

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

Dmitry Shapiro*

This paper develops a theoretical framework to explain the limited effect of business development programs (BDPs) on entrepreneurs' profits. We argue that a mismatch between a BDP's narrow focus on business-promoting strategies and the wider context in which microentrepreneurs operate can limit the impact of business training. In our framework, entrepreneurs are ambiguity-averse and have multiple sources of income (e.g., business and wage incomes.) We show that for a sufficiently ambiguity-averse entrepreneur with multiple income sources, efficient training can result in a decline in expected profit. Notably, when the wider context (multiple income sources, ambiguity-aversion) is considered, the business-training impact is limited and can result in a post-training expected profit decline. This limited impact is caused by the diversifying role that the business income plays in household finances.

JEL Classification: O12, O16, D1

Keywords: Business Development Programs, Ambiguity Aversion, Microentrepreneurship

I. Introduction

Muhammad Yunus, in his *“Banker to the Poor”*, argued that teaching microentrepreneurs is a waste (Yunus, 1999). One cannot improve loan use since

Received: April 23, 2019. Revised: Aug. 28, 2019. Accepted: Oct. 11, 2019.

* Associate Professor, Department of Economics and SNU Institute of Economic Research, 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, two anonymous referees and attendants of various economic seminars 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. This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A5A8027545).

borrowers already use loans efficiently. Indeed, 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) tested whether microentrepreneurs maximize their profit given constraints and found that “... [*many microentrepreneurs*’] *activities prove to be generating an economic loss*” (p. 510). De Mel, McKenzie and Woodruff (2008) found that real returns on capital vary with borrowers’ entrepreneurial ability, indicating that not everyone has the innate ability to do the best with what he or she 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 connections that help to build successful businesses (Banerjee, 2013).

One well-recognized way to make loan use more efficient is the use of business-training programs to improve microentrepreneurs’ business knowledge (Prediger and Gut, 2014). However, the effect of business-training programs is mixed. Meta studies have shown that, while entrepreneurship programs do have a 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 effects of business training on profits. Karlan and Valdivia (2011) reported that the training of female entrepreneurs in Peru led to a noticeable improvement in “bad months” and less noticeable improvement or even a decline in good months. Karlan, Knight and Udry (2012) studied the effect of training on a group of tailors in Ghana. While the business literacy of the tailors in their sample increased, their profits declined. Bruhn and Zia (2011) trained 445 clients in Bosnia and Herzegovina. They found 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. Finally, Drexler et al. (2014) reported that only simplistic training — which consists mostly of basic rules of thumb — improves profits while complex training 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 increases in business literacy after training. Giné and Mansuri (2014) specifically noted that “*business training did lead to an increase in business knowledge, so lack of understanding is not the issue*” (p. 19). This limited impact of training does not appear to be due to improved accounting. For example, Drexler et al. (2014) found that although there was a reduction in mistakes and more consistency across measures of how people calculate profits or sales, it did not affect the main results. Further, de Mel et al. (2014) compared self-reported profits to revenue and cost figures and controlled for detailed measures of accounting practices as a further robustness check. They found no significant evidence that training changes reporting.

McKenzie and Woodruff (2014) argued that issues such as sample size and

sample heterogeneity make it harder to detect the effect of training on Profitability. In a follow-up paper, de Mel, McKenzie and Woodruff (2014) addressed those issues using a large and homogeneous sample of 1252 female entrepreneurs in Sri Lanka. The authors found no training's impact on the Profitability of existing businesses but a positive impact on the Profitability of new businesses and concluded that *"the lack of impacts in most of the existing literature ... may not be just due to power issues"* (p. 200). They further conjectured that business training programs might be less effective than previously thought.

Finally, another explanation suggested in the empirical literature is that one reason for the weakness of BDPs and the business-training they provide is related to their narrow focus on business-promoting strategies, which ignores the wider 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 businesses (Banerjee and Duflo, 2011). Verrest (2013) argued that BDPs *"are relevant to only a minority of entrepreneurs"* due to variations in household vulnerability or a lack of business ambition (p. 58). Second, microentrepreneurs do not view their business activities solely as a way to bring in more money. Instead, they consider them as a valuable diversification tool for dealing with irregularities in income sources (Krishna, 2004); as a way to reduce the household's vulnerability to negative shocks, such as job-loss or illness (Ellis, 2000); or as a strategy 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 the complex reality in which microentrepreneurs run their businesses can be responsible for a limited, or even negative, impact of business training on microentrepreneurs' profits.

To capture the wider context in which microentrepreneurs operate, we introduce two assumptions. Our first assumption is that the microentrepreneur has multiple sources of income. One source of income is profit from the business activity. This depends on the amount of capital invested and the state of nature. Other income sources can include farming, wage employment, temporary migration or income from informal risk-sharing arrangements. Non-business income does not depend on the capital investment but can depend on the state of nature. It is well documented in the literature that households commonly rely on multiple income sources,¹ yet this assumption is rarely used in the theoretical microfinance literature,

¹ For example, 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

where non-business incomes are typically normalized to zero.² As we show in this paper, disregarding multiple sources of in-come results in a loss of generality. The effectiveness of training differs depending on whether the microentrepreneur has one or multiple sources of income available. In particular, only in a setting with multiple income sources can the post-training expected profit decline.

Our second assumption is that the microentrepreneur has two objectives. The first is to maximize expected income. We will refer to this as the business-oriented ambition. The second objective is to maximize the “rainy day” income or, more formally, the worst-case income. We will refer to this as the livelihoods-oriented ambition. We model the microentrepreneur’s utility as a weighted average of the two objectives: business-oriented ambition (maximizing expected income) and livelihoods-oriented ambition (maximizing worst-case income). Mathematically, our setup follows the framework of Gilboa and Schmeidler (1989) in that the microentrepreneur is modeled as an ambiguity-averse decision-maker who maximizes her minimum expected utility.³

The microentrepreneur takes a business training course, which introduces a new business practice. The new business practice affects the Profitability of the business activity only and does not affect non-business income sources. The new practice is superior to the old one, and its impact is stronger in states where capital is more productive. The training’s total impact on the expected profit is a sum of two effects. The first effect is the *profit improvement effect*. It is equal to the difference in post-training and pre-training expected profits at the pre-training level of capital in-vestment. It is always positive due to the superiority of the new practice. The second effect is the *capital adjustment effect*. It takes into account that the

wages (Table 3.1, Ellis 2000). A survey of households in Mamone, a poor 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%, and other activities (beer brewing, basket weaving, carpentry) for 18.5% (Valentine, 1993).

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) and Ghosh and Ray (2001) and extending to more recent papers, including Chowdhury (2005), Ahlin and Waters (2014), de Quidt et al. (2016) and Shapiro (2015).

³ The literature has documented the role of ambiguity-aversion on the willingness to adopt new technologies and practices, and the effect is distinct from risk-aversion. Engle-Warnick et al. (2007) documented that farmers in Peru use a traditional variety of potato with low expected yield, which, nonetheless, generates enough potatoes to feed the farmer’s family. This is despite the availability of new varieties of potatoes, such as the *Papa Caprio*, which provide substantial yield improvement. They show that it was ambiguity-aversion and not risk-aversion that was responsible for the crop adoption decision. Similarly, Barham et al. (2014) examined the 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 the adoption of the less ambiguous genetically-modified corn seeds.

microentrepreneur will adjust her capital investment after the adoption of the new practice. It is equal to the difference in post-training expected profit at post-training and pre-training levels of capital investment. Unlike the profit improvement effect, it can be either positive or negative. If, after the training the microentrepreneur invests more than before the training, it is positive. Otherwise, it is negative.

Having a negative capital adjustment effect undermines the efficiency of the business training and limits its impact. In our model, the business training has a stronger effect in higher states, so in order to fully reap the benefits of the new business practice one should invest more than before the training. When the capital adjustment effect is negative, however, the microentrepreneur does the exact opposite and invests less, thereby limiting the training's impact.

Our main results are as follows. First, we show that for the capital adjustment effect to be negative it is necessary to consider the full context — multiple ambitions and multiple income sources — in which the microentrepreneur operates. If the microentrepreneur has only business-oriented ambition (i.e., she is an expected-income maximizer) or her only income source is business profit, then the capital adjustment effect is positive. Second, we derive sufficient conditions for the capital adjustment effect to be negative. The livelihoods-oriented ambition (i.e. ambiguity-aversion) should be sufficiently strong and the non-business income sources should be sufficiently diversifying. Given the ambiguity-averse focus on the worst-case outcome, sufficiently diversifying means that non-business income should be high enough in the worst state for the business activity. Finally, we show that not only can the capital adjustment effect be negative but it can also outweigh the always-positive profit improvement effect resulting in a post-training decline in expected profit.

To see the intuition, consider a livelihoods-oriented entrepreneur whose objective is to maximize her worst-case income. The business training has stronger effect in states where the capital is more profitable and does not impact non-business income sources. Thus, for any given capital investment, the post-training worst-case state(s) is lower than the pre-training worst-case state(s). Since optimal capital investment is lower in states with lower capital Profitability, a livelihoods-oriented microentrepreneur — instead of taking advantage of the improved Profitability by investing more — invests less, which results in the negative capital adjustment effect. The role of the non-business income sources is that the state with the lowest capital Profitability is not necessarily the state with the lowest income. In particular, this means that a post-training worst-case state can have strictly lower capital Profitability than a pre-training worst-case state such that the capital adjustment effect is strictly negative.

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. 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 the strength, or lack thereof, of the business-oriented ambition. Third, we show that the holistic modeling of the microentrepreneurial decision is crucial to understanding how efficient training can have a mixed-to-negative impact. Only when multiple income sources *and* multiple objectives are introduced can the training have negative impact on expected profit.

The rest of the paper is organized as follows. Section 2 develops a basic setup and presents model's assumptions regarding microentrepreneurs' preferences and technology. Section 3 introduces training and studies its effect on the microentrepreneurs' profit. In subsection 3.1 we consider a simple functional form to model the training's effect. General results are presented in subsection 3.2. The explicit example of post-training profit decline is given in subsection 3.3. All proofs are in the Appendix.

II. Basic Setup

Consider a microentrepreneur who has access to multiple sources of income that include profit from business activities as well as non-business income sources. We assume that profit from business activities, $\pi(s, K)$, depends on the amount of capital invested, K , and the state of nature, s . Funds from non-business income sources, $h(s)$, do not depend on K but can depend on the state of nature. Given $\pi(s, K)$ and $h(s)$ the microentrepreneur's income is $I(s, K) = \pi(s, K) + h(s)$. States are the only source of uncertainty and are labeled by integer numbers from 1 to n , $s \in \{1, \dots, n\}$. The probability of state s is p_s . We will use $K^*(s)$ to denote the capital level that maximizes income and profit in a given state s : $K^*(s) = \arg \max_K I(s, K) = \arg \max_K \pi(s, K)$.

We assume that profit and income functions have the following properties:

- (A1) (*Regularity*) For any s and $K \geq 0$, the state-profit function, $I(s, K)$, is a concave, single-peaked, differentiable function of K .
- (A2) (*Complementarity*) $I_K(s', K) > I_K(s, K)$ when $s' > s$.
- (A3) (*States-ranking*) $I(1, K^*(1)) < \dots < I(n, K^*(n))$.
- (A4) (*Initial Conditions*) $0 \leq \pi(1, 0) \leq \dots \leq \pi(n, 0)$.

Assumption (A1) imposes common technical conditions, such as concavity, differentiability with respect to K and a uniqueness of maximum. Since $h(s)$ does not depend on K , $\pi(s, K)$ also satisfies (A1). Assumption (A2) requires that the capital has a higher marginal Profitability at higher states. Since $h(s)$ does

not depend on K , $\pi(s, K)$ also satisfies (A2). An immediate corollary from (A2) is that $K^*(1) < \dots < K^*(n)$. (A3) ensures that not only capital has higher marginal Profitability in higher states but also that higher states have higher upside potential. Finally, (A4) requires that higher states have higher profit when $K = 0$. Note that (A4) is the only assumption that $\pi(s, K)$ has to satisfy but $I(s, K)$ does not have to. (A1)-(A3) are satisfied by both $\pi(s, K)$ and $I(s, K)$. In particular, from (A2) and (A4) it follows that the profit function satisfies (A3): $\pi(1, K^*(1)) < \dots < \pi(n, K^*(n))$.

The microentrepreneur's utility is the weighted average of the expected income and the worst-case income,

$$U(K) = (1 - \eta) \sum_s p_s I(s, K) + \eta \min_s I(s, K), \quad (1)$$

which she maximizes with respect to K . Parameter η is exogenously given, and we discuss its interpretation later in this section.

As follows from (1), the microentrepreneur is not an expected-profit maximizer. Instead, her utility is a combination of two objectives that, following Verrest (2013), we will refer to as a *business-oriented* ambition and a *livelihoods-oriented* ambition. We define a business-oriented ambition as an expected income maximization. Since $h(s)$ does not depend on K , maximizing expected income is equivalent to maximizing expected profit. We define a livelihoods-oriented ambition as maximizing the “rainy day” income, which is the income in the worst-case state, $\min_s I(s, K)$.⁴ We will use the terms “*rainy day income*” and “*worst-case income*” interchangeably throughout this paper.

Objective function (1) is a special case of the ambiguity-aversion preferences that were developed by Gilboa and Schmeidler (1989).⁵ Consider an ambiguity-averse microentrepreneur who does not know the objective distribution of states, $\{p_s\}$, and instead assumes that it belongs to a set of priors $\mathcal{Q} = \{q : q_s \geq (1 - \eta)p_s, q_s \geq 0, \sum_s q_s = 1\}$. As in Gilboa and Schmeidler (1989), the microentrepreneur maximizes

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 theoretical literature that applies the ambiguity-aversion framework to problems in development economics is sparse. The two most commonly used ambiguity-aversion frameworks are those of Gilboa and Schmeidler (1989) and Klibanoff et al. (2005), which is a generalization of the Gilboa and Schmeidler (1989). For example, the Gilboa and Schmeidler (1989) framework is used by Bryan (2019) to model farmers' reaction to partial insurance products. The Klibanoff et al. (2015) framework is used by Elabed and Carter (2015) to model farmers' willingness to pay for microinsurance when they are ambiguity-averse.

her minimum expected utility, that is, she chooses K to maximize the expected income under the *worst* prior in \mathcal{Q} :

$$\max_K \min_{q \in \mathcal{Q}} \sum_s q_s I(s, K). \quad (2)$$

It is straightforward to verify that (2) is equivalent to maximizing (1):

$$\begin{aligned} \max_K \min_{q \in \mathcal{Q}} \sum_s q_s I(s, K) &= \max_K \left\{ (1-\eta) \sum_s p_s I(s, K) + \eta \min_s I(s, K) \right\} \\ &= \max_K U(K),^6 \end{aligned} \quad (3)$$

where $\{q_s\}$ is a prior from \mathcal{Q} and $\{p_s\}$ is the objective distribution.

Parameter η has two mathematically equivalent interpretations. The first one, used to derive (3), is the multiple prior interpretation. It captures the ambiguity via the decision-makers' beliefs about a set of priors, \mathcal{Q} , with a higher η corresponding to a larger set. 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 the decision weight interpretation. It captures the ambiguity via decision weight with a higher η corresponding to a higher degree of ambiguity aversion. When $\eta=0$, the microentrepreneur is an expected-income maximizer; when $\eta=1$, the microentrepreneur is the worst-case income maximizer. 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

Let K_η^* be the capital level that maximizes $U(K)$, K_w be the capital level that maximizes the worst-case income, $K_w = \arg \max_K \min_s I(s, K)$, and K^* be the capital level that maximizes expected income. Since $\{h(s)\}$ does not depend on K , K^* maximizes both expected income and expected profit. By definition, $K^* = K_0^*$ and $K_w = K_1^*$. Let $s_w(K)$ denote a worst-case state (any if there is more than one), i.e., a state with the lowest $I(s, K)$. Finally, let $I_w(K) := I(s_w(K), K)$ denote the worst-case income for a given K .

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

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 This discussion is based on Baillon et al. (2018) who develop a method to experimentally separate the two interpretations. Their analysis is based on a more general class of preferences axiomatized in Chateauneuf et al. (2007). Equation (3) in our model is a special case of equation (3) in Baillon et al. (2018) with $a_i = \eta$ and $b_i = -\eta$.

- i) $K_w = K^*(1)$, or
- ii) there exist two states, $s < s'$, such that $I(s, K_w) = I(s', K_w) \leq I(t, K_w)$ for any t , and $I_K(s, K_w) < 0 < I_K(s', K_w)$.

Proposition 2 If $K_w < K^*$ then K_η^* is a decreasing function of η and $K_\eta^* \in [K_w, K^*]$.

Propositions 1 and 2 characterize properties of the optimal capital choice. Proposition 1 looks at case of $\eta = 1$ and shows that unless $K_w = K^*(1)$ there exist at least two worst-case states given K_w . Proposition 2 shows K_η^* is a decreasing function of η when $K_w < K^*$. One can immediately adjust its proof to the case of $K_w > K^*$ when K_η^* is an increasing function of η . Both Propositions are proved in the Appendix and their proofs are straightforward applications of (A1)-(A3).

Whether $K_w = K^*(1)$ or not will prove to be important for our analysis, and it is determined by properties of non-business income sources, $\{h(s)\}$. From (A2) and (A4), it follows that $\pi(1, K) = \min_s \pi(s, K)$ for every K , so that state 1 is uniformly the worst state for business profits. Thus, without non-business income sources, i.e. when $I(s, K) = \pi(s, K)$, it is always the case that $K_w = K^*(1)$. In order for $K_w \neq K^*(1)$ there must exist state s such that $\pi(1, K^*(1)) + h(1) \geq \pi(s, K^*(1)) + h(s)$. Non-business income in state 1, $h(1)$, should be large enough to substitute for low income from business activities. A necessary condition for $K_w \neq K^*(1)$ is $h(1) > \min_s h(s)$. In particular, $\{h(s)\}$ cannot be zero, which is a common assumption in the literature; nor can it be a constant or an increasing function of s .

In terms of empirical evidence, on the one hand, it is well documented that many risks affecting the 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 poor households coping with income fluctuations, are widespread. (Ambrus et al., 2014). Risk-sharing is routinely mentioned as the most common way for households to deal with negative shocks, including death, sickness, crime and court cases and shocks in income generating activities (De Weerd and Dercon, 2006; Mazzucato, 2009). Moreover, they are substantial enough to be successfully applied for smoothing household consumption. While household income in developing countries varies greatly, consumption is remarkably smooth (Fafchamps and Lund, 2003). Also, in addition to informal risk-sharing arrangements, governmental programs can provide an income boost in a disastrous state. For example, in Botswana, the government drought relief program during 1985-86 created wage employment opportunities, substituting for the decreased share of livestock in crops in income

portfolios (Valentine, 1993).

As an example of $\{h(s)\}$ where $h(1) > \min_s h(s)$, consider a microentrepreneur whose only non-business income sources are income from an informal risk-sharing arrangement with distant family members or other villagers and employment income. State 1 is a bad state for the microentrepreneur, so if state 1 is realized, she receives help through her risk-sharing arrangement, $h(1) > 0$. State 2 is an intermediate state where the business is profitable enough that no help from risk-sharing is needed, but there are no employment opportunities, $h(2) = 0$. State 3 is a good state for the business, and the microentrepreneur also receives positive labor income, $h(3) > 0$. Another example is when the only non-business income source is a governmental subsidy via disaster relief programs, where state 1 is the disaster state. Then, one would have $h(1) > 0$ and $h(2) = \dots = h(n) = 0$. Finally, it could 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.

III. Business Training

Assume that the microentrepreneur can take a training course that introduces her to a new business practice or a new technology.⁸ We will use the superscript *new* to refer to variables and functions related to the new practice. For example, $\pi^{new}(s, K)$ is the profit function under the new practice; K_η^{*new} is the capital level that maximizes the microentrepreneur's post-training objective function, and so on. We assume that the business-training and the post-training profit function satisfy the following assumptions (BT1)-(BT4).

- (BT1) (*Regularity*) $\pi^{new}(s, K)$ is a single-peaked, concave, differentiable function of K .
- (BT2) (*Zero cost*) The cost of the training and implementing the new practice is zero.
- (BT3) (*Business focus*) Training has no effect on non-business activities, $h^{new}(s) = h(s)$ for every s .
- (BT4) (*Profit improvement*) $\pi^{new}(s, K) > \pi(s, K)$ for any s and $0 < K \leq K_\eta^{*new}(n)$.

⁸ The scope and level of training vary between different BDPs. In Karlan, Knight and Udry (2012), for example, the training was on a small scale and involved targeted lessons, such as keeping time and transaction records, separating business and personal money. On the other hand, de Mel, McKenzie and Woodruff (2014) 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 three-to five-day training courses and covers topics such as organization of sta., record keeping and stock control, marketing and financial planning.

Assumption (BT1) imposes the same regularity condition on $\pi^{new}(s, K)$ that assumption (A1) imposed on $\pi(s, K)$. Assumption (BT2) states that there is no cost associated with taking the training and no cost associated with implementing the new practice. In particular, it is (weakly) optimal for the microentrepreneur to undertake the training. Assumption (BT3) states that the business-training affects the income from business activities only and does not affect incomes from other activities, $\{h(s)\}$. Finally, assumption (BT4) assumes that the new business practice is superior to the old one, and it improves profit in every state and for a sufficiently large domain of K . The upper bound $K^*(n)$ is chosen so that whatever happens outside of $[0, K^{*new}(n)]$ does not affect the microentrepreneur's decision-making since $K_\eta^* \in [K_w, K^*] \subseteq [K^*(1), K^*(n)]$. It is imposed so as not to exclude a large group of possible profit improvements, such as the one used in Section 3.1 and multiplicative improvements in Section 3.2. We assume that $K > 0$ rather than $K \geq 0$ to allow for the possibility that $\pi^{new}(s, 0) = \pi(s, 0)$.

Given our assumptions it is straightforward to verify that the training increases the microentrepreneur's utility. Indeed,

$$U(K_\eta^*) < U^{new}(K_\eta^*) \leq U^{new}(K_\eta^{*new}). \quad (4)$$

Here, the first inequality follows from (BT3) and (BT4). Indeed, $h^{new}(s) = h(s)$ by (BT3) and $\pi^{new}(s, K_\eta^*) > \pi(s, K)$ by (BT4) and the fact that $K_\eta^* \leq K^*(n)$. Therefore, $I^{new}(s, K_\eta^*) > I(s, K_\eta^*)$ which implies the first inequality. The second inequality is due to the fact that K_η^{*new} is the optimal capital level for the post-training utility function. We summarize the reasoning above in Proposition 3.

Proposition 3 *If the training satisfies (BT1)-(BT4) then it strictly increases a microentrepreneurial utility.*

From a welfare perspective, this indicates that the business training is valuable, as it has a positive effect on microentrepreneurs' well-being regardless of its effect on expected profit. Furthermore, a microentrepreneur will prefer the new practice and will always adopt it. Microentrepreneurs tend to follow, at least in the short-run, the practices they learn during the training course, and this is well-documented in the literature. Table 8 in McKenzie and Woodruff (2014)'s survey summarizes the effects 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).

Next, we look at the effect of training on the expected profit. Given that income from non-business 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, $E_s \pi^{new}(s, K)$. The difference between the two is a constant, $E_s h(s)$ and, in

particular, if one increases (decreases) then so does another. The total effect of the training on expected profit can be disentangled into two effects:

$$E_s \pi^{new}(s, K_\eta^{*new}) - E_s \pi(s, K_\eta^*) = \underbrace{[E_s \pi^{new}(s, K_\eta^*) - E_s \pi(s, K_\eta^*)]}_{\text{profit improvement}} + \underbrace{[E_s \pi^{new}(s, K_\eta^{*new}) - E_s \pi^{new}(s, K_\eta^*)]}_{\text{capital adjustment}} \quad (5)$$

where K_η^* and K_η^{*new} are utility-maximizing capital levels before and after training, respectively. The first effect is the *profit improvement effect*, and it measures how much the profit will change given the pre-training capital investment. The profit improvement assumption ensures that it is positive. The second effect is the *capital adjustment effect*. After the training, K_η^* is no longer optimal, and the microentrepreneur will change her capital investment to K_η^{*new} . The capital adjustment effect measures how much the profit will change given the microentrepreneur's adjustment to the adoption of new business practices or a new technology. Unlike the profit improvement effect, the capital adjustment effect can be either positive or negative. If $K_\eta^{*new} > K_\eta^*$, then it is positive. Otherwise, it is negative.

Whenever the capital adjustment effect 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. In our model, the business-training will have stronger effect in higher states; so, in order to take full advantage of it, one needs to invest more than before the training. However, because of the microentrepreneur's ambiguity-aversion and availability of income from non-business activities, the microentrepreneur might do the exact opposite and invest less, thereby limiting the training's effect.

3.1. Effect of Business Training on Profit. A Special Case

In this subsection, we consider a special case where the business training improves Profitability by a fixed factor. We will consider a more general specification in the next subsection.

(BT5-I) (Non-negativity) $\pi(s, K) \geq 0$ for every s and every $K \in (0, K^*(n)]$.

(BT5-II) (λ -improvement) $\pi^{new}(s, K) = \lambda_s \pi(s, K)$, where $1 < \lambda_1 < \dots < \lambda_n$.

Assumption (BT5-I) is imposed to ensure that (BT4) is satisfied. Assumption (BT5-II) states that the training increases the profit in state s by a fixed factor, λ_s . (BT5-II) greatly simplifies the proofs due to its property that $K^*(s) = K^{*new}(s)$;

however, as the next section shows, it is not necessary for the main message of the paper. Condition $1 < \lambda_1 < \dots < \lambda_n$, and its generalization in the next subsection, are imposed so that the effect of training is stronger in higher states, which are the states where capital is more profitable. An example of this would be if trainees learn how to find cheaper suppliers or become more efficient at inventory management, which will have a stronger effect during good states when sales are higher.⁹

We begin the analysis of business training on expected profit with two benchmarks. The first is when the microentrepreneur has only business-oriented ambition. The second benchmark is when the microentrepreneur's only income source is the business income. As the next Proposition shows, in both benchmarks the capital adjustment effect is greater than or equal to zero. Given that the profit improvement effect is always positive, this means that the total effect is also positive.

Proposition 4 *Assume that the training satisfies (BT5-I) and (BT5-II). The capital adjustment effect is non-negative if either:*

- i) $\eta = 0$; or
- ii) *the microentrepreneur's only source of income is the business income.*

Corollary 1 *If the conditions of Proposition 4 are satisfied, then the business training has a positive total effect on expected profit.*

Intuitively, BDPs are designed to promote business-oriented strategies, such as business growth or production strengthening (Verrest, 2013). When $\eta = 0$, the microentrepreneur's only objective is to maximize her expected income, which is equivalent to maximizing her expected profit. The training's focus on improving profit from the business activities matches the microentrepreneur's objective, and post-training expected profit goes up. When profit is the only source of income, there is nothing to supplement the microentrepreneur's income in the worst-case state when the profit is low. As Proposition 4 shows, the best available option to maximize the worst-case income is to increase capital investment, which will result in a higher expected profit.

⁹ Brooks et al. (2018) empirically studied the effect of training using two treatments. The first is a standard business-training program used throughout Kenya. The second treatment is the so-called, mentor condition, where the entrepreneur is being mentored by a more experienced entrepreneur from the community. In the study, only the mentor treatment had a positive effect on the entrepreneurs' profit. As anecdotal evidence of why it worked, Brooks et al. (2018) mentioned Prudence, who was a participant of one the mentor treatments, and who used to purchase inventory from suppliers at the entrance of a market area. After training, she started to purchase at stalls deeper into the market and only after comparing prices. Her cost dropped from 250 Ksh to 100 Ksh as a result, while she kept her sales price exactly the same. Relating this to our paper and (BT5-II), if training results in a post-training reduction in marginal cost, it will have a stronger effect in states that are favorable to the business activity.

Consider now the case of a microentrepreneur who has multiple income sources *and* whose objective differs from maximizing her expected income, $\eta \neq 0$. When $\eta \neq 0$, the microentrepreneur puts the positive weight on income in the worst-case state. When business income is the only income source, state 1 is the worst-case state for every K . With multiple income sources, however, state 1 is not necessarily the worst-case state, as non-business income sources (e.g., an informal risk-sharing arrangement) can supplement the low income from business activities in state 1. This max changes how the microentrepreneur responds to the business training. Let $s_w^{\max}(K_\eta^*)$ be the highest worst-case state before the training. If, because of the non-business income sources, $s_w^{\max}(K_\eta^*) > 1$ then $s_w^{\text{new}}(K_\eta^{*\text{new}}) < s_w^{\max}(K_\eta^*)$, that is, new worst-case state(s) are strictly lower. By the complementarity, lower states need lower capital investment, which puts a downward pressure on the optimal post-training capital level and can result in the negative capital adjustment effect.

Proposition 5 formalizes the intuition above for the case of $\eta = 1$. It imposes two conditions. First, $K^*(1) < K_w$. By Proposition 1, $K^*(1) \neq K_w$ means there is state $s_w \neq 1$ at is a worst-case state given K_w . That is, as discussed above, non-business income in state 1, $h(1)$, is high enough so that $\pi(1, K^*(1)) + h(1) \geq \pi(s_w, K^*(1)) + h(s_w)$. The second condition is that $K_w < K^*$ (i.e., the microentrepreneur was under-investing prior to the training). Proposition 5 shows that, on the one hand, $K^* < K^{*\text{new}}$. To maximize expected income one needs to invest more after the training. On the other hand, $K_w^{\text{new}} < K_w$. To maximize the worst-case income, one should invest less after the training. When $\eta = 1$, the microentrepreneur maximizes her worst-case income, and so the capital adjustment effect is negative.

Proposition 5 Assume that the training satisfies (BT5-I) and (BT5-II). Let $\eta = 1$. If $K^*(1) \neq K_w$ and $K_w < K^*$, then the capital adjustment effect is negative.

We conclude this section by looking at the effect of BDPs on expected profit when $0 < \eta < 1$. Proposition 4 has shown that when $\eta = 0$ the capital adjustment effect is always positive. Proposition 5 has shown that if $\eta = 1$ and $K^*(1) < K_w < K^*$, then the capital adjustment effect is negative. A natural conjecture would be that when $K^*(1) < K_w < K^*$ the capital adjustment effect is a decreasing function of η , and there exists $\hat{\eta}$ such that it is positive when $\eta < \hat{\eta}$ and negative when $\eta > \hat{\eta}$. Similarly, one could conjecture that the total profit effect is also a decreasing function of η . It turns out that neither is correct. Proposition 6 shows that the capital adjustment effect is not a decreasing function of η , and the total effect is not necessarily a decreasing function of η . A consequence is that, even for this simple form of profit improvement, not much can be said about the signs of the capital adjustment effect and the total effect for intermediate values of η .

To see why the capital adjustment effect is not a decreasing function of η , consider its derivative with respect to η , $(E_s \pi^{\text{new}}(s, K_\eta^{*\text{new}}) - E_s \pi^{\text{new}}(s, K_\eta^*))'_\eta$, when

$\eta = 0$. Term $(E_s \pi^{new}(s, K_\eta^{*new}))'_\eta = 0$ when $\eta = 0$ because $K_\eta^{*new} = K^{*new}$ and $E_s \pi^{new}(s, K^{*new}) = 0$. Term $(E_s \pi^{new}(s, K_\eta^*))'_\eta$, as we show, is negative when $\eta = 0$ so that the sign of the derivative is positive. That is, when η is sufficiently close to zero, the capital adjustment effect is an increasing function of η . When η is sufficiently close to one, on the other hand, the capital adjustment effect is a decreasing function of η . Therefore, it is a non-monotone function of η .

With the total effect, the situation is slightly different. First, similarly to the capital adjustment effect, it can be an increasing function of η when η is sufficiently close to zero. To see how, consider the limit case when $\lambda_1 = \dots = \lambda_{n-1} = 1$ and $\lambda_n = \infty$. Then, the post-training choice of capital is not sensitive to changes in η , so $K_\eta^* = K^*(n)$ and the post-training expected profit does not change. The pre-training expected 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 . Second, also similarly to the capital adjustment effect, the total effect is a decreasing function of η when η is sufficiently close to one. Third, differently from the capital adjustment effect, the total effect can be a decreasing function of η for every $\eta \in [0, 1]$. For example, as long as λ_n / λ_1 is not too large, so that the extreme example above is not applicable, the total effect is a decreasing function of η .

Proposition 6 *Assume that the training satisfies (BT5-I) and (BT5-II). Assume also that $K_w < K^*$. Then, the capital adjustment effect is*

- i) *an increasing function of η for any η sufficiently close to zero.*
- ii) *a decreasing function of η for any η sufficiently close to one. The total effect of training on expected profit*
- iii) *can be an increasing function of η when η is sufficiently close to zero.*
- iv) *is a decreasing function of η for any η sufficiently close to one.*
- v) *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 a specific functional form of profit improvement, $\pi^{new}(s, K) = \lambda_s \pi(s, K)$. In this subsection, we extend the analysis to more general functional forms of profit improvement. We will refer to them as *additive* and *multiplicative* improvements.

(BTA) (Additive improvement) The post-training state-profit functions satisfy (BT1)-(BT4) and are such that

$$(\pi^{new}(t, K) - \pi(t, K))' \geq (\pi^{new}(s, K) - \pi(s, K))' \geq 0 \quad \text{when } t > s, \quad (6)$$

and $\pi^{new}(t,0) - \pi(t,0) \geq \pi^{new}(s,0) - \pi(s,0) \geq 0$ when $t > s$.

Under an additive improvement, $\pi^{new}(s,K) > \pi(s,K)$ for every s and every K , so it is stronger than what is required by (BT4). Also, notice that by definition the training with an additive improvement has a stronger effect in higher states (the first inequality in (6)) and for higher levels of K (the second inequality in (6)).

One example of the business training that satisfies (BTA) is $\pi^{new}(s,K) = \lambda_s + \pi(s,K)$, where $\lambda_n \geq \dots \geq \lambda_1 \geq 0$. For another example, assume that the pre-training profit function is given by $\pi(s,K) = F(s,K) - RK$, where $F(s,K)$ is a standard production function such that $F_K > 0, F_{KK} < 0$ and $F_K(t,K) > F_K(s,K)$ when $t > s$. The post-training profit function is $\pi^{new}(s,K) = \lambda_s F(s,K) - RK$. In this case, condition (6) becomes

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

It is satisfied if $\lambda_n \geq \dots \geq \lambda_1 \geq 1$. The second requirement of (BTA) is trivially satisfied.

(BTM-I) (Non-negativity) $\pi(s,K) \geq 0$ for every $s \in \{1, \dots, n\}$ and every $K \in [0, K^{*new}(n)]$.

(BTM-II) (Multiplicative improvement) The post-training state-profit functions satisfy (BT1)-(BT4) and are such that

$$\frac{\pi^{new}(s,K)}{\pi(s,K)} = g(s) \frac{\pi^{new}(1,K)}{\pi(1,K)}, \quad (7)$$

where $g(s) \geq 1$ is a weakly increasing function of s . $\frac{\pi^{new}(1,K)}{\pi(1,K)} > 1$ and is a weakly increasing function of K when $K \in [0, K^{*new}(n)]$.

By design the business-training is weakly more efficient in higher states because $g(s)$ is an increasing function of s . From equation (7), it follows that $\pi^{new}(s,K) / \pi(s,K)$ is an increasing function of K when $K \in [0, K^{*new}(n)]$, so in that domain the training has a stronger effect for higher K . We impose (BTM-I) for the same reason as we imposed (BT5-I). The multiplicative improvement improves the profit function by multiplying it by some factor. To ensure that it is an improvement, the pre-training profit function needs to be positive. It is straightforward to verify that under multiplicative improvement $K^{*new}(n) \geq K^*(n)$, which means that that multiplicative training satisfies (BT4).

An example of the multiplicative improvement is $\pi^{new}(s,K) = \lambda_s \pi(s,K)$, where $\lambda_s \geq 1$. The requirement that $g(s)$ is a weakly increasing function of s is

satisfied if $\lambda_n \geq \dots \geq \lambda_1$.

It turns out that when the post-training profit function satisfies either (BTA) or the two (BTM) assumptions then the equivalent of Proposition 4 holds. When the microentrepreneur is risk-neutral or has only one source of income the training will have a positive effect on expected profit. The proof of Proposition 7, while more technical, follows the same steps as that of Proposition 4.

Proposition 7 *Assume that the training satisfies either (BTA) or (BTM-I) and (BTM-II). The capital adjustment effect is non-negative if one of the two conditions hold:*

- i) $\eta = 0$; or
- ii) *the microentrepreneur's only source of income is the business income.*

Corollary 2 *If conditions of Proposition 7 are satisfied, then the business training has a positive total effect on expected profit.*

Finally, we derive sufficient conditions for the capital adjustment effect to be negative, which is a generalization of Proposition 5. Just like in Proposition 5, it is necessary that $K^*(1) < K_w$. Non-business incomes sources must provide sufficient cushion to the business income in state 1. However, in a more general setting of this subsection, an additional condition is needed. Since $K^{*new}(s)$ is not necessarily equal to $K^*(s)$, it is possible to have a training so efficient at improving the marginal Profitability of capital 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 a restriction on how much the training can improve capital's marginal Profitability. Finally, the last condition of Proposition 8 is analogous to the assumption $\lambda_1 < \dots < \lambda_n$ from the previous subsection.

Proposition 8 *Let $\eta = 1$ and $K^*(1) < K_w < K^*$. Let s_w be the lowest pre-training worst-case state given K_w , and let $K^{*new}(s_w) < K_w$. If (BTA) is satisfied and all inequalities in (6) are strict, then the capital adjustment effect is negative. Similarly, if (BTM-I) and (BTM-II) are satisfied and $g(s)$ is a strictly increasing function of s , then the capital adjustment effect is negative.*

3.3. Example of Post-Training Profit Decline

As we discussed earlier, whenever the capital adjustment effect is negative, it undermines the effectiveness of the business training. Instead of taking advantage of improved Profitability and expanding her business by investing more, the microentrepreneur finds it safer to invest less, thereby limiting the training's impact. In fact, the negativity of the capital adjustment effect can be large enough to outweigh the positive profit improvement effect and result in a lower post-training

expected income and expected profit.

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 profit, employment income, and income from informal risk-sharing arrangements. The microentrepreneur has the endowment of labor normalized to 1. Labor can be used for business activities and for employment. The microentrepreneur has no endowment of capital but can borrow it at rate R . Capital can be used for business activities only. Income from risk-sharing arrangements does not require any inputs.

The timing is as follows. First, the microentrepreneur decides how much capital to invest into her business. Second, the state of nature, s , is realized. Given s , the microentrepreneur decides how to divide her labor endowment between business activities and employment. Finally, the microentrepreneur earns employment income and business profit according to her capital choice and labor allocation. She also receives income from her risk-sharing arrangements. Note that the choice of labor allocation is flexible and can be adjusted to the state of nature. The capital investment, on the other hand, cannot, as it is made before the uncertainty is realized.

If the microentrepreneur splits the labor between employment and business activities as $(1-L, L)$ and makes capital investment, K , then her profit in state s is $sF(K, L) - RK = s(\sqrt{K} + \sqrt{L}) - RK$. Her employment income in state s is $w(1-L)$. We assumed that $w_s = w$ such that wages are neither positively nor negatively correlated with the state of nature.¹⁰ Her income from a risk-sharing arrangement in state s is A_s . We assume that the income from the risk-sharing arrangement has an expected payment of 0, $E_s A_s = 0$. Specifically, we assume that $A_s = A > 0$ when $s \leq 3$, and $A_s = -A < 0$ when $s \geq 4$.¹¹

Conditional on realized state s , the microentrepreneur's objective is to maximize her total income

¹⁰ 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) showed that many households use business activities as a diversification tool against possible negative 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.

¹¹ An implicit assumption here is that informal insurance payments cannot be used for investment, which is generally consistent with the empirical evidence showing that the most common reason for accepting such payments is to meet immediate consumption needs rather than for investment purposes. Only 3.8% of all gifts and 18.4% of informal loans are used for investment purposes (Fafchamps and Lund, 2003).

$$\max_{L_s} s(\sqrt{K} + \sqrt{L_s}) - RK + w(1 - L_s) + A_s,$$

so that $L_s^* = \min\{\frac{s^2}{4w^2}, 1\}$. The objective is total income and not utility because this decision is made after the uncertainty (i.e., the state of nature) is realized so that the microentrepreneur no longer faces any risk. The state-income function $I(s, K)$, therefore, is

$$I(s, K) = \underbrace{s(\sqrt{K} + \sqrt{L_s^*}) - RK}_{=\pi(s, K)} + \underbrace{w(1 - L_s^*) + A_s}_{=h(s)}. \quad (8)$$

Before the state of nature is realized, the microentrepreneur chooses capital to maximize her utility

$$U(K) = (1 - \eta)E_s I(s, K) + \eta \min_s I(s, K),$$

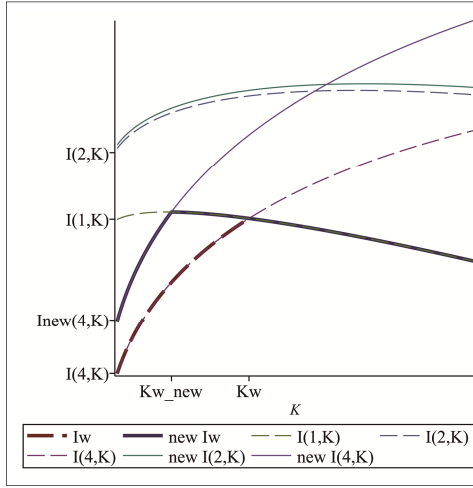
which is the same as (1).

Consider an additive improvement where the post-training profit is given by $s\lambda_s(\sqrt{K} + \sqrt{L}) - RK$, where $1 \leq \lambda_1 \leq \dots \leq \lambda_6$. We use 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 $\lambda_6 = 1.17$. One can verify that when $\eta = 1$, then not only is the capital adjustment effect negative, but, most importantly, the post-training expected income and post-training expected profit are also lower than pre-training expected income and pre-training expected profit, respectively.

Figures 1 and 2 visualize the example.¹² When $\eta = 1$ then only income in the worst-case states matter, which are states 1 and 4. States 2, 3, 5 and 6 are never worst-case states. Thus, in order to keep Figure 1 tractable, we plot $I(s, K)$ and $I^{new}(s, K)$ in state 2 only but not in states 3, 5 and 6. As Figure 1 shows, training makes state 4 more profitable but results in the negative capital adjustment effect, as $K_w^{new} \approx 0.54 < K_w \approx 1.17$. The negative capital adjustment effect dominates the profit improvement effect. Post-training expected income $E_s I^{new}(s, K_w^{new}) \approx 6.49 < E_s I(s, K_w) \approx 6.51$. Post-training expected profit is lower as well: $E_s \pi^{new}(s, K_w^{new}) \approx 6.37 < E_s \pi(s, K_w) \approx 6.39$.

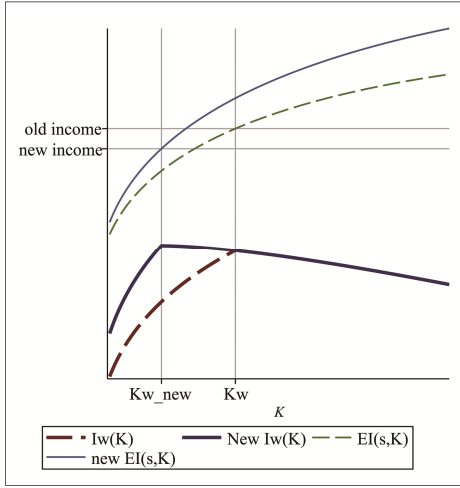
¹² For both figures, K is on the horizontal axis and income is on the vertical axis. In both Figures, the values of K are from interval $[0, 3]$. The values on the vertical axis are from $[2.2, 7.5]$ for Figure 1 and $[2.2, 8.5]$ for Figure 2. The numbers are picked merely for illustrative purposes.

[Figure 1]



Note: Thin dashed and solid curves are income functions in different states that are labeled using the vertical axis. Solid lines are post-training income functions, while dashed lines are pre-training income functions. Since $I(1,K) = I^{new}(1,K)$, the dashed line is used for both. The solid line above $I(2,K)$ is $I^{new}(2,K)$. The thick dashed line is $I_w(K) = \min\{I(1,K), I(4,K)\}$, the thick solid line is $I_w^{new}(K) = \min\{I^{new}(1,K), I^{new}(4,K)\}$.

[Figure 2]



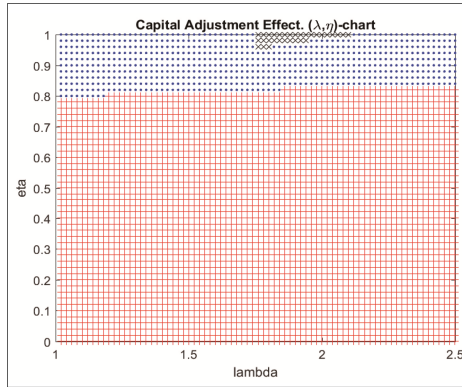
Note: Thin lines are expected incomes before and after the training. Thick lines are the worst-case incomes before and after the training. Dashed lines are for the pre-training expected income and worst-case income functions. Solid lines are for the post-training expected income and worst-case incomes. When $\eta = 1$ the microentrepreneur's choice of capital changes from Kw to Kw_new . New expected income is lower.

Figures 3 and 4 are plotted to study how robust the example above is to perturbations in parameters. In total, there are 13 parameters: $\{p_s\}_{s=1}^6$ and $\{\lambda_s\}_{s=1}^6$ and η . For the purpose of visualization, we parameterize $\{p_s\}$ and $\{\lambda_s\}$ to make them functions of one-dimensional variables p and λ , respectively. We set the probability of state 1 to be equal to p and the probabilities of states 2 through 6 equal to $(1-p)/5$. When $p = 1/6$, all states are equally likely. We define the effect in state s as $\hat{\lambda}_s = 1 + (\lambda_s - 1)(\lambda - 1)$, where $\{\lambda_s\}$ are parameters used to build Figure 1 and 2. When $\lambda = 1$, then the training has effect, $\hat{\lambda}_s = 1$. When $\lambda = 2$, then $\hat{\lambda}_s = \lambda_s$.

Figure 3 shows that a high-degree of ambiguity-aversion is needed in order to have the negative capital adjustment effect: for the total effect to be negative, one needs to have a very high η and an intermediate range of λ . When η is set equal to 1, as on Figure 4, the capital adjustment effect is negative for almost all parameter values. When p gets close to 1, then $K^* \rightarrow K^*(1)$ so that the condition $K_w < K^*$ in Proposition 8 is no longer satisfied, and the capital

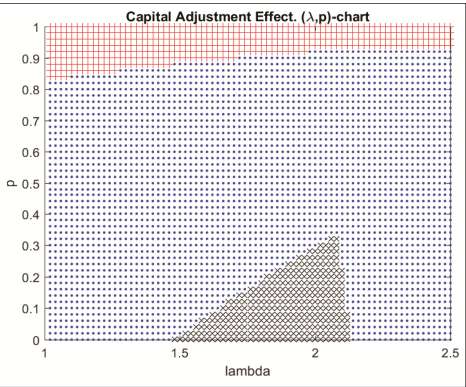
adjustment effect can be positive. The total effect is negative for low values of p , and intermediate values of λ . When λ is high then the profit improvement effect is also high and outweighs the capital adjustment effect. When λ is low, the capital adjustment effect is too small to outweigh the profit improvement effect. Thus, the total effect can be negative only for intermediate values of λ .

[Figure 3]



Note: The figure is built for $p = 1/6$. 'x' signs correspond to (λ, η) values where the post-training expected profit declines; '.' signs correspond to (λ, η) values where the capital adjustment effect is negative but the total effect is positive; '+' signs correspond to (λ, η) values where the capital adjustment is positive.

[Figure 4]



Note: The figure is built for $\eta = 1$. 'x' signs correspond to (λ, p) values where the post-training expected profit declines; '.' signs correspond to (λ, p) values where the capital adjustment effect is negative but the total effect is positive. '+' signs correspond to (λ, p) values where the capital adjustment is positive.

IV. Concluding Remarks

This paper provides a theoretical framework to understand the mixed impact of business training. We rely on a holistic view of a microentrepreneur as someone whose livelihood and goals are more complex than just being an entrepreneur. We model it using two assumptions. First, the microentrepreneur has several sources of income in addition to income from business activities. Second, the microentrepreneur has other ambitions in addition to maximizing business income.

The impact of the training varies depending on the microentrepreneurs' ambitions and the environment in which they operate. This is consistent with the observation that "*BDPs have been more successful for some entrepreneurs than the others*" (Verrest, 2013, p. 58). We further show that the reason behind the limited effect of business training is that BDPs' focus on growing microentrepreneurs' businesses ignores the non-business aspects of microentrepreneurs' livelihoods. For

microentrepreneurs who have a strong business-oriented ambition or whose only income source is profit from business activities, the training effect is always positive. When the microentrepreneur has other goals beyond profit maximization and other income sources beyond business activities, the training impact can be limited and even negative.

There are several limitations of our approach that are left for future research. First, we focus on one factor that could be responsible for the post-training expected profit decline. However, there are other factors that could lead to the same outcome (e.g., inefficiency of the training). Second, our framework cannot be used to explain the success of some of the training treatments studied in the literature. Drexler et al. (2014) showed that a simplistic rule-of-thumb training worked better than the more complex one that is commonly used by BDPs. Brooks et al. (2018) introduced the mentor treatment as an alternative to a standard business-training program, where trainees were mentored by a more experienced entrepreneur from the community. They showed that the mentor treatment, and only the mentor treatment, had a positive effect on the profit. Finally, our model is static and thus cannot be used to explain some dynamic phenomena documented in the literature. For instance, Karlan, Knight, and Udry (2012) reported that Ghana tailors switched to a new practice after the training but then abandoned it one year after the training stopped. The static framework in our paper could not be used to explain this short-run switch and the medium-run reversal.

Appendix. Proofs

Proof of Proposition 1: First, we prove the statement about K_w . 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 the one hand, $K_w \neq K^*(s')$. This is because

$$I(s', K^*(s')) > I(1, K^*(1)) > I(1, K^*(s')) \geq \min_s I(s, K^*(s')). \quad (9)$$

The first inequality is by the *states-ranking* assumption. The second inequality follows from $K^*(s') \neq K^*(1)$, and single-peakedness of $I(1, \cdot)$. The last inequality holds because income in state 1 is weakly greater than income in the worst-case state. 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 (9) 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.¹³ 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: By (A1), $U(K)$ is a concave function of K because the sum of concave functions is a concave function, and the minimum of concave functions is a concave function. Therefore, it has left and right derivatives and the

¹³ 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 .

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

$$(1-\eta)E_s I_K(s, K_\eta^*) + \eta \cdot (I_w(K_\eta^* -))' \geq 0 \geq (1-\eta)E_s I_K(s, K_\eta^*) + \eta \cdot (I_w(K_\eta^* +))'. \quad (10)$$

If the utility function is differentiable at K_η^* , then (10) becomes

$$(1-\eta)E_s I_K(s, K_\eta^*) + \eta \cdot (I_w(K_\eta^*))' = 0. \quad (11)$$

We prove that K_η^* is a decreasing function of η by contradiction.¹⁴ Assume it is not. Then there exist $\eta_1 < \eta_2$ such that $K_{\eta_1}^* < K_{\eta_2}^*$. Then

$$\begin{aligned} 0 &\geq (1-\eta_1)E_s I_K(s, K_{\eta_1}^*) + \eta_1 \cdot (I_w(K_{\eta_1}^* +))' \\ &> (1-\eta_1)E_s I_K(s, K_{\eta_2}^*) + \eta_1 \cdot (I_w(K_{\eta_2}^* -))' \\ &\geq (1-\eta_2)E_s I_K(s, K_{\eta_2}^*) + \eta_2 \cdot (I_w(K_{\eta_2}^* -))', \end{aligned}$$

which, is a contradiction since by (10) $K_{\eta_2}^*$ cannot be optimal given η_1 . Here, the first inequality follows from (10) and the fact that $K_{\eta_1}^*$ is optimal given η_1 , and the second inequality follows from the fact that the utility function is strictly concave and $K_{\eta_1}^* < K_{\eta_2}^*$. As for the last inequality, note that by (10), when $K^* > K_w$ then $K^* \geq K_\eta^* \geq K_w$ for every η . Therefore, $E_s I_K(s, K_{\eta_2}^*) \geq 0$ and $(I_w(K_{\eta_2}^* -))'_K \leq 0$. Given that $\eta_2 > \eta_1$, the last inequality is clearly satisfied. ■

Proof of Proposition 4: Part i): this part is trivial since, by definition, K^{*new} maximizes $E_s \pi^{new}(s, K)$ and, therefore, the capital adjustment effect is non-negative.

Part ii): when the microentrepreneur's only source of income is the business income then $I(s, K) = \pi(s, K)$. Given $\pi(1, 0) \leq \dots \leq \pi(n, 0)$ and the complementarity assumption, state 1 is the worst-case pre-training state for every K . The microentrepreneur's pre-training utility therefore is $(1-\eta)E_s \pi(s, K) + \eta\pi(1, K)$. From (BT5-I) and $\lambda_1 < \dots < \lambda_n$ follows that after the training state 1 is the worst-case when $K \in [0, K^*(n)]$. Thus, on this interval the microentrepreneur's post-training utility therefore is $(1-\eta)E_s \pi^{new}(s, K) + \eta\pi^{new}(1, K)$.

We prove that $K_\eta^* \leq K_\eta^{*new} \leq K^{*new}$ which combined with concavity of expected profit would guarantee that the capital adjustment effect is non-negative. First, we

¹⁴ It is not a strictly decreasing function of η . When the worst-case state is unique, as in (11), it is a strictly decreasing function of η . When it is not unique, as in (10), it is weakly decreasing. That is, there is a range of η 's that would correspond to the same optimal K_η^* .

prove $K_\eta^{*new} \leq K^{*new}$. Proof by contradiction. Assume not, $K_\eta^{*new} > K^{*new}$. On the one hand, $K_\eta^{*new} \leq K^*(n)$ and, therefore, it satisfies the following FOC:

$$(1-\eta)E_s \lambda_s \pi_K(s, K_\eta^{*new}) + \eta \lambda_1 \pi_K(1, K_\eta^{*new}) = 0. \quad (12)$$

On the other hand, if $K_\eta^{*new} > K^{*new}$ then (12) cannot be satisfied. Indeed, if $K_\eta^{*new} > K^{*new}$ then $E_s \lambda_s \pi_K(s, K_\eta^{*new}) < 0$. Furthermore, note that $\lambda_s \pi_K(s, K)$ has the same sign as $\pi_K(s, K)$. By the complementarity assumption $\pi_K(n, K) > \dots \pi_K(1, K)$ and, therefore, $E_s \lambda_s \pi_K(s, K_\eta^{*new}) < 0$ implies that $\lambda_1 \pi_K(1, K_\eta^{*new}) < 0$. Thus the LHS of (12) should be negative which is a contradiction. One can use the same reasoning to show that $K_\eta^* \leq K^*$.

Second, we prove that $K_\eta^* \leq K_\eta^{*new}$. Let τ be the lowest state such that $\pi_K(\tau, K_\eta^*) \geq 0$. By complementarity, $\pi_K(s, K_\eta^*) > 0$ when $s > \tau$ and $\pi_K(s, K_\eta^*) < 0$ when $s < \tau$. Then,

$$\begin{aligned} (1-\eta)E_s \lambda_s \pi_K(s, K_\eta^*) + \eta \lambda_1 \pi_K(1, K_\eta^*) &= (1-\eta)E_s (\lambda_s - \lambda_1) \pi_K(s, K_\eta^*) \geq \\ &\geq (1-\eta)(\lambda_\tau - \lambda_1) E_s \pi_K(s, K_\eta^*) \geq 0, \end{aligned}$$

and, therefore, $K_\eta^* \leq K_\eta^{*new}$. The last inequality follows from $\lambda_1 < \dots < \lambda_n$ and that $K_\eta^* \leq K^*$. ■

Proof of Proposition 5: By Proposition 1, if $K_w \neq K^*(1)$ there exist two states, $s < s'$ such that $I(s, K_w) = I(s', K_w) \leq I(t, K_w)$ for all other states t , and $I_K(s', K_w) < 0 < I_K(s'', K_w)$. Since $\{h(s)\}$ does not depend on K , $I_K(s, K) = \pi_K(s, K)$ for every s and K . In particular, $\pi_K(s', K_w) < 0 < \pi_K(s'', K_w)$.

First, we show that $K_\eta^{new} < K_w$. From $\lambda_1 < \dots < \lambda_n$ and (BT5-I) follows that $I^{new}(s', K_w) = \lambda_s \pi(s', K_w) + h(s') < \lambda_t \pi(t, K_w) + h(t) = I^{new}(t, K_w)$ for every $t > s'$. Thus, any post-training worst-case state given K_w is less than or equal to s' . Take one and denote it as s_w^{new} . From complementarity and $s_w^{new} \leq s'$ follows that $\pi_K(s_w^{new}, K_w) < 0$. For any $K > K_w$ then

$$\min_t I^{new}(t, K) \leq \lambda_{s_w^{new}} \pi(s_w^{new}, K) + h(s_w^{new}) < \lambda_{s_w^{new}} \pi(s_w^{new}, K_w) + h(s_w^{new}) = \min_t I^{new}(t, K_w).$$

Here, the first inequality is due to the fact that the worst-case income given K is less or equal than the income in s_w^{new} . The second inequality follows from the fact that $K > K_w$ and $\pi_K(s_w^{new}, K_w) < 0$. The final equality follows from the fact that s_w^{new} is the worst-case state given K_w . Therefore, if $K > K_w$ then $I_w^{new}(K) < I_w^{new}(K_w)$ so $K > K_w$ cannot be optimal capital for the worst-case post-training

come. Moreover, K_w is no longer the optimal for the worst-case post-training income either. Since all worst-case states are less or equal than s' it means that the corresponding state-income functions are decreasing at K_w . Then K slightly below K_w will give strictly higher worst-case income. It proves that $K_w^{new} < K_w$.

Next, we show that $K^* < K^{*new}$. Recall that since $h(s)$ does not depend on K it means that K^* maximizes both expected income and expected profit. Similarly, K^{*new} maximizes both post-training expected income and post-training expected profit. The post-training expected profit is single-peaked and concave. Therefore, to show $K^{*new} > K^*$ it is sufficient to show that $E_s \lambda_s \pi_K(s, K^*) > 0$, and we use here that derivatives with respect to K of expected income and expected profit are equal. Let τ be the smallest state such that $\pi_K(\tau, K^*) \geq 0$. Then

$$\begin{aligned} E_s \lambda_s \pi_K(s, K^*) &= \sum_{s < \tau} \lambda_s p_s \pi_K(s, K^*) + \sum_{s \geq \tau} \lambda_s p_s \pi_K(s, K^*) > \\ &> \lambda_\tau \sum_{s < \tau} p_s \pi_K(s, K^*) + \lambda_\tau \sum_{s \geq \tau} p_s \pi_K(s, K^*) > 0 \end{aligned}$$

Here we used (BT5-II), the fact the first sum is the summation of negative terms and the second sum is the summation of positive terms, and that K^* satisfies $E_s \pi_K(s, K^*) = 0$.

Thus, on the one hand, $K_w^{new} < K_w$, and on the other hand $K^* < K^{*new}$, and by assumption $K_w < K^*$. Then $E_s \pi^{new}(s, K_w^{new}) - E_s \pi^{new}(s, K_w) = E_s I^{new}(s, K_w^{new}) - E_s I^{new}(s, K_w) < 0$, where the inequality follows from the concavity of $E_s I(s, K)$ and $K_w^{new} < K_w < K^* < K^{*new}$. Thus, the capital adjustment effect is negative. ■

Proof of Proposition 6: First, we establish two lemmas.

Lemma 5.1 *When $K > K_w$ ($K^*(n) \geq K > K_w^{new}$) state 1 is a pre-training (post-training) worst-case state.*

Proof: First, consider pre-training income function. If $K_w = K^*(1)$ then by complementarity state 1 is the only worst-case state when $K > K_w$. 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, $I_K(s, K_w) < 0$, i.e. $K^*(s) < K_w$. By definition of $K^*(1)$ and the states-ranking assumption: $I(1, K^*(s)) < I(1, K^*(1)) < I(s, K^*(s))$. By complementarity, $I(1, K^*(s)) < I(s, K^*(s))$ implies that $I(1, K) < I(s, K)$ for every $K > K^*(s)$. This is contradiction to the fact that $K_w > K^*(s)$ and state s is a worst-case state given K_w .

By (A2) and (A4) state 1 has the lowest profit $\pi(s, K) > \pi(1, K)$. Thus as long as $\pi(1, K) > 0$ — which by (BT5-I) happens when $K \leq K^*(n)$ — from $I(s, K) > I(1, K)$ follows that $I^{new}(s, K) > I^{new}(1, L)$. It completes the proof. ■

Lemma 5.1 implies that $U(K) = (1-\eta)E_s I(s, K) + \eta I(1, K)$ when $K \geq K_w$ and $U^{new}(K) = (1-\eta)E_s I^{new}(s, K) + \eta I^{new}(1, K)$ when $K^*(n) \geq K \geq K_w^{new}$. Furthermore, the pre-training (post-training) utility is differentiable if $K > K_w$ ($K^*(n) > K > K_w^{new}$).

Lemma 5.2 $K_\eta^*(K_\eta^{*new})$ is a strictly decreasing functions of η if $K_\eta^* \neq K_w$ ($K_\eta^{*new} \neq K_w^{new}$).

Proof: When $K_\eta^* \neq K_w$ then $K_\eta^* > K_w$ (Proposition 2). The pre-training utility function is differentiable at K_η^* and $U'(K_\eta^*) = 0$. By the implicit function theorem:

$$\begin{aligned} \frac{\partial K_\eta^*}{\partial \eta} &= - \frac{-E_s \pi_K(s, K_\eta^*) + \pi'(1, K_\eta^*)}{(1-\eta)E_s \pi_{KK}(s, K_\eta^*) + \eta \pi_{KK}(1, K_\eta^*)} \\ &= - \frac{1}{1-\eta} \frac{\pi_K(1, K_\eta^*)}{(1-\eta)E_s \pi_{KK}(s, K_\eta^*) + \eta \pi_{KK}(1, K_\eta^*)} < 0. \end{aligned} \quad (13)$$

Here we use that $I_K(s, K) = \pi_K(s, K)$ since $h(s)$ does not depend on K , and we simplified the numerator using $U'(K_\eta^*) = 0$. The numerator is negative because $K_\eta^* > K^*(1)$, and the denominator is negative by concavity of state-profit functions. Similarly,

$$\frac{\partial K_\eta^{*new}}{\partial \eta} = - \frac{1}{1-\eta} \frac{\lambda_1 \pi_K(1, K_\eta^{*new})}{(1-\eta)E_s \lambda_s \pi_{KK}(s, K_\eta^{*new}) + \eta \lambda_1 \pi_{KK}(1, K_\eta^{*new})} < 0 \quad (14)$$

where we use that $K_\eta^{*new} \leq K^*(n)$. This completes the proof. ■

i) and iii): Let η^1 be such that $(1-\eta^1)E_s \pi_K(s, K_w) + \eta^1 \pi_K(1, K_w) = 0$, and η^2 be such that $(1-\eta^2)E_s \lambda_s \pi_K(s, K_w^{new}) + \eta^2 \lambda_1 \pi_K(1, K_w^{new}) = 0$. Both η^1 and η^2 are positive as long as $K_w < K^*$. Then for any $\eta < \min\{\eta^1, \eta^2\}$ it is the case $K_\eta^* > K_w$ and $K_\eta^{*new} > K_w^{new}$.

When $\eta < \min\{\eta^1, \eta^2\}$ the derivative of the total effect on the expected profit is

$$\begin{aligned} (E_s \pi^{new}(s, K_\eta^{*new}) - E_s \pi(s, K_\eta^*))'_\eta &= \frac{\partial K_\eta^{*new}}{\partial \eta} E_s \lambda_s \pi_K(s, K_\eta^{*new}) \\ &\quad - \frac{\partial K_\eta^*}{\partial \eta} E_s \pi_K(s, K_\eta^*). \end{aligned} \quad (15)$$

Consider a limit when $\lambda_n \rightarrow \infty$. Then $K_\eta^{*new} \rightarrow K^*(n)$, and

$$\frac{\partial K_\eta^{*new}}{\partial \eta} E_s \lambda_s \pi_K(s, K_\eta^{*new}) \rightarrow -\frac{1}{1-\eta} \frac{\lambda_1 \pi_K(1, K^*(n)) \pi_K(n, K^*(n))}{(1-\eta) \pi_{KK}(n, K^*(n))} = 0,$$

where the last equality follows from the fact that $K^*(n)$ maximizes $\pi(n, K)$. Thus, the first term in (15) converges to zero and the second term does not depend on λ 's and is negative. Therefore, the limit of (15) is positive and, in particular, (15) is positive for sufficiently large but finite λ_n .

Next look at the capital adjustment effect:

$$\begin{aligned} (E_s \pi^{new}(s, K_\eta^{*new}) - E_s \pi^{new}(s, K_\eta^*))'_\eta &= \frac{\partial K_\eta^{*new}}{\partial \eta} E_s \lambda_s \pi_K(s, K_\eta^{*new}) \\ &\quad - \frac{\partial K_\eta^*}{\partial \eta} E_s \lambda_s \pi_K(s, K_\eta^*). \end{aligned} \quad (16)$$

We proved in Proposition 5 that $K^* < K^{*new}$. Let $\eta \rightarrow 0$ so that $K_\eta^{*new} \rightarrow K^{*new}$ and $K_\eta^* \rightarrow K^*$. Then the first term in (16) converges to zero since $E_s \lambda_s \pi_K(s, K^{*new}) = 0$ and the second term does not, since $K^* \neq K^{*new}$. Therefore, (16) is positive when $\eta = 0$ and, by continuity is positive for sufficiently small η .

ii) and iv): We need to consider two cases. First, $K_w = K^*(1)$. Then $K_\eta^{new} = K^*(1)$ and

$$0 > \frac{\partial K_\eta^*}{\partial \eta} = \frac{E_s \pi_K(s, K^*(1))}{\pi_{KK}(1, K^*(1))} > \frac{E_s (\lambda_s / \lambda_1) \pi_K(s, K^*(1))}{\pi_{KK}(1, K^*(1))} = \frac{\partial K_\eta^{*new}}{\partial \eta},$$

at $\eta = 1$. One can then use (15) and (16) to verify that the derivative of the total effect and the derivative of the capital adjustment effect are negative at $\eta = 1$. By continuity there exists η^0 such that for any $\eta > \eta^0$ both are decreasing functions of η .

The second case, $K_w \neq K^*(1)$. Let η^1 and η^2 be defined as in part i). Then $K_\eta^* = K_w$ for any $\eta \geq \eta^1$, and $E_s \pi(s, K_\eta^*)$ is constant function of η for any $\eta > \eta^1$. Similarly $K_\eta^{*new} = K_w^{new}$ for any $\eta \geq \eta^2$, and $E_s \lambda_s \pi(s, K_\eta^{*new})$ is constant function of η for any $\eta > \eta^2$. Notice that $\eta^1 < 1$ as otherwise $K_w = K^*(1)$. One can use it to show that $\eta^1 < \eta^2$. Indeed,

$$(1 - \eta^1) E_s \lambda_s \pi_K(s, K_w) + \eta^1 \lambda_1 \pi_K(1, K_w) = (1 - \eta^1) E_s (\lambda_s - \lambda_1) \pi_K(s, K_w) > 0. \quad (17)$$

The equality follows from the definition of η^1 , namely, $(1-\eta^1)E_s\pi_K(s, K_w) + \eta^1\pi_K(1, K_w) = 0$. The inequality follows from the fact that $K_w < K^*$ and can be proved using the same technique as in the second part of Proposition 5. From (17) follows that $K_{\eta^1}^{*new} > K_w$. In Proposition 5 we established that $K_w > K_w^{new}$. Therefore, $\eta^2 > \eta^1$.

When $\eta \geq \eta^2$ then $K_{\eta}^* = K_w$ and $K_{\eta}^{*new} = K_w^{new}$ and they do not depend on η . Both capital adjustment effect and total effects are constants. When $\eta \in [\eta^1, \eta^2]$ then $K_{\eta}^* = K_w$ and does not depend on η but $K_{\eta}^{*new} > K_w^{new}$ and by Lemma 5.2 is a decreasing function of η . Then both the capital adjustment effect and the total effect are strictly decreasing functions of η . Combining the two results we get that the capital adjustment effect and the total effects are weakly decreasing functions of η when $\eta > \eta_1$.

v): Let $\lambda_1 = \dots = \lambda_n$. Then $K_{\eta}^* = K_{\eta}^{*new}$ for every η and one can use (15) to immediately verify that $(E_s\pi^{new}(s, K_{\eta}^{*new}) - E_s\pi(s, K_{\eta}^*))'_{\eta}$ 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 $(E_s\pi^{new}(s, K_{\eta}^{*new}) - E_s\pi(s, K_{\eta}^*))'_{\eta}$ is also non-positive. That completes the proof. ■

Proof of Proposition 7: Part i): this part is trivial since, by definition, K^{*new} maximizes $E_s\pi^{new}(s, K)$ and, therefore, the capital adjustment effect is non-negative.

Part ii): the proof follows the same steps as the proof of Proposition 4. It is, however, more complicated as $K^{*new}(s)$ is not necessarily equal to $K^*(s)$.

When $h(s) = 0$ then $I(s, K) = \pi(s, K)$ for every s . State 1 is the worst-state case before and after the training for every K . For the pre-training case, the statement immediately follows from (A2) and (A4). For the post-training case, the statement follows from Lemma 5.3 below. Therefore, the microentrepreneur's pre-training utility and post-training utilities are $(1-\eta)E_s\pi(s, K) + \eta\pi(1, K)$ and $(1-\eta)E_s\pi^{new}(s, K) + \eta\pi^{new}(1, K)$ respectively when $K \leq K^{*new}(n)$.

Lemma 5.3 Let s_w and s_w^{new} be the lowest worst states given K under the pre- and post-training state-profit functions respectively. For the additive improvement, $s_w^{new} \leq s_w$. For the multiplicative improvement $s_w^{new} \leq s_w$ if $K \leq K^{*new}(n)$.

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, 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 expected profit at state t is weakly higher, which implies that the new lowest worst-case state is s_w or lower. ■

We will prove that $K_\eta^* \leq K_\eta^{*new} \leq K^{*new}$ which combined with the concavity of expected profit would guarantee that the capital adjustment effect is non-negative. First, we prove that $K_\eta^{*new} \leq K^{*new}$. K_η^{*new} satisfies the FOC $(1-\eta)E_s \pi_K^{new}(s, K_\eta^{*new}) + \eta \pi_K^{new}(1, K_\eta^{*new}) = 0$. By Lemma 5.4 below, $K_\eta^{*new} > K^{*new}(1)$ when $\eta < 1$. 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 $K_\eta^{*new} < K^{*new}$ when $\eta < 1$. Then by continuity, $K_\eta^{*new} \leq K^{*new}$ when $\eta \leq 1$.

Lemma 5.4 $K_\eta^{*new} > K^{*new}(1)$ for every $\eta < 1$.

Proof. The proof is based on the following fact. Let s_w be a pre-training worst-case state given $K \leq K^{*new}(n)$. If $\pi_K^{new}(s_w, K) \geq 0$ then $\pi_K^{new}(s, K) > 0$ for any $s > s_w$. In the case of pre-training profit function, it would immediately follow from complementarity. But post-training profit function does not necessarily satisfy complementarity which is why it has to be proved.

First, we consider an additive improvement. From (6) 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 $\pi_K^{new}(s_w, K) \geq 0$ by the assumption. Thus $\pi_K^{new}(s, K) > 0$.

In the case of a multiplicative improvement

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

The first inequality holds because $\pi_K(s, K) > \pi_K(s_w, K)$, which is the complementarity assumption, and because profit functions are positive. The last inequality follows from two facts: the third line is non-negative because it is equal to $\pi_K^{new}(s_w, K)$ 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.

State 1 is the worst post-training case and, therefore, from $\pi_K^{new}(1, K^{*new}(1)) = 0$ follows that $\pi_K^{new}(s, K^{*new}(1)) > 0$ for every s . Plugging $K = K^{*new}(1)$ into the FOC we get $(1 - \eta)E_s \pi_K(s, K^{*new}(1)) + \eta \pi_K(1, K^{*new}(1)) > 0$ when $\eta < 1$. Thus, by concavity $K_\eta^{*new} > K^{*new}(1)$ when $\eta < 1$. ■

Now, we prove that $K_\eta^* \leq K_\eta^{*new}$. To do that we take derivative of the post-training utility function at K_η^* and show that it is non-negative. Concavity then would imply $K_\eta^* \leq K_\eta^{*new}$.

Consider an additive improvement. The derivative of the post-training utility function is

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

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 a multiplicative improvement, the post-training utility can be written as

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

Here we use that $K_\eta^* \leq K^*(n) \leq K^{*new}(n)$ so that state 1 is the worst-case state. Its derivative is

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

The first term is positive because multiplicative improvement requires that $\pi^{new}(1, K)/\pi(1, K)$ is an increasing function of K . Since K_η^* maximizes the pre-training utility the second term is equal

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

We will show that expression in parenthesis is positive whenever $\eta < 1$. Let τ be the lowest state such that $\pi_K(\tau, K_\eta^*) \geq 0$. The pre-training profit function satisfies complementarity and, therefore,

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

The last inequality follows from the earlier established fact that $K_\eta^* < K^*$ and that from the definition of the multiplicative improvement follows that $g(\tau) \geq 1$. Thus $K_\eta^* < K_\eta^{*new}$ when $\eta < 1$ and by continuity $K_\eta^* \leq K_\eta^{*new}$ when $\eta \leq 1$.

This completes the proof since $K_\eta^* \leq K_\eta^{*new} \leq K^{*new}$ implies that the capital adjustment effect is non-negative. ■

Proof of Proposition 8: Because of concavity of the utility function, 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 $K_w^{new} < K_w$. Since $K_w \neq K^*(1)$ it follows from Proposition 1 that there exist two states, $s < s'$, such that $I(s, K_w) = I(s', K_w) \leq I(t, K_w)$ for all other states t , and $I'_K(s, K_w) < 0 < I'_K(s', K_w)$. Let s_w and s_w^{new} the lowest worst states given K_w before and after the training respectively. By Lemma 5.3 $s_w \geq s_w^{new}$.

Proposition 8 requires that $K^{*new}(s_w) < K_w$. One can show that from $K^{*new}(s_w) < K_w$ follows that $K^{*new}(s_w^{new}) < K_w$. In the case of an additive improvement it directly follows from the definition. Indeed, since $s_w \geq s_w^{new}$ we have

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

From $K_w > K^{*new}(s_w)$ follows that $I_K^{new}(s_w, K_w) < 0$, which combined with the inequality above implies that $I_K^{new}(s_w^{new}, K_w) < 0$ and, therefore, $K^{*new}(s_w^{new}) < K_w$. In the case of a multiplicative improvement, we use the states-ranking assumption. If $s_w = s_w^{new}$ then we are done. The case $s_w > s_w^{new}$ is impossible. Indeed, if $s_w > s_w^{new}$ 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 states-ranking assumption. The second inequality follows from the fact that $K^*(s_w^{new})$ is optimal in s_w^{new} . By definition of s_w , $I'_K(s_w, K_w) < 0$ and, therefore, $K_w > K^*(s_w)$. One can then use complementarity to conclude that $I(s_w, K_w) > I(s_w^{new}, L_w)$. 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 income given K is less or equal than the income at s_w^{new} . The second inequality comes from the fact that $I^{new}(s_w^{new}, \cdot)$ declines when $K > K_w$. Thus $K_w^{new} \leq K_w$.

Furthermore, K_w is no longer the optimal worst-case capital either. By definition of s_w , $I(s_w, K_w) \leq I(t, K_w)$ for every t , including $t > s_w$. By the Proposition's assumption, all inequalities in (6) 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 .¹⁵ Take any

¹⁵ Note that this result is different from Lemma 5.3 which is only about the two lowest worst-case states. This result is about at all post-training worst-case states. It requires stronger assumptions which are that either inequalities in (6) are strict in the case of the additive improvement, and $g(s)$ is

post-training worst-case state given K_w , and denote it as \hat{s}_w^{new} . We know that $\hat{s}_w^{new} \leq s_w$. Then $K^{*new}(\hat{s}_w^{new}) < K_w$ by the exact same reasoning as in the beginning of the proof of Proposition 8. Thus all state functions that correspond to the worst-case states are strictly decreasing 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 7 it was established that $K_\eta^* \leq K_\eta^{*new}$ for any $\eta < 1$. When all inequalities in (6) are strict in the case of the additive improvement, or $g(s)$ is strictly increasing in the case of multiplicative improvement it is trivial to show that $K_\eta^* < K_\eta^{*new}$ for every $\eta < 1$. In particular, the inequality holds for $\eta = 0$ which means that $K^* < K^{*new}$. This completes the proof. ■

strictly increasing in the case of multiplicative improvement.

References

- Ahlin, C., and B. Waters (2014), “Dynamic Microlending under Adverse Selection: Can it Rival Group Lending?” *Journal of Development Economics*.
- Ambrus, A., M. Mobius, and A. Szeidl (2014), “Consumption Risk-sharing in Social Networks,” *American Economic Review*, 104(1), 149–182.
- Baillon, A., H. Bleichrodt, U. Keskin, O. l’Haridon, and C. Li (2018), “The Effect of Learning on Ambiguity Attitudes,” *Management Science*, 64(5), 2181–2198.
- 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.
- Banerjee, A., and E. Duflo (2011), “Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty,” *Public Affairs*.
- Barham, B. L., J. P. Chavas, D. Fitz, V. R. Salas, and L. Schechter (2014), “The Roles of Risk and Ambiguity in Technology Adoption,” *Journal of Economic Behavior & Organization*, 97, 204–218.
- Bateman, M., and H. J. Chang (2009), “The Microfinance Illusion,” Available at SSRN 2385174.
- Brooks, W., K. Donovan, and T. R. Johnson (2018), Mentors or Teachers? Microenterprise Training in Kenya *AEJ: Applied Economics*, 10 (4).
- Bruhn, M., and B. Zia (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.
- Bryan, G. (2019), “Ambiguity Aversion Decreases the Impact of Partial Insurance: Evidence from African Farmers,” forthcoming in *Journal of the European Economic Association*.
- Chateauneuf, A., J. Eichberger, and S. Grant (2007), “Choice under Uncertainty with the Best and Worst in Mind: Neo-additive Capacities,” *Journal of Economic Theory*, 137(1), 538–567.
- Cho, Y., and M. Honorati (2014), “Entrepreneurship Programs in Developing Countries: A Meta Regression Analysis,” *Labour Economics*, 28, 110–130.
- Chowdhury, P. R. (2005), “Group-lending: Sequential Financing, Lender Monitoring and Joint Liability,” *Journal of Development Economics*, 77(2), 415–439.
- De Mel, S., D. McKenzie, and C. Woodruff (2008), “Returns to Capital in Microenterprises: Evidence from a Field Experiment,” *The Quarterly Journal of Economics*, 123(4), 1329–1372.
- De Mel, S., D. McKenzie, and C. Woodruff (2014), “Business Training and Female Enterprise Start-up, Growth, and Dynamics: Experimental Evidence from Sri Lanka,” *Journal of Development Economics*, 106, 199–210.
- De Quidt, J., T. Fetzner, and M. Ghatak (2016), “Group Lending without Joint Liability,” *Journal of Development Economics*, 121, 217–236.
- De Weerd, J., and S. Dercon (2006), “Risk-sharing Networks and Insurance Against Illness,” *Journal of Development Economics*, 81(2), 337–356.
- Drexler, A., G. Fischer, and A. Schoar (2014), “Keeping it Simple: Financial Literacy and

- Rules of Thumb,” *American Economic Journal: Applied Economics*, 6(2), 1–31.
- Elabed, G., and M. R. Carter (2015), “Compound-risk Aversion, Ambiguity and the Willingness to Pay for Microinsurance,” *Journal of Economic Behavior & Organization*, 118, 150–166.
- Ellis, F. (2000), *Rural Livelihoods and Diversity in Developing Countries*, Oxford University Press.
- Engle-Warnick, J., J. Escobal, and S. Laszlo (2007), *Ambiguity Aversion as a Predictor of Technology Choice: Experimental Evidence from Peru* (No. 2007s-01). CIRANO.
- Fafchamps, M., and S. Lund (2003), “Risk-sharing Networks in Rural Philippines,” *Journal of Development Economics*, 71(2), 261–287.
- Ghosh, P., and D. Ray (2001), “Information and Enforcement in Informal Credit Market,” Unpublished Manuscript, Department of Economics, New York University.
- Gilboa, I., and M. Marinacci (2016), “Ambiguity and the Bayesian Paradigm,” In *Readings in Formal Epistemology* (pp. 385–439), Springer, Cham.
- Gilboa, I., and D. Schmeidler (1989), “Maxmin Expected Utility with a Nonunique Prior,” *Journal of Mathematical Economics*, 18(2), 141–153.
- Giné, X., and G. Mansuri (2014), “Money or Ideas? A Field Experiment on Constraints to Entrepreneurship in Rural Pakistan,” *A Field Experiment on Constraints to Entrepreneurship in Rural Pakistan (June 1, 2014)*, World Bank Policy Research Working Paper, (6959).
- Karlan, D., and M. Valdivia (2011), “Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions,” *Review of Economics and Statistics*, 93(2), 510–527.
- Karlan, D., R. Knight, and C. Udry (2012), *Hoping to Win, Expected to Lose: Theory and Lessons on Micro Enterprise Development* (No. w18325), National Bureau of Economic Research.
- Klibanoff, P., M. Marinacci, and S. Mukerji (2005), “A Smooth Model of Decision Making under Ambiguity,” *Econometrica*, 73(6), 1849–1892.
- Krishna, A. (2004), Escaping Poverty and Becoming Poor: Who Gains, who Loses, and why?” *World Development*, 32(1), 121–136.
- 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.
- McKenzie, D., and C. Woodruff (2014), “What are we Learning from Business Training and Entrepreneurship Evaluations Around the Developing World?” *The World Bank Research Observer*, 29(1), 48–82.
- Prediger, S., and G. Gut (2014), “Microcredit and Business-training Programs: Effective Strategies for Micro-and Small Enterprise growth?” GIGA Focus, Number 3, 2014.
- Shapiro, D. A. (2015), “Microfinance and Dynamic Incentives,” *Journal of Development Economics*, 115, 73–84.
- Valdivia, M. (2012), *Training or Technical Assistance for Female Entrepreneurship? Evidence from a Field Experiment in Peru*, GRADE working paper.
- Valentine, T. R. (1993), “Drought, Transfer Entitlements, and Income Distribution: The

- Botswana Experience,” *World Development*, 21(1), 109–126.
- Verrest, H. (2013), “Rethinking Microentrepreneurship and Business Development Programs: Vulnerability and Ambition in Low-income Urban Caribbean Households,” *World Development*, 47, 58–70.
- Yunus, M. (1999), *Banker to the Poor*, New York: Public Affairs, 1999.