 6: If $K_w \neq K^*(1)$ there exist two states, $s < s'$ such that $\pi(s, K_w) = \pi(s', K_w) \leq \pi(t, K_w)$ for all other states t, and $\pi'_K(s, K_w) < 0 < \pi'_K(s', K_w)$. Since $\lambda_1 < \cdots < \lambda_n$, it follows that $\lambda_s\pi(s, K_w) < \lambda_t\pi(t, K_w)$ for every $t > s$. Thus, the new worst-state given K_w, denote it s^{new}_w, is less or equal than s, $s^{new}_w \leq s$. From complementarity and $s^{new}_w \leq s$ follows that $\pi(s^{new}_w, K)$ is decreasing at K_w. Therefore, $K > K_w$ cannot be the new optimal worst-case capital: $\min_t \lambda_t\pi(t, K) \leq \lambda_{s^{new}_w}\pi(s^{new}_w, K) < \lambda_{s^{new}_w}\pi(s^{new}_w, K_w) = \min_t \lambda_t(t, K_w)$. Moreover, K_w is no longer the optimal worst-case capital either. Since all worst-case states are less or equal than s it means that the corresponding state-profit functions are decreasing at K_w. Then K slightly below K_w will give strictly higher worst-case profit. It proves that $K^{new}_w < K_w$.

Next, we study the impact of training on K^*. The post-training expected profit is single-peaked and concave. Therefore, to show $K^{*new} > K^*$ it is sufficient to show that $E_s\lambda_s\pi'_K(s, K_w) > 0$. Let τ be the largest state such that $\pi'_K(\tau, K^*) < 0$. By weak complementarity τ is the state such that $K^*(\tau) < K^* \leq K^*(\tau + 1)$. Then

$$\begin{aligned}E_s\lambda_s\pi'_K(s,K^*) &= \sum_{s\leq\tau}\lambda_s p(s)\pi'_K(s,K^*)+\sum_{s>\tau}\lambda_s p(s)\pi'_K(s,K^*) > \\ &> \lambda_\tau\sum_{s\leq\tau}p(s)\pi'_K(s,K^*)+\lambda_{\tau+1}\sum_{s>\tau}p(s)\pi'_K(s,K^*)>0.\end{aligned}$$

Here I used that the first sum is the summation of negative terms and the second sum is the summation of positive terms, that $\lambda_\tau < \lambda_{\tau+1}$ and that K^* satisfies the first-order condition: $E_s\pi'_K(s, K^*) = 0$.

Thus, on the one hand, $K^{new}_w < K_w$ and on the other hand $K^* < K^{*new}$, and by assumption $K_w \leq K^*$. But then $E_s\pi(s, K^{new}_w) - E_s\pi(s, K_w) < 0$ which means that the capital adjustment effect is negative. ■

Proof of Proposition 7 : Step 1. We show that for any $K > K_w$ ($K > K^{new}_w$) the worst-case state is state 1 for pre-training (post-training) state-profit functions.

For the pre-training profit if $K_w = K^*(1)$ then by complementarity state 1 is the worst state when $K > K_w$, and we are done. If $K_w > K^*(1)$ then there are at least two worst-case states given K_w. The lowest worst-case state must be 1. Assume not. Assume it is $s > 1$. By Proposition 1, $\pi(s, K_w)$ is declining at K_w, i.e. $K^*(s) < K_w$. By the better assumption and complementarity:

$\pi(1, K^*(s)) < \pi(1, K^*(1)) < \pi(s, K^*(s))$. Since derivative of $\pi(1, K)$ is less than the derivative of $\pi(s, K)$ and since $K^*(s) < K_w$ it means that $\pi(1, K_w) < \pi(s, K_w)$. Thus state s cannot be the worst state for $K \geq K_w$. Contradiction.

For the post-training case the proof is slightly different because state-profit functions might not satisfy complementarity anymore. Nonetheless, since $\lambda_1 < \cdots < \lambda_n$ they still satisfy the better-assumption and $K^{*new}(s) = K^*(s)$ for any s. First, $K_w^{new} \geq K^*(1)$. Otherwise, all state-profit functions would be strictly increasing at K_w^{new} and, therefore, K_w^{new} cannot be optimal worst-case capital. K_w^{new} cannot be equal to $K^*(s)$ for $s > 1$ because from $\lambda_1 < \lambda_s$ and the better-assumption follows that $\pi^{new}(1, K^*(s)) < \pi^{new}(s, K^*(s))$. Then by the same logic as in Proposition 1 there must be at least one worst-case state, s^0, given K_w^{new} such that the corresponding state profit-function is declining at K_w^{new}. Then $K^{*new}(s^0) < K_w^{new}$. Assume to the contrary that $s^0 > 1$. For the pre-training state-profit functions, $\pi(1, K^*(s^0)) < \pi(s^0, K^*(s^0))$ and $\pi(1, K) < \pi(s^0, K)$ for any $K > K^*(s^0)$. Since $\lambda_1 < \lambda_{s^0}$ it means that the same must hold for the post-training state-profit function which means state $s^0 > 1$ cannot be the worst-case state at K_w^{new}.

Step 1 implies that for the pre- and post-training utility function is a weighted average of expected utility and utility at state 1, $(1-\eta)E_s U(s, K) + \eta U(1, K)$ when $K \geq K_w$ (or $K \geq K_w^{new}$). Furthermore, it is differentiable everywhere except for $K = K_w$ ($K = K_w^{new}$) if $K_w \neq K^*(1)$.

Step 2: We show that if $K_\eta^* \neq K_w$ ($K_\eta^{*new} \neq K_w^{new}$) are decreasing functions of η if . When $K_\eta^* \neq K_w$ the pre-training utility function is differentiable and $K_\eta^* > K_w$ and satisfies the FOC. By the implicit function theorem:

$$\frac{\partial K_\eta^*}{\partial \eta} = -\frac{-E_s \pi'(s, K_\eta^*) + \pi'(1, K_\eta^*)}{(1-\eta)E_s \pi''(s, K_\eta^*) + \eta \pi''(1, K_\eta^*)} = -\frac{1}{1-\eta} \frac{\pi'(1, K_\eta^*)}{(1-\eta)E_s \pi''(s, K_\eta^*) + \eta \pi''(1, K_\eta^*)} < 0. \quad (5.4)$$

Here we use that $I'(s, K) = \pi'(s, K)$ since $h(s)$ does not depend on K. Then the numerator is negative because $K_\eta^* > K^*(1)$, and the SOC the denominator is negative by concavity of profit functions.

Similarly, for the post-training profit

$$\frac{\partial K_\eta^{*new}}{\partial \eta} = -\frac{1}{1-\eta} \frac{\lambda_1 \pi'(1, K_\eta^{*new})}{(1-\eta)E_s \lambda_s \pi''(s, K_\eta^{*new}) + \eta \lambda_1 \pi''(1, K_\eta^{*new})} < 0. \quad (5.5)$$

Step 3: If K_η^* is such that both pre- and post-training profits are differentiable then the derivative of the total effect on the expected profit with respect to η is

$$(E_s U^{new}(s, K_\eta^{*new}) - E_s U(s, K_\eta^*))'_\eta = \frac{\partial K_\eta^{*new}}{\partial \eta} E_s \lambda_s \pi'_K(s, K_\eta^{*new}) - \frac{\partial K_\eta^*}{\partial \eta} E_s \pi'_K(s, K_\eta^*). \quad (5.6)$$

From (5.5) one can see that for a given $\eta < 1$, if λ's are sufficiently high then one can make $\partial K_\eta^{*new}/\partial \eta$ to be arbitrarily close to zero. In this case the first term in (5.6) is arbitrarily close to zero while the second term is positive. This proves i), which is that it is possible for the total effect to be an increasing function of η.

Now we prove existence of η^0 from the statement of the theorem. Without loss of generality assume that $\lambda_1 = 1$, otherwise divide (5.4) and (5.5) by λ_1 and define $\hat{\lambda}_s = \lambda_s/\lambda_1$. We need to consider two cases. First, $K_w = K^*(1)$ in which case pre- and post-training utilities are differentiable everywhere. Then at $\eta = 1$

$$0 > \frac{\partial K_\eta^*}{\partial \eta} = \frac{E_s \pi'(s, K^*(1))}{\pi''(1, K^*(1))} > \frac{E_s \lambda_s \pi'(s, K^*(1))}{\pi''(1, K^*(1))} = \frac{\partial K_\eta^{*new}}{\partial \eta}.$$

Plug the expression for derivatives into (5.6) to get that the derivative of the total profit effect is negative at $\eta = 1$. Thus by continuity there exists η^0 such that for any $\eta > \eta^0$ the total effect is decreasing function of η.

The second case, $K_w \neq K^*(1)$. Let η^0 be such that $(1-\eta^0)E_s\pi'_K(s, K_w) + \eta^0\pi'_K(1, K_w) = 0$, and η^1 such that $(1-\eta^1)E_s\lambda_s\pi'_K(s, K_w^{new}) + \eta^1\lambda_1\pi'_K(1, K_w^{new}) = 0$ Then $K_\eta^* = K_w$ for any $\eta \geq \eta^0$, and $E_s\pi(s, K_\eta^*)$ is constant function of η for any $\eta > \eta^0$. Similarly $K_\eta^{*new} = K_w^{new}$ for any $\eta \geq \eta^1$, and $E_s\pi(s, K_\eta^*)$ is constant function of η for any $\eta > \eta^1$.

Notice that $\eta^0 < 1$ as otherwise $K_w = K^*(1)$. One can use it to show that $\eta^0 < \eta^1$. Indeed,

$$(1-\eta^0)E_s\lambda_s\pi'_K(s, K_w) + \eta^0\lambda_1\pi'_K(s, K_w) = (1-\eta^0)E_s(\lambda_s - \lambda_1)\pi'_K(s, K_w) > 0.$$

That implies that $K_{\eta^0}^{*new} > K_w > K_w^{new}$ and, therefore, $\eta^1 > \eta^0$. When $\eta \in [\eta^0, \eta^1]$ then the total effect is a strictly decreasing function of η, and it is constant when $\eta > \eta^1$. Thus the total effect is weakly decreasing function for any $\eta > \eta^0$. This proves ii).

Step 4: Now we prove the last part. Let $\lambda_1 = \cdots = \lambda_n$. Then $K_\eta^* = K_\eta^{*new}$ for every η and one can use (5.6) to immediately verify that $(E_sU^{new}(s, K_\eta^{*new}) - E_sU(s, K_\eta^*))'_\eta$ is non-positive for any η. By continuity then there exists $\Lambda^0 > 1$ such that as long as $\lambda_n/\lambda_1 < \Lambda^0$ then $(E_sU^{new}(s, K_\eta^{*new}) - E_sU(s, K_\eta^*))'_\eta$ is also non-positive. That completes the proof. ■

Proof of Proposition 9: The proof follows the same steps as the proof of Proposition 5. It is, however, more complicated as $K^{*new}(s)$ is not necessarily equal to $K^*(s)$.

When the microentrepreneur has a single income source, her income is equal to her profit. Since $\pi(1,0) = \cdots = \pi(n,0) = 0$ the complementarity implies that before the training state 1 is the worst-case state for every K. It follows from Lemma 5.1 below that then state 1 remains to be the worst-case after training. The microentrepreneur's utility, both pre- and post-training, therefore is weighted average of expected profit and $\pi(1, K)$.

Lemma 5.1 *Let s_w and s_w^{new} be the lowest worst states given K under the pre- and post-training state-profit functions respectively. Then $s_w^{new} \leq s_w$.*

Proof. First, consider the multiplicative improvement. Take any $t > s_w$. By definition, for any state t

$$\pi^{new}(t, K) = \frac{g(t)}{g(s_w)} \frac{\pi(t, K)}{\pi(s_w, K)} \pi^{new}(s_w, K).$$

Since s_w is the worst-case state given K, we have that $\pi(t, K) \geq \pi(s_w, K)$. By assumption, $g(s)$ is a weakly increasing function, and so $g(t)/g(s_w) \geq 1$ for every $t > s_w$. Therefore, $\pi^{new}(t, K) \geq \pi^{new}(s_w, K)$ when $t > s_w$. It means that the new lowest worst-case state is s_w or lower.

Now consider the additive improvement. Take any $t > s_w$. Then $\pi(t, K) \geq \pi(s_w, K)$. Then

$$\begin{aligned} \pi^{new}(t, K) - \pi^{new}(s_w, K) &= [\pi^{new}(t, 0) - \pi^{new}(s_w, 0)] + \int_0^K [\pi^{new}(t, k) - \pi^{new}(s_w, k)]'_K \, dk \geq \\ &\geq [\pi(t, 0) - \pi(s_w, 0)] + \int_0^K [\pi(t, k) - \pi(s_w, k)]'_K \, dk = \\ &= \pi(t, K) - \pi(s_w, K) \geq 0. \end{aligned}$$

It means that for any $t > s_w$, the post-training profit at state t is weakly higher, which then implies that the new lowest worst-case state is s_w or lower. ■

Next, we prove that $K_\eta^* < K_\eta^{*new} \leq K^{*new}$ which then would guarantee that the capital adjustment effect is positive. First, we prove that $K_\eta^{*new} \leq K^{*new}$. K_η^{*new} satisfies the FOC $(1-\eta)E_s \pi'^{new}(s, K_\eta^*) + \eta \pi'^{new}(1, K_\eta^*) = 0$. If $K_\eta^{*new} > K^{*new}(1)$ then the second term in the FOC is negative and, therefore, the first term must be positive. That means that $K_\eta^{*new} < K^{*new}$. Thus to prove that $K_\eta^{*new} \leq K^{*new}$ the only thing left is to show that $K_\eta^{*new} > K^{*new}(1)$. Notice that for the post-training case it needs to be proved since post-training state-profit functions do not necessarily satisfy the complementarity condtion.

Lemma 5.2 *$K_\eta^{*new} > K^{*new}(1)$ for every $\eta < 1$.*

Proof. The proof is based on the following fact. Let s_w be the pre-training worst-case state given K. If $\pi'^{new}(s_w, K) \geq 0$ then $\pi'^{new}(s, K) > 0$ for any $s > s_w$. To establish it consider, first, an additive improvement. From (3.3) follows

$$(\pi^{new}(s, K) - \pi^{new}(s_w, K))' \geq (\pi(s, K) - \pi(s_w, K))'.$$

By complementarity of the pre-training state-profit state functions $(\pi(s, K) - \pi(s_w, K))' > 0$, and $\pi'^{new}(s_w, K) \geq 0$ by the assumption. Thus $\pi'^{new}(s, K) > 0$.

For multiplicative improvement

$$\begin{aligned}
\pi'^{new}(s,K) &= \frac{g(s)}{g(s_w)}\left(\pi(s,K)\left(\frac{\pi^{new}(s_w,K)}{\pi(s_w,K)}\right)' + \pi'(s,K)\frac{\pi^{new}(s_w,K)}{\pi(s_w,K)}\right) \\
&> \frac{g(s)}{g(s_w)}\left(\pi(s,K)\left(\frac{\pi^{new}(s_w,K)}{\pi(s_w,K)}\right)' + \pi'(s_w,K)\frac{\pi^{new}(s_w,K)}{\pi(s_w,K)}\right) \\
&= \frac{g(s)}{g(s_w)}\left(\pi(s_w,K)\left(\frac{\pi^{new}(s_w,K)}{\pi(s_w,K)}\right)' + \pi'(s_w,K)\frac{\pi^{new}(s_w,K)}{\pi(s_w,K)}\right) + \\
&+ \frac{g(s)}{g(s_w)}(\pi(s,K)-\pi(s_w,K))\left(\frac{\pi^{new}(s_w,K)}{\pi(s_w,K)}\right)' \geq 0.
\end{aligned}$$

The first inequality holds because $\pi'(s,K) > \pi'(s_w,K)$, which is the complementarity assumption, and that profit functions are positive. The last inequality follows from two facts: the third line is non-negative because it is equal to $\pi'^{new}(s_w,K)$ which, by assumption, is greater or equal than zero. The last line is non-negative because state s_w is the pre-training worst-case state given K, and the derivative of $\pi^{new}(s_w,K)/\pi(s_w,K)$ is non-negative.

Since state 1 is the worst post-training case, from $\pi'^{new}(1,K_w^{new}) = 0$ follows that $\pi'^{new}(s,K_w^{new}) > 0$ for ever s, and, therefore, $K^{*new}(s) > K_w^{new} = K^{*new}(1)$. But then K_η^{*new} must be strictly greater than $K^{*new}(1)$ for every $\eta < 1$. ■

Now we establish that $K_\eta^* < K_\eta^{*new}$. To do that we take derivative of the post-training utility function at K_η^* and show that it is positive. Concavity then would imply $K_\eta^* < K_\eta^{*new}$. First consider the additive improvement. The derivative of the post-training utility function is

$$\begin{aligned}
(U^{new}(K_\eta^*))' &= (U^{new}(K_\eta^*) - U(K_\eta^*))' + U(K_\eta^*)' = (U^{new}(K_\eta^*))' = \\
&= \big((1-\eta)E_s(\pi^{new}(s,K_\eta^*) - \pi(s,K_\eta^*)) + \eta(\pi^{new}(1,K_\eta^*) - \pi(1,K_\eta^*))\big)' > 0.
\end{aligned}$$

Here we took into account that K_η^* is optimal for pre-training utility and the last inequality is by the definition of the additive improvement.

In the case of the multiplicative improvement, the post-training utility can be written as

$$(1-\eta)E_s\pi^{new}(s,K_\eta^*) + \eta\pi^{new}(1,K_\eta^*) = \frac{\pi^{new}(1,K_\eta^*)}{\pi(1,K_\eta^*)}\Big((1-\eta)E_s g(s)\pi(s,K_\eta^*) + \eta\pi(1,K_\eta^*)\Big).$$

Its derivative is

$$\left(\frac{\pi^{new}(1,K_\eta^*)}{\pi(1,K_\eta^*)}\right)'\Big((1-\eta)E_s g(s)\pi(s,K_\eta^*) + \eta\pi(1,K_\eta^*)\Big) + \frac{\pi^{new}(1,K_\eta^*)}{\pi(1,K_\eta^*)}\Big((1-\eta)E_s g(s)\pi(s,K_\eta^*) + \eta\pi(1,K_\eta^*)\Big)'.$$

The first term is positive because multiplicative improvement requires that $(\pi^{new}(1,K)/\pi(1,K))$

is an increasing function of capital. Next, we prove that the second term is positive.

$$\Big((1-\eta)E_s g(s)\pi(s,K_\eta^*)+\eta\pi(1,K_\eta^*)\Big)' = \Big((1-\eta)E_s(g(s)-1)\pi(s,K_\eta^*)\Big)' + \\ + \Big((1-\eta)E_s\pi(s,K_\eta^*)+\eta\pi(1,K_\eta^*)\Big)'.$$

The second term above is equal to zero because this is the FOC that determines K_η^*. Finally, we show that the first term in the expression above is positive. Let τ be such that $\pi_K'(\tau-1,K_\eta^*) < 0 \le \pi_K'(\tau,K_\eta^*)$. Such τ exists because pre-training state-profit functions satisfy complementarity, and $K_\eta^* > K^*(1)$. Then

$$\left(\sum_s (g(s)-1)E_s\pi(s,K_\eta^*)\right)' = \sum_{s\le\tau}(g(s)-1)p_s\pi_K'(s,K_\eta^*)+\sum_{s>\tau}(g(s)-1)p_s\pi_K'(s,K_\eta^*) > \\ > (g(\tau)-1)\sum_{s\le\tau}p_s\pi_K'(s,K_\eta^*)+(g(\tau)-1)\sum_{s>\tau}p_s\pi_K'(s,K_\eta^*) = \\ = (g(\tau)-1)E_s\pi_K'(s,K_\eta^*) > 0.$$

The last inequality follows from the earlier established fact that $K_\eta^* < K^*$ and that from the definition of the multiplicative improvement follows that $g(\tau) > 1$.

This completes the proof since $K_\eta^* < K_\eta^{*new} \le K^{*new}$ implies that the capital adjustment effect is positive. ■

Proof of Proposition 10: Because of concavity, to show that the capital adjustment effect is negative it is sufficient to show that $K_w^{new} < K_w$ and $K^* < K^{*new}$.

First, we prove that $K_w^{new} < K_w$. If $K_w \neq K^*(1)$, then by Proposition 1 there exist two states, $s < s'$ such that $I(s,K_w) = I(s',K_w) \le I(t,K_w)$ for all other states t, and $I_K'(s,K_w) < 0 < I_K'(s',K_w)$. Let s_w and s_w^{new} be the lowest worst states given K_w before and after the training respectively. One can show that from $K^{*new}(s_w) < K_w$ follows that $K^{*new}(s_w^{new}) < K_w$. In the case of the additive improvement it directly follows from the definition. Since $s_w \ge s_w^{new}$ we have

$$(I^{new}(s_w,K_w)-I^{new}(s_w^{new},K_w))' \ge (I(s_w,K_w)-I(s_w^{new},K_w))' \ge 0.$$

By assumption $K_w > K^{*new}(s_w)$, which means that $I'^{new}(s_w,K_w) < 0$, which combined with the inequality above implies that $I'^{new}(s_w^{new},K_w) < 0$ and, therefore, $K^{*new}(s_w^{new}) < K_w$. In the case of the multiplicative improvement we will need to use the *better*-assumption. If $s_w = s_w^{new}$ then we are done. The case $s_w > s_w^{new}$ is impossible. Indeed, assume that $s_w > s_w^{new}$. Then $I(s_w,K^*(s_w)) > I(s_w^{new},K^*(s_w^{new})) > I(s_w^{new},K^*(s_w))$. The first inequality is by the *better*-assumption. The second inequality follows from the fact that $K^*(s_w^{new})$ is optimal in s_w^{new}. Since $K_w > K^*(s_w)$ one can use complementarity to conclude that $I(s_w,K_w) > I(s_w^{new},K_w)$. But then s_w cannot be the worst-case state given K_w, which is a contradiction.

From $K^{*new}(s_w^{new}) < K_w$ follows that $K > K_w$ cannot be the new optimal worst-case capital:

$I_w^{new}(K) = \min_t I^{new}(t, K) \leq I^{new}(s_w^{new}, K) < I^{new}(s_w^{new}, K_w) = \min_t I^{new}(t, K_w) = I_w^{new}(K_w)$. The first inequality comes from the fact that the lowest profit given K is less or equal than the profit at state s_w^{new}. The second inequality comes from the fact that $\pi(s_w^{new}, \cdot)$ declines when $K > K_w$. Thus $K_w^{new} \leq K_w$.

One can further show that K_w is no longer the optimal worst-case capital either. Before training state s_w was the worst-case state and, therefore, $I(s_w, K_w) \leq I(t, K_w)$ for every t, including $t > s_w$. By assumption, all inequalities in (3.3) are strict in the case of the additive improvement, and $g(s)$ is strictly increasing in the case of the multiplicative improvement. Therefore, $I^{new}(t, K_w) > I^{new}(s_w, K_w)$ for any $t > s_w$. Then, it must be the case that all worst states for K_w are less than or equal to s_w. Take any worst-case state given K_w, and denote it as $\hat{s}_w^{new}$. We know that $\hat{s}_w^{new} \leq s_w$.[13] Then $K^{*new}(\hat{s}_w^{new}) < K_w$ by the exact same reasoning as in the beginning of the proof of Proposition 10. Thus all state functions that correspond to the worst-case states decline at K_w. But then, K slightly below K_w will give strictly higher worst-case income, and K_w is no longer optimal. That proves that $K_w^{new} < K_w$.

Now we need to show that $K^* < K^{*new}$. In the proof Proposition 9 it was established that $K_\eta^* < K_\eta^{*new}$ for any $\eta < 1$. In particular, the inequality holds for $\eta = 0$ which means that $K^* < K^{*new}$. This completes the proof.■

References

[1] Ahlin, C., & Waters, B. (2014). Dynamic microlending under adverse selection: Can it rival group lending?. *Journal of Development Economics*.

[2] Ambrus, A., Mobius, M., & Szeidl, A. (2014). Consumption risk-sharing in social networks. *American Economic Review*, 104(1), 149-82.

[3] Baillon, A., Bleichrodt, H., Keskin, U., l'Haridon, O., & Li, C. (2017). The effect of learning on ambiguity attitudes. *Management Science*, 64(5), 2181-2198.

[4] Banerjee, A. V. (2013). Microcredit Under the Microscope: What Have We Learned in the Past Two Decades, and What Do We Need to Know? *Annu. Rev. Econ.*, 5(1), 487-519.

[5] Banerjee, A., & Duflo, E. (2011). Poor economics: A radical rethinking of the way to fight global poverty. *Public Affairs*.

[6] Barham, B. L., Chavas, J. P., Fitz, D., Salas, V. R., & Schechter, L. (2014). The roles of risk and ambiguity in technology adoption. *Journal of Economic Behavior & Organization*, 97, 204-218.

[7] Bateman, M., & Chang, H. J. (2009). The microfinance illusion. Available at SSRN 2385174.

[13] To avoid confusion, note that $\hat{s}_w^{new} \leq s_w$ is not by Lemma 5.1, as the Lemma was applicable to the lowest worst-case state only, which is s_w^{new}. That $\hat{s}_w^{new} \leq s_w$ follows from the established earlier fact that $I^{new}(t, K_w) > I(s_w, K_w)$ for any $t > s_w$.

[8] Berry, S. (1993). *No condition is permanent: The social dynamics of agrarian change in sub-Saharan Africa.* University of Wisconsin Pres.

[9] Betcherman, G., Dar, A., & Olivas, K. (2004). *Impacts of active labor market programs: New evidence from evaluations with particular attention to developing and transition countries.* Social Protection, World Bank.

[10] Brooks, W., Donovan, K., & Johnson, T. R. (2018). Mentors or Teachers? Microenterprise Training in Kenya *AEJ: Applied Economics*, 10 (4).

[11] Bruhn, M., & Zia, B. (2011). Stimulating managerial capital in emerging markets: the impact of business and financial literacy for young entrepreneurs. *World Bank Policy Research Working Paper Series*, Vol. 2011

[12] Cho, Y., & Honorati, M. (2014). Entrepreneurship programs in developing countries: A meta regression analysis. *Labour Economics*, 28, 110-130.

[13] Chowdhury, P. R. (2005). Group-lending: Sequential financing, lender monitoring and joint liability. *Journal of development Economics*, 77(2), 415-439.

[14] Collins, D., Morduch, J., Rutherford, S., & Ruthven, O. (2009). *Portfolios of the poor: How the world's poor live on $2 a day.* Princeton: Princeton University Press.

[15] Dar, A., & Tzannatos, Z. (1999). *Active labor market programs: A review of the evidence from evaluations.* Social Protection, World Bank.

[16] De Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in microenterprises: evidence from a field experiment. *The Quarterly Journal of Economics*, 123(4), 1329-1372.

[17] 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.

[18] De Quidt, J., Fetzer, T., & Ghatak, M. (2016). Group lending without joint liability. *Journal of Development Economics*, 121, 217-236.

[19] De Weerdt, J., & Dercon, S. (2006). Risk-sharing networks and insurance against illness. *Journal of Development Economics*, 81(2), 337-356.

[20] Di Mauro, C., (2008) Uncertainty aversion vs. competence: An experimental market study. *Theory Decision* 64:301-331.

[21] 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.

[22] Ellis, F. (2000). *Rural livelihoods and diversity in developing countries.* Oxford university press.

[23] Engle-Warnick, J., Escobal, J., & Laszlo, S. (2007). *Ambiguity aversion as a predictor of technology choice: Experimental evidence from Peru* (No. 2007s-01). CIRANO.

[24] Fafchamps, M., & Lund, S. (2003). Risk-sharing networks in rural Philippines. *Journal of Development Economics*, 71(2), 261-287.

[25] Fox, C. R., & Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *The Quarterly Journal of Economics*, 585-603.

[26] Fox, C. R., & Weber, M. (2002). Ambiguity aversion, comparative ignorance, and decision context. *Organizational Behavior and Human Decision Processes*, 88(1), 476-498.

[27] Ghosh, P., & Ray, D. (2001) "Information and enforcement in informal credit market." Unpublished Manuscript, Department of Economics, New York University

[28] Giné, X., & Mansuri, G. (2014). Money or Ideas? A Field Experiment on Constraintsto Entrepreneurship in Rural Pakistan. *A Field Experiment on Constraintsto Entrepreneurship in Rural Pakistan (June 1, 2014). World Bank Policy Research Working Paper*, (6959).

[29] Heath, C., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of risk and uncertainty*, 4(1), 5-28.

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

[31] Karlan, D., Knight, R., & Udry, C. (2012). *Hoping to win, expected to lose: Theory and lessons on micro enterprise development* (No. w18325). National Bureau of Economic Research.

[32] Kellett, P., & Tipple, A. G. (2000). The home as workplace: a study of income-generating activities within the domestic setting. *Environment and Urbanization*, 12(1), 203-214.

[33] kweb Kilka, M. & Weber, M. (2001) What determines the shape of the probability weighting function under uncertainty? *Management Science*, 47(12):1712-1726.

[34] Krishna, A. (2004). Escaping poverty and becoming poor: who gains, who loses, and why?. *World development*, 32(1), 121-136.

[35] LeRoy, S. F., & Werner, J. (2001). *Principles of financial economics.* Cambridge University Press, first edition.

[36] Lloyd-Jones, T., & Rakodi, C. (2014). *Urban Livelihoods: A People-centred Approach to Reducing Poverty.* Routledge.

[37] Mazzucato, V. (2009). Informal insurance arrangements in Ghanaian migrants' transnational networks: The role of reverse remittances and geographic proximity. *World Development*, 37(6), 1105-1115.

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

[39] Morduch, J. (1999). Between the state and the market: Can informal insurance patch the safety net?. *The World Bank Research Observer*, 14(2), 187-207.

[40] Prediger, S., & Gut, G. (2014) "Microcredit and Business-Training Programs: Effective Strategies for Micro-and Small Enterprise Growth?", GIGA Focus, Number 3, 2014.

[41] Shapiro, D. A. (2015). Microfinance and dynamic incentives. *Journal of Development Economics*, 115, 73-84.

[42] Valdivia, M. (2012). *Training or technical assistance for female entrepreneurship? Evidence from a field experiment in Peru.* GRADE working paper.

[43] Valentine, T. R. (1993). Drought, transfer entitlements, and income distribution: The Botswana experience. *World Development*, 21(1), 109-126.

[44] Verrest, H. (2013). Rethinking microentrepreneurship and business development programs: vulnerability and ambition in low-income urban Caribbean households. *World Development*, 47, 58-70.

[45] World Bank (2013), *Global Financial Development Report 2014. Financial Inclusion*, Washington D.C.: The World Bank

[46] Yunus, M., (1999) *Banker to the Poor*, New York: Public Affairs, 1999.