

**AI Mentorship and the Language of Business:  
Building Financial Capability among Microenterprises in Developing  
Economies**

Marcela Aguilar  
Rice University  
[kma6@rice.edu](mailto:kma6@rice.edu)

December 2025

I am especially thankful to my dissertation committee for their guidance: Yael Hochberg, K. Ramesh (Chair), Shiva Sivaramakrishnan, and Xun Tang. I acknowledge the comments and suggestions provided by Cathy Shakespeare, Eric Floyd, the participants at the “Conference on Accounting and Information Frictions in Microenterprises” held at Rice University and the workshop participants at the National University of Singapore. The empirical design and hypotheses presented here are guided by the pre-analysis plan registered at OSF ([osf.io/r945e](https://osf.io/r945e)). This study received exemption from Rice University’s IRB under the category of minimal risk (IRB-FY2025-292). I thank Ramesh for being the Principal Investigator of the RCT and for facilitating the process with care and precision. I also gratefully acknowledge the support of the International Labour Organization (ILO), through its Office in San José, Costa Rica, and the Department of Employers' Activities (ACT/EMP). I am particularly thankful to Randall Arias, Senior Specialist at ACT/EMP, for his invaluable guidance and encouragement throughout the project. I appreciate the collaboration of the Chamber of Commerce and Industry of Tegucigalpa, Honduras (CCIT) for their openness to academically evaluate an innovative approach to business training for microenterprises, specially to Blanca Salgado and Karla Ruiz for their continuous support and enthusiasm.

# **AI Mentorship and the Language of Business: Building Financial Capability among Microenterprises in Developing Economies**

## **Abstract**

This paper examines whether AI mentorship framed around accounting as the *language of business* can build financial capability among microentrepreneurs in developing economies. Using a randomized controlled trial with microenterprises in Honduras, I evaluate the impact of an AI-based mentorship program that delivers personalized business guidance through WhatsApp. The AI mentor, powered by a large language model and adapted from the International Labour Organization's *Start and Improve Your Business curriculum*, provides real-time, on-demand support tailored to each entrepreneur's needs. The experimental design includes a treatment arm of AI mentor that frames accounting as a *language of business*, integrating financial reasoning across decision areas, and a control arm of AI mentor that teaches accounting as a stand-alone skill, as done in most business training programs. The treatment increases the salience and perceived relevance of financial reasoning, confirming that the framing mechanism operates as intended. Based on high-frequency conversational data, I show that entrepreneurs exposed to the accounting-framed AI mentor display greater integration of financial reasoning into marketing and business strategy decision-making processes, thereby opening the *black box* of how entrepreneurs acquire and apply business knowledge. They are also more likely to develop growth-oriented and forward-looking business plans, demonstrating the impact on actionable outcomes. When compared to traditional in-person training (third arm in the experiment), both versions of AI mentors lead to a greater perception of business growth, higher-quality business plans and greater adoption of cost analysis and recordkeeping practices, particularly among entrepreneurs starting from weaker managerial baselines. Overall, my results suggest that framing accounting as the language of business helps entrepreneurs internalize financial reasoning and that AI-based business mentorship can match or exceed human-led approaches while scaling at low marginal cost.

**Keywords:** Entrepreneurship, education, field experiment, business training, AI-based mentor, developing economies, microenterprises, recordkeeping, informal economy, accounting

## **I. Introduction**

Using a randomized control trial (RCT) of microenterprises in Honduras, this study examines whether an AI-based personalized and scalable digital mentoring intervention can build financial capability among microentrepreneurs in developing economies. I investigate what constitutes the essential financial mindset for these entrepreneurs and whether they benefit more from learning accounting within an integrated framework as a *language of business* that is “of enormous help to anyone evaluating the worth of a business and tracking its progress” (Buffett, 1987) rather than through the siloed instruction that characterizes most microentrepreneur training programs.

Microenterprises generate up to 70–80% of employment in developing economies and contribute between one-third and nearly half of their total GDP (ILO, 2019; World Bank, 2021). Despite their macroeconomic significance, most microenterprises operate with weak management practices, limited managerial capacity, and underdeveloped systems for financial decision-making (Aguilar et al., 2025; McKenzie & Woodruff, 2014). In response, governments and development organizations collectively spend over \$1 billion each year on business training programs, with the expectation that enhancing the capabilities of microbusinesses will promote their economic success and, in turn, contribute to broader economic development (McKenzie, 2021). Yet, despite the popularity of these interventions, rigorous evaluations reveal modest and often insignificant effects (McKenzie et al., 2021).

Most existing studies continue to evaluate classroom-based training programs that aim to simplify business concepts for entrepreneurs, often through “rule-of-thumb” lessons, designed to make the material more practical and accessible. While these approaches may improve comprehension, there is no evidence that they transform the underlying reasoning entrepreneurs

use in managing their businesses. Specifically, although accounting is often described as the *language of business* and central to firm decision-making, it is typically taught in business training programs as a stand-alone or siloed technical skill rather than integrated into broader business reasoning, thereby limiting its adoption and perceived value (Anderson et al., 2018; Arraiz et al., 2019; Drexler et al., 2014; Ramesh, 2025). My first research question (RQ1) examines whether training interventions can systematically alter entrepreneurs' financial decision-making framework, that is, how they process accounting information and incorporate it into managerial decisions.

Answering this question requires moving beyond traditional evaluation methods, which typically focus on reduced-form outcomes such as sales or profits captured in post-treatment surveys (Anderson et al., 2019; de Mel et al., 2009; Karlan et al., 2025; Small Firm Diaries, 2024). These measures are highly volatile and subject to reporting error, making it difficult to observe how entrepreneurs actually think and apply knowledge during the treatment period. Observing such intermediate mechanisms offers more precise insight into how entrepreneurs build their managerial capabilities. My study directly captures these cognitive and behavioral processes through high-frequency conversational data between entrepreneurs and the AI mentor, which helps unpack the *black box* of the learning process to assess whether entrepreneurs internalize accounting as a *language of business* and apply this framework in their managerial decisions.

The AI-based mentor in this project is powered by a large language model (GPT-4) and built upon the International Labour Organization's *Start and Improve Your Business* (SIYB-ILO) curriculum, providing personalized guidance via WhatsApp.<sup>1</sup> Overall, the AI-mentor is designed

---

<sup>1</sup>The AI Mentor is deployed through [Atomic AI's platform](#), which enables the implementation of different treatment variations (i.e., group 1 and group 2) and supports systematic data collection for the experiment. It employs a Retrieval-Augmented Generation (RAG) architecture that leverages an initial profile of each entrepreneur to deliver a personalized and dynamic mentoring accessible via WhatsApp, a widely used platform in developing economies (Otis et al., 2024; Ramdas & Sungu, 2022). The system offers

to address the key limitations of most business training programs: it is scalable (Al-Ubaydli et al., 2017; List, 2023), capable of delivering content of consistent quality (McKenzie et al., 2021), and acting as a knowledge-on-demand tool (Camuffo et al., 2019; Mollick & Mollick, 2022), thereby delivering consistent quality at low marginal cost.<sup>2</sup>

To test my RQ1, I partner with ILO and the Chamber of Commerce of Tegucigalpa (CCIT), Honduras, and implement a business training program through an RCT with the two versions of the AI mentor, identical in quality, scalability, and SIYB-ILO curricula<sup>3</sup>: the control arm with an AI mentor that is designed for a siloed approach to accounting education for business training (G1) and a treatment arm with an AI mentor that provides guidance that is couched in the financial/accounting language, explicitly linking their business decisions to financial analysis (G2). Additionally, all program participants have an incentive to engage, as an external donor offers a seed capital competition for entrepreneurs who complete the program. This design allows me to isolate the effects of framing and examine whether integrating accounting into broader business reasoning strengthens entrepreneurs' financial acumen.<sup>4</sup> By doing so, my experiment directly addresses a central gap in the development-economics literature on business training: the inability to isolate the causal role of accounting in driving microenterprise performance (Bernard & Sutherland, 2025).

As a validation of the experimental design, first, I examine whether the treatment framing increases the salience and perceived relevance of financial reasoning, as framing alone may be

---

24/7, context-aware support, responding to entrepreneurs' questions in real time to fully enable knowledge-on-demand through multiple interaction modes, including text, images, and voice notes.

<sup>2</sup> In Honduras, the business training program currently offered by the Chamber of Commerce and analyzed in this study "Formaliza Tu Negocio" has an estimated average cost of \$1,600 per participant (ILO, 2023). By contrast, the average cost per participant with the AI mentor is approximately \$40, see section VII for more details about the cost effectiveness of the AI mentor.

<sup>3</sup> SIYB program is a standard business training program that has been implemented in more than 100 countries (Van Lieshout & Mehtha, 2017).

<sup>4</sup> Because most microbusinesses operate as projects in place focused on current cash flows, rigorous accounting and financial analysis are essential for tracking and improving performance (McConnell & Muscarella, 1985, p. 400 ; Myers, 1977, p.171).

insufficient if accounting remains abstract or disconnected from entrepreneurs' immediate challenges. I find that the framing mechanism is functioning as intended: entrepreneurs in the treatment group significantly increase their demand for accounting information, sending ten additional messages, equivalent to roughly 58 percent of the control-group mean. As a placebo test, I find no corresponding increase in messages related to marketing content, suggesting that the higher demand for accounting information is consistent with the intended treatment. Unlike a vaccination study, where all treated participants receive identical intervention, my setting allows entrepreneurs to choose the aspects of the AI mentor's content most relevant to their needs. Consistent with this design, the treatment effect of the accounting-framed AI mentorship is concentrated in areas directly tied to decision-making, with 80 percent of the overall increase in accounting-related demand due to higher engagement with profitability analysis, rather than procedural aspects of accounting. This experimental validation shows that entrepreneurs can be nudged to engage more deeply with financially relevant knowledge during training.

However, the validation test does not show whether the treatment had the intended benefit (RQ1). To test this formally, I construct an accounting integration measure to quantify the extent to which accounting concepts (e.g., margins, cash flows, costs) co-occur with discussions of key functional areas of the business, such as marketing, business strategy, and business planning.<sup>5</sup> A higher integration indicates that accounting is not being treated as an isolated or compliance task, but rather as a cognitive framework used to reason through functional and managerial challenges. If the accounting-focused AI-based mentorship fosters this shift, I would expect to observe greater integration of financial reasoning across decision areas.

---

<sup>5</sup> See Appendix C for a real example of a conversation with accounting integration.

I find that entrepreneurs using the accounting-framed AI mentor are more likely to integrate financial reasoning across functional domains of their businesses, sending on average about 11 more accounting-integrated messages than those using the general AI mentor, representing 78 percent of the sample mean of the control group. Entrepreneurs in the treatment group more frequently integrated accounting insights when engaged in conversations relating to marketing and business strategy.<sup>6</sup> Overall, the evidence highlights that when microentrepreneurs internalize accounting as the language of business, they begin to use it as a cognitive framework for evaluating trade-offs across functional areas, rather than as a compliance exercise detached from managerial decision-making.

I next examine whether this cognitive shift translates into concrete behavioral responses. Specifically, I test whether entrepreneurs exposed to the accounting-framed mentorship are more likely to formulate actionable growth plans relative to those in the control group. Using high-frequency conversation data, I show that treated entrepreneurs are significantly more likely to articulate growth-oriented actions related to both the intensive margin (extensive margin), with 47 (90) percent more likely than the sample mean.<sup>7</sup> Moreover, when restricting attention to long-term plans, I find larger effects, indicating that accounting, when framed as the language of business, not only shapes immediate operational adjustments but also fosters strategic, forward-looking decision-making by microenterprises. Taken together, these findings suggest that when entrepreneurs learn to interpret accounting in an integrative decision framework, it meaningfully shifts both the content and horizon of their business actions.

---

<sup>6</sup> As a placebo test, I find no significant difference between control and treatment when entrepreneurs are engaging with the AI mentors solely on accounting issues in an unintegrated fashion.

<sup>7</sup> See Appendix D for a real example of an entrepreneur's conversation and actions identified.

Unlike my AI mentors, the traditional human-led training programs are incapable of contemporaneously capturing the cognitive and behavioral processes that are key to understanding the development of managerial capabilities, which is a key contribution of my study. However, a natural follow-up question is whether scalable, knowledge-on-demand AI-based training models can generate better or comparable business outcomes than traditional human-led training, and whether an AI model that emphasizes financial acumen provides incremental benefits (RQ2).

While the AI-based mentor is specifically designed to address well-documented limitations of traditional training, the theoretical direction of its impact remains unclear. On one hand, building on recent evidence that timely, well-structured advice plays a critical role in shaping entrepreneurial decision-making and contrasting with predefined chatbots and SMS-based interventions that often suffer from low engagement and limited personalization, the AI-based mentor model offers continuous, adaptive support tailored to individual needs (Cohen & Koning, 2024; de Barros et al., 2024; Mehmood, 2024). On the other hand, the absence of human interaction may reduce motivation, perceived credibility, or emotional support, especially in contexts where trust and social reinforcement play a key role in learning (Tomy & Wittenberg-Moerman, 2024). This tension creates genuine uncertainty about whether AI-driven training can fully substitute for or surpass the effects of traditional in-person instruction, particularly among microentrepreneurs operating with limited digital familiarity or facing structural barriers.

To test this RQ2, my RCT includes a third arm: a traditional human-led training program (G3), which serves as a benchmark to assess the relative effectiveness of the AI mentor variants (G1 and G2) based on pre- and post-treatment survey data, as done in most studies of training programs (McKenzie et al., 2021). Entrepreneurs in G3 received the standard training offered by the ILO and the CCIT, based on the same SIYB content. Therefore, the primary difference between

this group and the AI mentor groups is the delivery method. However, given that this test is based on survey data obtained only at two points in time, it is subject to the measurement limitations discussed in extant research (Anderson et al., 2019; de Mel et al., 2009; Fafchamps et al., 2012; Karlan et al., 2025; McKenzie et al., 2021).

In terms of outcomes, I consider three types: perceived business growth and traditional reduced-form outcomes (sales/profits), adoption of business practices, and the quality of business plans. Managerial perceptions may be less affected by measurement error and the short treatment duration that limit the usefulness of reduced-form outcomes, but they are inherently subjective and may not fully capture economic realities. Although the adoption of business practices does not provide direct evidence of productivity gains, it aligns with the study's focus on understanding the intermediate building blocks underlying the development of managerial capabilities.

In my empirical analysis, I find that entrepreneurs in the accounting-focused AI mentorship group (G2) report significantly higher *perceived business growth* relative to those in the human-led training group (G3), with an estimated increase of 0.16, after controlling for pre-treatment values (McKenzie, 2012).<sup>8</sup> Despite this perceived growth, monthly sales and profits remain unchanged, consistent with the well-documented volatility and measurement error inherent in these reduced-form indicators, especially over the short duration of the intervention (de Mel et al., 2009). Regarding *business practices*, participation in the accounting-focused AI mentorship (G2) increases the likelihood of entrepreneurs enhancing their cost analysis practices by 28 percent relative to the human-led training group (G3). The effects are even stronger among entrepreneurs whose baseline business practices were below the mean: they are more than 25 percent more likely to regularly calculate profits and losses and maintain systematic accounting records compared with

---

<sup>8</sup> While the accounting-framed AI mentor outperforms the human-led program, a formal test of the difference between the two AI models indicates that their coefficients are not statistically distinguishable at conventional levels.

the human-led group.<sup>9</sup> These findings align with earlier evidence that the accounting-focused AI mentor (G3) increases demand for accounting knowledge, especially profitability analysis, and that this engagement translates into tangible improvements in managerial practices.

I next provide descriptive evidence on the quality of business plans submitted by entrepreneurs across training modalities. At the beginning of the program, participants were required to prepare a comprehensive business plan using a standardized CCIT template to qualify for a seed capital competition.<sup>10</sup> Unlike the outcomes analyzed in RQ2, this analysis provides descriptive evidence on the overall managerial capabilities of the entrepreneurs as evidenced by the quality of their business plans. I employed a Large Language Model as an independent “judge” (LLMaJ) to systematically evaluate all submitted business plans (Csaszar et al., 2024).<sup>11</sup> The results show that, relative to the human-led training, entrepreneurs in both AI mentorship produce higher-quality business plans. The standard AI (accounting-focused AI) raises the total score by 11 (4) points out of 120 total points, with large, significant gains in Marketing, Cost, Profits, Legal, and Human Resources. This analysis illustrates a novel approach to studying the development of higher-order managerial capabilities among microentrepreneurs, moving beyond atomistic productivity measures toward more integrative indicators of managerial quality.

Finally, I examine which aspects of the program entrepreneurs find most valuable across the three arms. Entrepreneurs in both AI groups valued the mentor’s continuous support and flexible access, unlike the fixed schedule of the human-led training. The human-led group, however, offered stronger networking opportunities and a clearer sense of progress through its

---

<sup>9</sup> See Appendix F for examples of how entrepreneurs keep their records.

<sup>10</sup> The plan encompassed multiple dimensions, marketing, cost analysis, sales, profitability, legal structure, and human resources (see more in Appendix E) and required Excel-based forecasting, a demanding task for entrepreneurs with limited formal education or digital experience.

<sup>11</sup> The LLMaJ received each plan along with the structured evaluation form and rubric, producing section-level scores and concise qualitative comments. To validate the automated evaluation, CCIT coaches independently reviewed a subset of business plans, and their comments were highly consistent with those of the LLM, confirming alignment between human and model-based evaluations.

structured sessions. Future research could examine whether a human-AI hybrid training program can offer a better cost-benefit trade-off in enhancing the business capabilities of microentrepreneurs.

The most directly related study is Otis et al. (2024), which tests whether generative AI improves business outcomes relative to no support.<sup>12</sup> In contrast, my study examines whether the way business knowledge is framed, specifically, presenting accounting as the language of business, shapes how entrepreneurs internalize financial concepts and apply them to managerial decision-making. It also evaluates whether an AI-based mentor can serve as a feasible and effective alternative to traditional human-led training.

Building on this foundation, this study makes several contributions. It contributes to the accounting literature by empirically testing whether accounting, when positioned as a cross-functional logic, can shape managerial decision-making in microenterprises and by identifying the essential financial mindset as the ability to use accounting as a cognitive framework. Conducting an RCT through an AI-based model enables the systematic observation of mentor-entrepreneur interactions, opening the black box of the learning process that would otherwise be difficult, if not impossible, to capture in traditional settings. AI is transforming not only what we can study but also how we conduct research (Chang et al., 2024; Charness et al., 2023). It also responds to broader calls in the literature to explore the “AI-enabled entrepreneurial process” (Giuggioli & Pellegrini, 2022). In doing so, it addresses long-standing challenges in traditional business training

---

<sup>12</sup> Their study relies on generic ChatGPT output without Retrieval-Augmented Generation (RAG) or structured content, and it does not incorporate user profiling or memory. Additionally, they use a conversational interface where entrepreneurs primarily select from pre-defined options rather than engaging in open-ended dialogue, which constrains the richness of interaction. On average, entrepreneurs sent only four messages over a 2.5-month period. In addition, the tool does not support multimodal inputs such as audio or images. These features may particularly limit accessibility for entrepreneurs with lower literacy levels or those less comfortable with text-based digital tools.

programs, particularly their limited application, one-size-fits-all structure, and inability to scale personalized guidance in a cost-effective way.

## II. Related Literature

This paper builds on the extensive literature evaluating business training programs for microenterprises in developing economies and identifies four structural barriers that limit their effectiveness.<sup>13</sup> First, most programs rely on skilled human trainers, are delivered in small cohorts, and are costly to replicate (Al-Ubaydli et al., 2017; List, 2023). These models have yet to resolve a central question: how can training reach a broad cross section of entrepreneurs while maintaining instructional quality and achieving meaningful learning outcomes?

Second, training quality remains difficult to define and to standardize. Even if business training programs could be scaled, curricula differ in scope and pedagogy, often emphasizing short-term outcomes rather than how entrepreneurs learn and apply knowledge. The evidence in the extant literature increasingly suggests that “more” content is not necessarily “better”; however, even well-intentioned simplifications can backfire when they are overly generic (e.g., Cole et al., 2024; Batista et al., 2022). Together, these findings underscore that effective training depends less on the amount or simplicity of generic content and more on how entrepreneurs’ actually learn and apply knowledge. Third, educational research emphasizes the beneficial role of *active knowledge construction* vs. passive information reception (Mollick & Mollick, 2023; Muralidharan & Singh, 2025; Roscoe & Chi, 2007). Yet most business training programs rely on static, classroom-based

---

<sup>13</sup> These findings are not due to a lack of experimentation: training interventions have taken a wide range of forms, including classroom-based group courses (e.g., Giné & Mansuri, 2021; Karlan & Valdivia, 2011), peer interactions (e.g., Cai & Szeidl, 2018; Chatterji et al., 2019; Dimitriadis & Koning, 2022, 2025), role model mentoring (e.g., Bakhtiar et al., 2022; Brooks et al., 2018), simplified financial literacy modules (e.g., Anderson et al., 2018; Arraiz et al., 2019; Drexler et al., 2014), personalized consulting (e.g., Anderson & McKenzie, 2022; Bruhn et al., 2018), digital delivery mechanisms (Davies et al., 2024; Dell’Acqua et al., 2023), structured decision-making processes (Camuffo et al., 2019, 2024), and assessments of business viability (Lee et al., 2024).

formats that fail to deliver knowledge on demand, even though entrepreneurs' learning needs emerge in real-time.

Fourth, a critical yet underexamined limitation is the marginal role of accounting in business training. Although accounting is frequently referred to as the *language of business*, it is rarely presented as a strategic tool for decision-making in business training programs, which weakens its perceived managerial relevance and undermines its sustained adoption over time (Drexler et al., 2014; Anderson et al., 2018; and Arraiz et al., 2019). More generally, while accounting features in most business training curricula, it is often taught in isolation, siloed from other functional areas, thereby limiting its integration into broader business reasoning (Karlan et al., 2025; McKenzie et al., 2021).

This paper addresses these limitations by designing and testing a scalable, AI-based mentorship model that delivers financial acumen through personalized and real-time interaction. Recent research in developed economies and large firms shows that large language models (LLMs) can outperform human analysts in financial forecasting using standardized statements, even without detailed context (Kim et al., 2025). Similarly, Tucker, Wang, and Zhao (2024) find that AI adoption boosts firms' sales, efficiency, and profitability, though gains depend on internal capabilities. Dell'Acqua et al. (2023) further show that LLMs can significantly improve task performance and output quality in professional settings, but only when tasks lie within the AI's frontier of competence. This suggests strong potential for AI mentorship in microenterprises with limited financial expertise, while also underscoring the importance of guided use. Extended discussions and empirical foundations for these arguments are presented in Appendix A.

### III. Research Questions and Hypothesis Development

The first research question focuses on whether it is feasible to foster financial acumen among entrepreneurs, i.e., the ability to interpret business decisions through financial logic and reasoning. In doing so, I shift the traditional emphasis of business training program evaluations from business outcomes to the underlying learning mechanisms and cognitive processes through which entrepreneurs acquire financial capability. Specifically, my first research question is as follows:

***RQ 1. Can training lead entrepreneurs to perceive accounting as the language of business?***

The underlying hypothesis for this is that AI mentorship, by framing accounting as the *language of business*, can shape how entrepreneurs approach decision-making by encouraging them to adopt financial logic as a core cognitive framework in their everyday interactions. If successful, entrepreneurs will come to view accounting not merely as a record-keeping tool but as a language for reasoning through their business decisions. This shift should manifest in trade-off thinking and cost-benefit reasoning as they engage with the AI when evaluating business options. A caveat is that microentrepreneurs may rely on intuitive, rule-of-thumb decision-making under constraints; adopting a more deliberate financial logic may impose a cognitive load or conflict with entrenched habits and limited financial literacy, potentially limiting the extent of the intended cognitive shift (DellaVigna, 2009; Gigerenzer & Gaissmaier, 2011; Kahneman, 2011; Lusardi & Mitchell, 2014; Mani et al., 2013). Moreover, the effectiveness of the AI mentor may be constrained if the guidance feels disconnected from their day-to-day decisions. However, I argue that personalized framing and on-demand knowledge transfer can influence how financial acumen is acquired, contextualized, and applied across business functions, leading to the following hypothesis:

***H<sub>1</sub>: AI mentorship framed through accounting fosters a cognitive shift in how entrepreneurs approach decision-making, increasing their use of financial logic across functional areas.***

The second research question examines the effectiveness of a custom-designed, AI-based mentorship model as an alternative to traditional human-led business training for microenterprises in developing economies:

***RQ 2. Is AI-based business training more effective for microenterprises than human-led business training?***

The focus of the question is whether AI mentorship, delivered via WhatsApp using a standardized curriculum, can match or exceed human-led instruction in improving perceived business growth, business practices, and business-plan quality.<sup>14</sup> While the AI mentor is designed to address the well-known limitations of conventional training, such as limited scalability, inconsistent quality, and lack of personalization, the theoretical direction of its impact is far from certain. On one hand, the AI system offers continuous access to business knowledge, personalized feedback at scale, and the potential to provide a consistent, high-quality training. On the other hand, it lacks the human connection and trust-building mechanisms that are often critical for engagement and credibility in microenterprise settings. Prior research highlights the importance of motivation, accountability, and emotional support provided by trainers or peers that facilitate learning, especially among entrepreneurs with limited formal education or digital exposure (Campos et al., 2017; Dimitriadis & Koning, 2025; Glaub et al., 2014). Given the lack of theoretical direction, I state in null form the following hypothesis:

***H<sub>2</sub>: There is no difference in business outcomes between the AI-led training and the human-led training.***

---

<sup>14</sup> Throughout the paper, references to the “AI mentor” or “AI-based mentorship” refer to the customized, curriculum-anchored system designed in this project, which delivers personalized, interactive guidance via WhatsApp. See Section II and Appendix B for full details on design and implementation. In this context, “on-demand” refers to interactive guidance delivered by an AI mentor via WhatsApp, not static or self-paced video content.

Building on this, a corollary hypothesis posits that AI-led training, framed through accounting, will deliver superior outcomes relative to the standard AI mentor. The expectation is that presenting accounting as an integrated decision-making framework enhances entrepreneurs' ability to evaluate trade-offs, manage costs, and interpret performance across functional areas, ultimately deepening their understanding of the overall business model.

#### **IV. Experimental Design**

To test my hypotheses, I implement a randomized controlled trial (RCT) with three experimental arms. All arms draw on the ILO's Start and Improve Your Business (SIYB) curriculum, which allows the training content to be held constant while introducing two sources of experimental variation: (1) the cognitive framing of the material, i.e., whether accounting principles are embedded as the language of business across decision-making domains ( $H_1$ ), and (2) the delivery mode of the training, i.e., human-led versus AI-led mentorship ( $H_2$ ). This section outlines the institutional background, sample characteristics, randomization procedure, and the training interventions.

##### **A. Institutional Background and the AI-Based Mentor**

This study builds on *Formaliza tu Negocio*, a business training program targeting microenterprises in the informal economy in Honduras.<sup>15</sup> The program is implemented by the Chamber of Commerce and Industry of Tegucigalpa (CCIT) in collaboration with the International Labour Organization (ILO). Since 2021, CCIT has delivered standardized online training based on the globally recognized ILO's SIYB methodology. The program is delivered by CCIT coaches who are formally certified by the ILO to train entrepreneurs under the SIYB methodology. The SIYB curriculum is split into six areas of enterprise management: (1) Marketing, (2) Costing and Pricing,

---

<sup>15</sup> For more information about the program, see <https://formalizatunegocio.ccit.hn/> and ILO (2023), which features it as a case study on training entrepreneurs in the informal economy.

(3) Record Keeping, (4) Business Planning and Strategy, (5) Operations and Inventory Control, and (6) Human Resource Management.<sup>16</sup>

Building on this existing training infrastructure, and in partnership with CCIT and the ILO, I design and test a personalized, AI-based mentorship system that delivers interactive business guidance via WhatsApp, which is commonly used by microentrepreneurs in the region.<sup>17</sup> The mentor is built on GPT-4, integrated with the SIYB curriculum, and supports multimodal inputs, including text, voice messages, and images. Entrepreneurs can initiate questions and continue prior conversations with the AI mentor, which draws on a dynamic business profile constructed from baseline survey data and updated through ongoing interactions. The system customizes responses based on prior conversations, observed learning gaps, and topic preferences, enabling a highly personalized coaching experience. Unlike general-purpose models, this AI mentor is domain-specific, curriculum-anchored, and tailored to the entrepreneur's business and socioeconomic profile. It emphasizes on-demand learning and reinforcement over time.<sup>18</sup>

## **B. Participants**

Eligible participants are owners of small businesses in Honduras with fewer than 15 workers, operating in the retail or service sectors. The experimental subjects are business owners, the primary decision-makers in microenterprises, who typically oversee multiple functions due to limited access to specialized staff. To qualify, businesses must operate from fixed premises (to verify their location in Honduras), have at least six months of operational history (to ensure they

---

<sup>16</sup> In the traditional Improve Your Business training, accounting is taught through a Record Keeping module that focuses on maintaining basic books and preparing simple financial statements, treating accounting as an administrative task rather than a decision-making tool. See the SIYB manual for more information on the [record keeping](#) module. CCIT has also included a legal module on the training to support those informal firms that want to formalize their business.

<sup>17</sup> Whatsapp is one of the mobile applications with most users in Honduras, with estimations of 90% (Media Landscapes, 2017).

<sup>18</sup> For a detailed explanation of the system architecture, including the retrieval-augmented generation process, memory structure, and learning analytics, see Appendix B.

are active enterprises rather than business ideas), and use a WhatsApp-enabled smartphone, as it is the channel for AI mentorship delivery.

Importantly, participants voluntarily enrolled in the free business training program without ex-ante knowledge of the specific training modality to which they would be assigned. In fact, they were not informed in advance about the existence of the AI mentor; advertisements for the *Formaliza Tu Negocio* program simply stated that entrepreneurs would receive training to improve their businesses. The program is offered through open calls, and entrepreneurs self-selected to apply based on their perceived need for business training.<sup>19</sup> This design ensures that the sample consists of motivated business owners actively seeking to improve their managerial practices. Applicants are then screened based on the eligibility criteria outlined above (see Figure 1 for a visual overview of the experimental timeline).

### **C. Treatment and Control Groups**

Eligible applicants were then randomly assigned to one of three groups and notified of their assignment. Each group attended a separate Zoom informational session that explained the structure of their respective training modality. This was followed by the treatment period, during which participants engaged with either the AI-based mentorship or the human-led training program. To minimize spillover effects, the groups used separate communication channels for program logistics. Upon completion of the training, participants were asked to complete a post-treatment survey, and those interested could also submit a business plan for evaluation (see Appendix E for more details).

---

<sup>19</sup> Participants were recruited through the social media channels of the Chamber of Commerce and Industry of Tegucigalpa (CCIT). See CCIT's [social media](#) platforms for the recruitment post. Interested entrepreneurs registered for the program by signing a consent form and completing a baseline survey, which is available on the [CCIT platform](#).

The experiment is designed to test two hypotheses, and each requires a distinct benchmark. For H<sub>1</sub>, which examines whether framing accounting as the *language of business* affects the managerial decision process, the standard AI mentor (G1) serves as the control and is compared with the accounting-focused AI (G2). For H<sub>2</sub>, which examines the effectiveness of training modalities, both AI mentor groups (G1 and G2) are benchmarked against the traditional human-led training group (G3).

Participants in Group 1 (G1) receive access to the AI-based mentor that delivers the SIYB curriculum through interactive, on-demand guidance via WhatsApp. Entrepreneurs can engage with the AI mentor at their own pace and according to their specific business needs. The AI mentor in Group 2 (G2) uses the same digital infrastructure and curriculum as Group 1 but differs solely in how the content is framed: accounting is presented as the language through which all business decisions are discussed. Specifically, G2 mentor encourages entrepreneurs to apply accounting logic, such as cost structure, break-even analysis, contribution margins, and cash flow planning, when evaluating decisions across all functional areas. For instance, when an entrepreneur seeks help crafting a social media post to announce a new hire, the G1 mentor offers standard marketing advice, such as emphasizing the employee's role and improving service perception. In contrast, the G2 mentor embeds the same discussion within a financial framework, prompting the entrepreneur to consider the full cost of hiring, expected returns in terms of revenue or efficiency, and the implications for profit margins. This framing aims to cultivate financial reasoning alongside practical application.

Group 3 (G3) follows the standard human-based training program implemented by CCIT since 2021, using the same SIYB-based content as the AI mentors. Entrepreneurs receive 30 hours of structured instructions via Zoom conducted by certified SIYB trainers, typically over three

weeks. A one-time personalized mentorship session with a human advisor is offered at the end of the training period to reinforce key lessons and offer targeted guidance.

#### **D. Randomization Procedures and Stratification**

To ensure internal validity and maximize statistical power, I implement a stratified randomization procedure prior to treatment assignment (Bruhn & McKenzie, 2009; Cytrynbaum, 2023).<sup>20</sup> Stratification helps maintain balance across key subgroups, reducing the likelihood that differences in outcomes are driven by pre-existing differences across groups, particularly in smaller sample sizes (Bruhn, 2009; Cytrynbaum, 2023; Glennerster & Takavarasha, 2013).

Stratification is based on three baseline covariates: gender (male/female), education level (low/high), and business stage (young/established), resulting in eight mutually exclusive strata.<sup>21</sup> Prior research shows that gender norms, education levels, and business maturity influence how entrepreneurs engage with and benefit from training (Bruhn, 2009; de Mel et al., 2014; Fiala, 2018; Field et al., 2010; Karlan et al., 2015; McKenzie et al., 2021). The stratification balances treatment across key sources of heterogeneity, reduces residual variance in outcome measures, and facilitates analysis of heterogeneous treatment effects for theoretically relevant subgroups (Duflo et al., 2006).

#### **V. Empirical Design**

Random assignment ensures baseline balance across arms, so average post-treatment differences are causally attributable to the intervention (Floyd & List, 2016). I assume SUTVA (no interference across firms) and no anticipation effect, meaning that participants do not alter their

---

<sup>20</sup> A sample size calculation using GPower indicates that a total sample of 245 provides 80% power to detect a moderate effect size. To further increase statistical power, baseline covariates are pre-specified in the outcome regression. Equal group sizes also facilitate the implementation of stratified randomization and post-stratified analyses.

<sup>21</sup> I define low education as having less than a secondary school education, and high education as having completed secondary school or higher. I classify young businesses as those operating for less than three years, while established businesses are those operating for more than three years.

behavior before receiving the intervention (*Ashenfelter dip*), consistent with my field experiment design.

Before testing  $H_1$ , I conduct a validation test of the experimental design and examine whether exposure to accounting framing generates a higher demand for accounting knowledge in general. If the framing is effective, it should manifest in more accounting-related questions, follow-ups, problem identification, and behavioral responses in G2 relative to G1. To test this, I estimate the following model for G1 and G2 using a Poisson regression:

$$Accounting\_Demand_{t,i} = \alpha + \beta_2 \cdot Treatment_2 + \gamma X_i + \varepsilon_{t,i}, \quad (1)$$

where  $Accounting\_Demand_{t,i}$  is the count of the entrepreneurs' total interactions relating to accounting during the AI mentorship period  $t$ , and  $Treatment_2 = 1$  for G2 (Accounting-integrated training) and 0 for G1 (Standard AI).  $X_i$  represent a vector of pre-determined, time-invariant covariates (gender, age, education, internet access, entrepreneurial motivation, and firm age) included to improve precision (Wooldridge, 2022).

Unlike traditional business training evaluations that rely on two or more discrete survey waves, the structure of my experiment allows me to observe entrepreneurs' interactions with the AI mentor on a daily basis (Hassan et al., 2025). Therefore, I estimate Equation (1) at both the entrepreneur level ( $n = 202$ ) and as a balanced panel at the daily level ( $n = 202 \times 30 = 6,060$  entity-days), including day-of-week fixed effects in the panel specification. As a placebo test, I examine whether the treatment also affects the demand for marketing content. I jointly estimate the following equations using Poisson regression within a Seemingly Unrelated Regression (SUR) framework (Roodman, 2011):

$$Accounting\_Demand_{t,i} = \alpha + \beta_2 \cdot Treatment_2 + \gamma X_i + \varepsilon_{1t,i},$$

$$Marketing\_Demand_{t,i} = \alpha + \beta_2 \cdot Treatment_2 + \gamma X_i + \varepsilon_{2t,i},$$

$$\begin{bmatrix} \varepsilon_{1t,i} \\ \varepsilon_{2t,i} \end{bmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}\right)$$

To test H<sub>1</sub>, I examine the integration of accounting concepts into broader managerial decision-making processes. This outcome quantifies the frequency with which accounting-related terms (e.g., costs, profits, margins, cash flow) appear *in tandem* with discussions of functional areas such as marketing, customer strategy, business planning, or supplier negotiations. This approach captures not merely whether entrepreneurs learn accounting concepts, but whether they actively apply accounting logic when engaging with diverse areas of their business, i.e., an indicator of deeper cognitive integration. Formally, using the two AI mentor groups, I estimate the following equation using Poisson regression:

$$Accounting\_Integration_{t,i} = \alpha + \beta_2 \cdot Treatment_2 + \gamma X_i + \varepsilon_{t,i} \quad (5)$$

where  $Accounting\_Integration_{t,i}$  is the count of AI mentor interaction logs using tag-level data, capturing the *co-occurrence* of accounting-related and management-function tags within each entrepreneur's queries or conversational turns. As before,  $Treatment_2$  is an indicator variable equal to 1 (0) if entrepreneur  $i$  is assigned to G2 (G1). Based on H<sub>1</sub>, I predict  $\beta_2 > 0$ , which would suggest not only attention to accounting concepts but a deeper shift in how entrepreneurs reason about and manage their business.

As part of RQ1, I also examine whether using accounting as the language of business translates into actionable managerial decisions. Specifically, I assess whether exposure to the accounting-framed mentorship prompts entrepreneurs to articulate and pursue one of four growth actions: expanding sales, opening a new establishment, improving operational efficiency, or introducing employee incentives. To measure this, I analyze the conversations between entrepreneurs and the AI mentor.<sup>22</sup> Using a large-language-model (LLM) classifier, I parse each

---

<sup>22</sup> See Appendix C for examples of conversations with accounting integration and Appendix D for tag examples.

daily conversation to detect the presence of action or plan statements, categorize them into the four business-growth domains and infer both the time horizon of the intended action. I aggregate the total number of identified actions for each entrepreneur over the full training period and estimate the following cross-sectional specification:

$$Business\_Actions_{t,i} = \alpha + \beta_2 \cdot Treatment_2 + \varepsilon_{t,i}, \quad (5)$$

where  $Business\_Actions_{t,i}$  represent the total number of actions by the entrepreneur in each of the four growth actions, and  $Treatment_2$  is an indicator variable equal to 1 if entrepreneur  $i$  is assigned to G2.

When testing  $H_2$ , I compare the human-led business training (G1) with the two AI-based mentorship treatments (G2 and G3). In this setting, the analysis no longer relies on conversational data but instead uses data from two surveys conducted at pre- and post-treatment. To estimate the treatment effects, I employ a baseline-adjusted regression, where the post-treatment outcome is regressed on the treatment indicator and the corresponding baseline measure to increase power by absorbing pre-existing variation across branches (Gertler et al., 2023; McKenzie, 2012). This approach increases statistical power relative to a difference-in-differences estimator when outcomes exhibit moderate autocorrelation for randomized experiments (McKenzie, 2012; 2014).<sup>23</sup> Formally, I estimate:

$$Y_{it} = \alpha + \beta_1 Treatment_1 + \beta_2 Treatment_2 + \delta Y_{i,0} + \theta' X_i + \varepsilon_{it}, \quad (6)$$

where  $Y_{it}$  is an outcome for the unit  $i$  at time  $t \in \{1, 2\}$  measuring business performance (sales and profits), growth perception, adoption of business practices, or the quality of the business plan.

$Treatment_1$  ( $Treatment_2$ ) is an indicator variable equal to 1 if entrepreneur  $i$  is assigned to G1

---

<sup>23</sup> In untabulated analyses, I estimate the treatment effects using a linear regression adjustment and a two-period difference-in-differences (DiD) framework following Wooldridge (2022). The results were consistent with the baseline-adjusted regression estimates reported here. Given the small number of survey waves and the efficiency gains of the baseline-adjusted regression under randomized designs, I focus on this specification as the main empirical model.

(G2), with human-led training (G1) being the benchmark.  $X_i$  represent a vector of pre-determined, time-invariant covariates (gender, owner age, education, internet access, entrepreneurial motivation, firm age) included to improve precision (Wooldridge, 2022). My  $H_2$  leads to the prediction that AI-based training performs similarly to human-led training ( $\beta = 0$ ). I also examine whether AI training with accounting integration yields superior outcomes ( $\beta_2 > \beta_1$ ).

I have pre- and post-treatment survey data for outcomes related to business performance, growth perception, and adoption of business practices, facilitating the estimation of Equation (6). Because no baseline measure of business plan quality is available, and the submission of the business plan was voluntary, I account for potential sample selection in the analysis. I estimate a two-stage selection model in which the first stage models the probability of submitting a business plan. The instrument is whether the entrepreneur was referred by another Chamber of Commerce to participate in the program, an exogenous factor plausibly affecting participation but not the subsequent quality of the plan. In the second stage, I estimate the effect of treatment assignment on the evaluated quality scores of each section of the business plan, conditional on selection.<sup>24</sup> Measures of business plan quality are derived from an LLM-based evaluation of the business plans.<sup>25</sup>

## VI. Results

### Baseline Characteristics

Table 1, Panel A, presents summary statistics across the three experimental groups. The groups appear balanced along observable characteristics such as gender, age, education,

---

<sup>24</sup> I am currently exploring additional approaches to address potential selection without relying solely on the instrumental variable strategy. These include applying the Altonji et al., 2005 method to assess the robustness of the results to selection on unobservables as well as estimating a copula-based selection model (Hasebe, 2022). These extensions are part of the ongoing refinement of the empirical estimation.

<sup>25</sup> Plans were evaluated using a Large Language Model acting as an independent judge (LLMaJ), which generated section-level scores. See more details for the business plan evaluation in Appendix E.

entrepreneurial training, and digital access. While minor variation exists in firm formality and digital adoption, no systematic baseline imbalances are evident, suggesting that randomization was implemented as designed. Untabulated results from a multinomial logistic regression of treatment assignment on the three stratification variables further confirm that none significantly predict treatment assignment. The joint Wald chi-square test yields a statistic of 0.61 with a corresponding *p-value* of 0.996, indicating that treatment status is orthogonal to the stratification variables.

Overall, the sample is composed primarily of women (around two-thirds across all groups), with an average age close to 37 years. Educational attainment is limited; roughly 45% of entrepreneurs have, at most, a secondary education, and only about 40% report any prior business training. This suggests that most participants operate with relatively modest levels of formal human capital. Half of the firms have been operating for three years and at a very small scale: owners employ only 2 workers on average, and have monthly sales around \$400, with around 70% operating informally without a tax registration. While nearly two-thirds of the entrepreneurs report having internet access, only less than 12% own a computer, underscoring structural constraints in technological adoption. Taken together, these characteristics portray a population of predominantly female microentrepreneurs managing small, informal businesses with limited formal training and low levels of digital infrastructure.

### **Experimental participation across treatments**

Panel A of Table 1 also provides descriptive statistics on survey completion (*Post Survey Completion*) and participation intensity (*Participation Intensity*) across the groups. Roughly 40% of the total sample completed the exit survey of the program, with broadly similar rates across groups (43% in G1, 45% in G2, 34% in G3). A chi-square test confirms that these differences are not statistically significant at conventional levels. In contrast, *participation intensity*, measured as

the within-group percentile rank of engagement throughout the program, differs meaningfully across treatments. Entrepreneurs in the AI groups are significantly more engaged than those in the human-led group, with participation rates averaging 39.6% in G1 and 42.6% in G2, compared to 28.6% in G3. A chi-square test strongly rejects the null of equal distributions across treatments.

Table 2 formally examines predictors of survey completion (Columns 1-2) and participation intensity (Columns 3-6) using linear probability models that control for individual and firm characteristics. The results confirm that AI-based mentoring significantly boosts participation. Entrepreneurs in G1 and G2 participate at rates 9 to 13 percentage points higher than those in the human training group. Similar patterns hold when restricting the sample to non-completers of the exit survey, suggesting that higher participation rates are not solely driven by those who completed the survey (Column 5-6).

Extant research finds that participation often declines over time and that keeping entrepreneurs engaged throughout multi-session interventions is difficult, particularly for small firms facing time and resource constraints (McKenzie, 2021). Against this backdrop, the higher participation rates observed in the AI-based groups, especially in G2 with its accounting focus, suggest that the digital and personalized format may offer a promising avenue for improving engagement in entrepreneurship development programs.

### **RQ 1. Does training lead owners to perceive accounting as the language of business?**

Table 1, Panel B, reports descriptive patterns of interaction with the AI mentors for group 1 and group 2. Entrepreneurs in the AI Accounting group engaged more actively, sending on average 89.6 messages relative to 71.1 in the AI Standard group. The AI Accounting mentor elicited stronger demand for accounting-related topics, particularly profitability and cost

management, with higher incidence rates of messages requesting accounting integration (39 vs. 25).

**i. Is Demand for Accounting Information Consistent with the Intended Treatment?**

Table 3 examines the effect of the AI accounting treatment (G2) on entrepreneurs' demand for accounting information. Consistent with the intended treatment, microentrepreneurs in Group 2 engaged in roughly ten more accounting-related conversations than those in Group 1, which is economically and statistically significant. Although not the focus of my study, descriptive evidence suggests that demand for accounting information is associated with entrepreneurs' structural constraints (education and internet access) and entrepreneurial motivation. Future research could explore how these factors predict receptiveness to AI-based business training and help target interventions more effectively.

As a placebo test, if the accounting-focused AI functions as intended, I should not observe a spillover effect into marketing. I jointly estimate the accounting demand and marketing demand equations in a linear SUR framework and find that the accounting-focused AI does not engender demand for marketing content, and the slope of the accounting demand measure is significantly greater than that of marketing demand by approximately nine additional messages (Table 3). This asymmetry validates the treatment's design: entrepreneurs exposed to the accounting-focused mentor direct more attention to financial guidance, not to unrelated domains.

**ii. What do participants choose from the accounting bento box? Disaggregation of Accounting Demand**

Table 4 Panel A provides additional descriptive evidence by decomposing accounting demand into subcategories to better understand the types of financial guidance entrepreneurs sought from the AI mentors. The evidence indicates that treatment effects are concentrated in profitability-related queries, such as break-even analysis, profit margins, and cost structures, where

the estimated marginal effects are both large and statistically significant. This suggests that entrepreneurs are most eager to engage with information directly tied to assessing firm performance and sustainability. A smaller but positive effect is also observed in cash management, covering topics like budgeting and loan management, though the magnitude of this effect is more modest. By contrast, basic accounting practices (e.g., recordkeeping, separating personal and business finances) do not show significant differences between the AI groups. Daily-level estimates in Panel B Table 4 reinforce these findings.

Taken together, these results reveal that when given the choice from the *accounting bento box*, entrepreneurs disproportionately select guidance on profitability and, to a lesser extent, cash management. This pattern highlights the centrality of financial performance in small business decision-making and validates the experimental design in capturing meaningful demand for high-value accounting information.

### **iii. Integrating Accounting into Business Decisions: A Test of the *Language of Business***

Tables 5 and 6 present the results for H<sub>1</sub>, which tests whether the AI Accounting mentor led entrepreneurs to embed accounting into broader managerial discussions, consistent with accounting functioning as the *language of business*. Table 5 provides the baseline evidence. Entrepreneurs assigned to the accounting-framed AI exhibit significantly higher rates of accounting integration, both in Poisson specifications and in the joint SUR estimation of the accounting participation and integration models. When compared to the standard AI group, the estimated marginal effect indicates that entrepreneurs send about 11 additional messages that integrate accounting content with managerial decisions in various functional areas, equivalent to 28 percent of the sample mean and 78 percent of the control group mean.

Table 6 disaggregates accounting integration by topic, revealing that entrepreneurs most often combined accounting with marketing and business strategy. These findings suggest that accounting concepts are not treated as isolated technical tools but are actively linked to sales generation and planning decisions. As a placebo test, I find that accounting as a stand-alone topic is not statistically significant (Column 5, Panel A, Table 6), implying that the treatment primarily operates by embedding financial language in multi-functional decision-making contexts, rather than simply intensifying stand-alone accounting discussions. Panel B, Table 6 replicates the analysis at the daily level and reinforce these findings.

Taken together, these findings indicate that the accounting-framed AI mentor does more than stimulate demand for financial knowledge in isolation. It also fosters the integration of accounting into other domains central to entrepreneurial success (e.g., marketing and strategy), thus validating the idea that accounting can serve as the *language of business*. This integrative role is particularly salient for entrepreneurs with fewer formal skills but greater digital access, highlighting the potential of AI mentorship to democratize the use of accounting as a decision-making tool.

#### **iv. Translating Accounting into Action: Do Entrepreneurs Apply Accounting Logic to Business Growth?**

Having established that the accounting-framed AI mentorship effectively increased both demand for accounting information and its cognitive integration into managerial reasoning, I now examine whether these shifts translate into tangible business actions. Table 7 reports the estimated treatment effects on entrepreneurs' business actions identified through LLM-based classification of daily mentor-entrepreneur conversations (see Appendix E for a real example). Each column captures the number of distinct growth-oriented actions articulated by the entrepreneur, grouped

into four domains: (1) sales expansion, (2) new branch establishment, (3) operational efficiency, and (4) worker incentive schemes.

Panel A presents results for all identified actions by firms in G1 and G2. Entrepreneurs in the accounting-focused AI mentor (Group 2) display substantially higher engagement in business actions in all the related areas, with particularly strong effects on sales expansion and operational efficiency, representing increases of roughly 55% and 35% relative to the respective sample means. These estimates suggest that financial framing engenders concrete plans to improve productivity and revenue generation. Panel B restricts the sample to longer-term actions contemplated by microenterprises (those with a stated horizon of three months or more). The effects are even larger: the accounting-framed group reports nearly twice as many longer-horizon sales and operational plans compared with the baseline group. This pattern implies that accounting framing not only stimulates short-run tactical responses but also promotes forward-looking planning consistent with formal business management practices.

## **RQ 2. Is AI-based business training more effective for microenterprises than human-led business training?**

In this section, I examine the effects of assignment to the AI treatment groups (G1 and G2) relative to the human training group (G3) on three outcomes: business growth, adoption of accounting and other business practices, and business plan quality.

### **i. Business Growth**

Table 8 presents the effects of the mentorship modalities on entrepreneurs' perceptions of business growth and on realized financial outcomes (sales/profits). In the full sample (Panel A), entrepreneurs in the accounting-focused AI group (G2) report a 16 percentage points higher perception of business growth relative to the human-led control group, indicating an early boost in confidence and perceived performance. However, monthly sales and profits show no statistically

significant improvements, reflecting the well-documented volatility and measurement noise of short-term financial indicators among microenterprises. In Panel B, which focuses on entrepreneurs starting with below-average business practices (Otis et al., 2024), the treatment effects are larger in magnitude: the AI Accounting group reports 40 percentage-point higher perceived growth. Financial outcomes remain noisy in this subgroup as well, with no statistically reliable changes in average sales or profits. Taken together, these results suggest that entrepreneurs quickly internalize the financial logic introduced by the AI mentor, which translates into stronger perceptions of business growth. However, the short duration of the intervention and the inherent variability of small-business revenues limit the detectability of measurable financial gains (McKenzie, 2021).

## **ii. Accounting and Other Business Practices**

Untabulated results show that all training groups, regardless of whether the training was delivered by AI or by humans, improved their overall business practices score by roughly 10 points. As expected, the main statistically significant differences across groups occur in accounting practices. Table 9 presents the treatment effects of the different mentorship modalities on entrepreneurs' adoption of business practices. Outcomes are grouped into two dimensions: *Accounting Practices* (Columns 1–3) and *Other Business Practices* (Column 4). Panel A shows results for the full sample, while Panel B restricts the analysis to entrepreneurs who began the program with below-average business practices. In the full sample, the AI Accounting mentor yields modest but positive effects across accounting-related outcomes. Entrepreneurs exposed to the AI Accounting treatment exhibit a 0.28 point improvement in cost-analysis practices, representing roughly a 9 percent improvement relative to the baseline mean.<sup>26</sup> Panel B highlights

---

<sup>26</sup> The economic significance is calculated as the treatment coefficient (0.275) divided by the human-led group's pre-treatment mean (2.97), indicating roughly a 9 percent improvement on the 1–4 cost-analysis scale.

that effects are much stronger among entrepreneurs who started the program with weaker business practices. The AI Accounting mentor substantially outperforms both the human-led and standard AI mentors across several dimensions: cost analysis improves by 77%, and the likelihood of maintaining systematic accounting records rises by 26%. The effects are not mechanical, as entrepreneurs in both AI groups were exposed to accounting content; however, only those in Group 2 learned accounting through its framing as the *language of business*. The lack of detectable changes in other business practices (Column 4) suggests that, for the average entrepreneur, the intervention's effects are concentrated in the accounting domain rather than in general management areas.

### **iii. Business Plans Evaluation<sup>27</sup>**

Table 10 explores whether the intervention affected entrepreneurs' ability to prepare higher-quality business plans, using detailed scores across nine dimensions.<sup>28</sup> As discussed in the empirical design section, to account for potential selection bias, since not all participants submitted a business plan, the analysis follows a two-stage selection framework. The results show that the AI Mentorship groups significantly improve the business plan quality in most of the dimensions. Entrepreneurs in G1 (AI Standard) scored nearly 11 points higher on average compared to the human-led training group, representing almost a 10% improvement relative to the 120 point maximum. Gains are particularly concentrated in marketing (increase of 4.86 points), human resources (3.30 points), cost (1.23 points), profit (0.32 points), and legal sections (1.97 points).

---

<sup>27</sup> This is a preliminary section. As mentioned in footnote 24, I am currently exploring additional approaches to address potential selection without relying solely on the instrumental-variable strategy, such as the methods proposed by Altonji et al. (2005) and Hasebe (2022).

<sup>28</sup> The business plan includes nine core dimensions: (i) Executive Summary (ii) Business Idea: product or service definition, market opportunity, client need, (iii) Market Research: understanding of customers, competitors, and market gaps, (iv) Commercialization Plan: product quality, pricing, distribution, promotion, personnel management, business image, (v) Sales Estimate: realism of projected revenues, (vi) Organizational Structure: clarity of roles and responsibilities, (vii) Formalization and Legal Aspects: regulatory compliance and fiscal awareness, (viii) Costing and Financial Planning: accuracy of cost, profit, and cash flow projections, and (ix) Initial Capital and Financing: justification of funding needs and sources.

These effects are statistically significant at the 1% level, suggesting that the AI mentor helps entrepreneurs articulate a more comprehensive view of their business. G2 (AI Accounting) also performs better than G3, though with smaller magnitudes in the overall evaluation but with higher impact of projected sales, suggesting that the accounting-focused treatment may encourage them to more carefully plan for revenue generation and pricing strategies.

Although it falls outside the main scope of my analysis, internet access emerges as another powerful predictor for the quality of the business plans. Entrepreneurs with internet access score almost 10 points higher overall.<sup>29</sup> The number of employees is also positively associated with higher scores in cost, profit, legal, and HR sections, which is intuitive given that larger firms likely face more complex management needs and have more to gain from formal planning.

#### **iv. Business Perspectives**

Table 11 reports treatment effects on entrepreneurs' preferences for various attributes of the training modalities. Column (1) shows that both AI groups perceived significantly greater support in addressing their business questions compared with the human-led training. This suggests that even a general AI mentor, without an accounting framing, can provide entrepreneurs a consistent and accessible source of advice throughout the program.

Interestingly, column (2) shows both AI groups experiencing a significantly lower perceived learning content when compared to the human training group. This suggests that while entrepreneurs valued having on-demand support, they may have found the AI's advice less structured or struggled to recognize that they were engaging in a learning process through their own context-specific questions. Column (3) captures flexibility, referring to convenience and

---

<sup>29</sup> While all training modalities in the experiment required internet connectivity, the relative cost of AI-based information retrieval is considerably lower than that of general web searches. For example, Björkegren et al. (2025) show that in Sierra Leone, the average web search consumes over 3,000 times more data than an AI-generated response, making AI approximately 87 percent less expensive per query.

scheduling format. Interestingly, only G3 entrepreneurs reported a statistically significant positive effect, even though both AI treatments offered the same 24/7 availability. One possible explanation is that G3 participants engaged more actively with the mentor, as suggested by their higher participation rates in Table 2, making the always-available format feel more valuable and better aligned with their workflow.

Finally, column (4) examines networking, which captures the value entrepreneurs place on peer interactions facilitated by the program. Consistent with the program design, where participants were not brought together in any live or group training sessions, both AI treatments show statistically significant negative effects, with G2 and G1 reducing perceived networking opportunities. These findings highlight a potential trade-off, while the AI mentor provides highly personalized, on-demand support, it may inadvertently crowd out opportunities or motivation for peer-to-peer learning that traditionally occurs in group-based programs. This is an important area for future research, as hybrid models that combine 24/7 AI availability with periodic peer interaction and structured sessions with human coaches may help capture the benefits of both individualized guidance and collaborative learning.

More importantly, program satisfaction metrics reveal meaningful differences. Entrepreneurs in the human-led training group (G3) reported the highest satisfaction, with 94.4% rating the program as “very useful” and 97.2% stating they would “definitely recommend” it to other business owners. Strikingly, the AI Accounting group (G2) came close to these levels, with 85.7% rating the program as “very useful” and 95.9% willing to recommend it. By contrast, the general AI group (G1) lagged behind, with only 56.5% rating the program as “very useful” and 76.1% willing to recommend it. These differences are statistically significant, highlighting that the accounting-focused AI not only improved business practices but also enhanced perceived value,

nearly matching the gold standard of human-led training, while outperforming the general AI model. Perhaps the AI accounting model encouraged entrepreneurs to rethink and analyze their decisions more carefully, promoting a more sustainable path to growth. These results, however, are based only on participants who completed the exit survey, and attrition may limit their representativeness.<sup>30</sup>

## VII. Cost Effectiveness

A key aspect for evaluating training interventions is not only impact, but also scalability and cost-effectiveness (Al-Ubaydli et al., 2017; Dinarte-Díaz et al., 2025; List, 2023; Muralidharan & Niehaus, 2017). Traditional entrepreneurship programs rely on scarce, skilled trainers and small cohorts; quality scales slowly along an upward-sloping supply of high-quality coaches. By contrast, the AI mentorship model delivers guidance via WhatsApp at negligible marginal cost per additional entrepreneur, maintaining consistent content quality at scale.

Following prior work (e.g., Bruhn et al., 2018; McKenzie & Woodruff, 2014), I benchmark costs using budget data from the Chamber of Commerce and Industry of Tegucigalpa (CCIT). According to ILO (2019), human-led SIYB training at CCIT costs approximately US\$1,600 per participant (trainer fees and administration). In contrast, the AI mentorship costs about US\$40 per participant, driven primarily by OpenAI token charges; server costs are excluded, but the marginal cost of expansion is effectively zero. Thus, the AI model attains comparable or larger improvements in business practices at less than 3% of the per-participant cost of the human-led program. Even under more conservative benchmarks from Van Lieshout & Mehta (2017), who report SIYB costs spanning US\$3,537 to US\$177 across 18 countries, the AI mentor remains

---

<sup>30</sup> In exploratory analyses, I examine participants' willingness to pay (WTP) for the program at baseline and at the endline survey. The share willing to pay more than \$102 USD increased from about 9% at baseline to roughly 22% at exit, with the accounting-framed AI group showing the largest upward shift. These results share the same limitation of being based only on participants who completed the exit survey, and attrition may limit their representativeness.

substantially cheaper. At the lower bound of US\$177, the AI mentorship is still around 4 times less expensive. To compare cost-effectiveness directly, I scale intent-to-treat effects by expenditures and report cost per 0.1 SD improvement in core accounting practices (cost tracking, profit monitoring, recordkeeping):

$$\text{Cost per 0.1 SD} = C \times 0.1 / |\hat{\beta}_{\text{SD}}|,$$

where  $C$  is per-participant cost and  $\hat{\beta}_{\text{SD}}$  the effect in SD units. Using Panel B, Table 8 estimates (below-mean baseline) and  $C = \$40$  for AI Accounting (G2), the implied costs are approximately US\$5.22 per 0.1 SD for cost analysis, US\$14.70 for profit calculation, and US\$15.21 for systematic recordkeeping. I also provide a simple payback calculation for context. Let  $C$  be program cost and  $\Delta\pi$  the monthly profit change; the payback horizon is:

$$\text{Payback months} = C / \Delta\pi$$

Using McKenzie (2020) as an illustrative scenario (not evidence for my AI setting), a firm with baseline monthly profits  $\pi_0 = \$100$  and a 10% lift implies  $\Delta\pi = \$10$ . A human-led program with  $C = \$177$  therefore pays back in around 18 months; the AI mentor with  $C = \$40$  pays back in around 4 months. I do not estimate profit effects for the AI mentor in this study as the post-training window is short and noisy, so these payback figures are illustrative.

## VIII. Conclusion

This paper introduces and evaluates a scalable, AI-based mentorship model designed to build financial acumen among microentrepreneurs. Three conclusions emerge. First, the treatment-validation tests show that entrepreneurs assigned to the accounting-framed AI mentor selectively increase their demand for accounting content (especially profitability and cash-flow topics) without commensurate increases in unrelated domains such as marketing. This confirms that the framing operates as intended. Second, mechanism tests indicate that the accounting-framed mentor

does more than stimulate topic-specific demand: it measurably increases the integration of financial logic into broader managerial conversations, most notably marketing and strategy, consistent with accounting functioning as a cross-functional decision language rather than a stand-alone compliance task. Third, relative to human-led training, the AI models delivers comparable or larger improvements in business practices and perceived growth, with the largest gains among owners starting from weaker managerial baselines. Additionally, the AI mentor represent a feasible solution as costs are an order of magnitude lower and the entrepreneurs value in similar magnitudes the experience with the AI mentor than with the human-led training. At roughly \$40 per participant vs. \$1,600 for classroom training, the AI model achieves substantial standardized improvements at less than 3% of per-participant cost.

My study also contributes methodologically. Embedding an RCT within an AI system generates high-frequency interaction data that make otherwise latent learning dynamics observable (e.g., what owners ask, when they ask, how they incorporate feedback, and how reasoning travels across domains). This design opens new possibilities for testing mechanisms (content, framing, behavioral channels) that are difficult to evaluate with traditional survey waves alone.

There are several limitations in the results. External validity is bounded by the setting (Honduras) and the digital delivery channel (WhatsApp). Interaction-based measures, while rich, are subject to variation in engagement intensity. Future work should test hybrid delivery models, (i.e., periodic, human-facilitated peer sessions layered onto 24/7 AI mentorship) as they can capture the strengths of individualized guidance while preserving networking and extrinsic motivation; and explore longer-run outcomes (profits, survival, formalization).

## References:

- Aguilar, M., Lind, G., & Ramesh, K. (2025). *The Role of Accounting in the Informal Economy* (SSRN Scholarly Paper No. 5199563). Social Science Research Network. <https://doi.org/10.2139/ssrn.5199563>
- Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113(1), 151–184. <https://doi.org/10.1086/426036>
- Al-Ubaydli, O., List, J. A., & Suskind, D. L. (2017). What Can We Learn from Experiments? Understanding the Threats to the Scalability of Experimental Results. *American Economic Review*, 107(5), 282–286. <https://doi.org/10.1257/aer.p20171115>
- Anderson, S. J., Chandy, R., & Zia, B. (2018). *Pathways to Profits: The Impact of Marketing vs. Finance Skills on Business Performance* The Impact of Marketing vs. Finance Skills on Business Performance on JSTOR. <https://www-jstor-org.ezproxy.rice.edu/stable/48748367?sid=primo>
- Anderson, S. J., Lazicky, C., & Zia, B. (2019). *Measuring the Unmeasured: Combining Technology and Behavioral Insights to Improve Measurement of Business Outcomes* (SSRN Scholarly Paper No. 3380729). Social Science Research Network. <https://papers.ssrn.com/abstract=3380729>
- Arraiz, I., Bhanot, S. P., & Calero, C. (2019). Less is More: Experimental Evidence on Heuristic-Based Business Training in Ecuador. *IDB Invest Working Paper*.
- Bernard, D., & Sutherland, A. (2025). *Private Firms and the Economic Role of Accounting: A Review of Empirical Research* (SSRN Scholarly Paper No. 5197461). Social Science Research Network. <https://doi.org/10.2139/ssrn.5197461>
- Björkegren, D., Choi, J. H., Budihal, D., Sobhani, D., Garrod, O., & Atherton, P. (2025). *Could AI Leapfrog the Web? Evidence from Teachers in Sierra Leone* (No. arXiv:2502.12397). arXiv. <https://doi.org/10.48550/arXiv.2502.12397>
- Bruhn, M. (2009). *Female-Owned Firms in Latin America: Characteristics, Performance, and Obstacles to Growth* (SSRN Scholarly Paper No. 1503816). <https://papers.ssrn.com/abstract=1503816>
- Bruhn, M., & McKenzie, D. (2009). In Pursuit of Balance: Randomization in Practice in Development Field Experiments. *American Economic Journal: Applied Economics*, 1(4), 200–232. <https://doi.org/10.1257/app.1.4.200>
- Buffett. (1987). *Chairman's Letter—1986*. <https://www.berkshirehathaway.com/letters/1986.html>
- Campos, F., Frese, M., Goldstein, M., Iacovone, L., Johnson, H. C., McKenzie, D., & Mensmann, M. (2017). Teaching personal initiative beats traditional training in boosting small business in West Africa. *Science*, 357(6357), 1287–1290. <https://doi.org/10.1126/science.aan5329>
- Camuffo, A., Cordova, A., Gambardella, A., & Spina, C. (2019). A Scientific Approach to Entrepreneurial Decision Making: Evidence from a Randomized Control Trial. *Management Science*, 66(2), 564–586. <https://doi.org/10.1287/mnsc.2018.3249>
- Chang, S., Kennedy, A., Leonard, A., & List, J. A. (2024). *12 Best Practices for Leveraging Generative AI in Experimental Research* (Working Paper No. 33025). National Bureau of Economic Research. <https://doi.org/10.3386/w33025>
- Charness, G., Jabarian, B., & List, J. A. (2023). *Generation Next: Experimentation with AI* (SSRN Scholarly Paper No. 4576689). Social Science Research Network. <https://doi.org/10.2139/ssrn.4576689>
- Cohen, S., & Koning, R. (2024). *<p>Advice And The Bayesian Entrepreneur</p>* (SSRN Scholarly Paper No. 5030623). Social Science Research Network. <https://doi.org/10.2139/ssrn.5030623>

- Csaszar, F. A., Ketkar, H., & Kim, H. (2024). *Artificial Intelligence and Strategic Decision-Making: Evidence from Entrepreneurs and Investors* (SSRN Scholarly Paper No. 4913363). Social Science Research Network. <https://doi.org/10.2139/ssrn.4913363>
- Cytrynbaum, M. (2023). *Optimal Stratification of Survey Experiments* (No. arXiv:2111.08157). arXiv. <https://doi.org/10.48550/arXiv.2111.08157>
- de Barros, O., Cabral, S., Lazzarini, S. G., & Nardi, L. (2024). *LOST IN AUTOMATION: THE IMPACT OF TECHNOLOGY-SUPPORTED TRAINING ON THE PERFORMANCE OF SMALL ENTREPRENEURS IN UNDERSERVED REGIONS* (SSRN Scholarly Paper No. 4803222). Social Science Research Network. <https://doi.org/10.2139/ssrn.4803222>
- de Mel, S., McKenzie, D. J., & Woodruff, C. (2009). Measuring microenterprise profits: Must we ask how the sausage is made? *Journal of Development Economics*, 88(1), 19–31. <https://doi.org/10.1016/j.jdeveco.2008.01.007>
- de Mel, S., McKenzie, D., & Woodruff, C. (2014). Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from Sri Lanka. *Journal of Development Economics*, 106, 199–210. <https://doi.org/10.1016/j.jdeveco.2013.09.005>
- Dell’Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). *Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality* (SSRN Scholarly Paper No. 4573321). Social Science Research Network. <https://doi.org/10.2139/ssrn.4573321>
- DellaVigna, S. (2009). Psychology and Economics: Evidence from the Field. *Journal of Economic Literature*, 47(2), 315–372. <https://doi.org/10.1257/jel.47.2.315>
- Dimitriadis, S., & Koning, R. (2022). Social Skills Improve Business Performance: Evidence from a Randomized Control Trial with Entrepreneurs in Togo. *Management Science*, 68(12), 8635–8657. <https://doi.org/10.1287/mnsc.2022.4334>
- Dimitriadis, S., & Koning, R. (2025). Networking Frictions and Entrepreneurial Learning in Developing Economies. *Management Science*, 71(4), 3127–3145.
- Drexler, A., Fischer, G., & Schoar, A. (2014). Keeping It Simple: Financial Literacy and Rules of Thumb. *American Economic Journal. Applied Economics*, 6(2), 1–31. <https://doi.org/10.1257/app.6.2.1>
- Duflo, E., Glennerster, R., & Kremer, M. (2006). *Using Randomization in Development Economics Research: A Toolkit* (No. t0333). National Bureau of Economic Research. <https://doi.org/10.3386/t0333>
- Fafchamps, M., McKenzie, D., Quinn, S., & Woodruff, C. (2012). Using PDA consistency checks to increase the precision of profits and sales measurement in panels. *Journal of Development Economics*, 98(1), 51–57. <https://doi.org/10.1016/j.jdeveco.2010.06.004>
- Fiala, N. (2018). Returns to microcredit, cash grants and training for male and female microentrepreneurs in Uganda. *World Development*, 105, 189–200. <https://doi.org/10.1016/j.worlddev.2017.12.027>
- Field, E., Jayachandran, S., & Pande, R. (2010). Do Traditional Institutions Constrain Female Entrepreneurship? A Field Experiment on Business Training in India. *American Economic Review*, 100(2), 125–129. <https://doi.org/10.1257/aer.100.2.125>
- Floyd, E., & List, J. A. (2016). Using Field Experiments in Accounting and Finance. *Journal of Accounting Research*, 54(2), 437–475. <https://doi.org/10.1111/1475-679X.12113>
- Gertler, P., Higgins, S., Scott, A., & Seira, E. (2023). *Using Lotteries to Attract Deposits* (Working Paper No. 31529). National Bureau of Economic Research. <https://doi.org/10.3386/w31529>

- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, 62(Volume 62, 2011), 451–482. <https://doi.org/10.1146/annurev-psych-120709-145346>
- Giuggioli, G., & Pellegrini, M. M. (2022). Artificial intelligence as an enabler for entrepreneurs: A systematic literature review and an agenda for future research. *International Journal of Entrepreneurial Behavior & Research*, 29(4), 816–837. <https://doi.org/10.1108/IJEBr-05-2021-0426>
- Glaub, M. E., Frese, M., Fischer, S., & Hoppe, M. (2014). Increasing Personal Initiative in Small Business Managers or Owners Leads to Entrepreneurial Success: A Theory-Based Controlled Randomized Field Intervention for Evidence-Based Management. *Academy of Management Learning & Education*, 13(3), 354–379.
- Glennerster, R., & Takavarasha, K. (2013). *Running Randomized Evaluations: A Practical Guide*. Princeton University Press. <https://doi.org/10.1515/9781400848447>
- Hasebe. (2022). *Endogenous models of binary choice outcomes: Copula-based maximum-likelihood estimation and treatment effects—Takuya Hasebe, 2022*. [https://journals.sagepub.com/doi/full/10.1177/1536867X221140943?casa\\_token=L4wG4Y5tBYcAAAAA%3Aw92xC3rGENnz5xmrGGazD7KI9ut9h8zqFiv0qterd9yN2QPCE5z000dOAA-egcXTQ8MjjehnqOZCwho](https://journals.sagepub.com/doi/full/10.1177/1536867X221140943?casa_token=L4wG4Y5tBYcAAAAA%3Aw92xC3rGENnz5xmrGGazD7KI9ut9h8zqFiv0qterd9yN2QPCE5z000dOAA-egcXTQ8MjjehnqOZCwho)
- Hassan, T. A., Hollander, S., Kalyani, A., van Lent, L., Schwedeler, M., & Tahoun, A. (2025). Text as Data in Economic Analysis. *Journal of Economic Perspectives*, 39(3), 193–220. <https://doi.org/10.1257/jep.20231365>
- ILO. (2019, October 10). *Small Matters: Global evidence on the contribution to employment by the self-employed, micro-enterprises and SMEs | International Labour Organization*. <https://www.ilo.org/publications/small-matters-global-evidence-contribution-employment-self-employed-micro>
- ILO. (2023). *Formaliza tu negocio: Caso de estudio: Honduras*.
- Kahneman, D. (2011). *Thinking, fast and slow* (p. 499). Farrar, Straus and Giroux.
- Karlan, D., de Bundel, F., Pariente, W., & Udry, C. (2025). *Entrepreneurship Meta Analysis*.
- Karlan, D., Knight, R., & Udry, C. (2015). Consulting and capital experiments with microenterprise tailors in Ghana. *Journal of Economic Behavior & Organization*, 118, 281–302. <https://doi.org/10.1016/j.jebo.2015.04.005>
- Kim, A., Muhn, M., & Nikolaev, V. (2025). *Financial Statement Analysis with Large Language Models* (No. arXiv:2407.17866). arXiv. <https://doi.org/10.48550/arXiv.2407.17866>
- List, J. A. (2023). The Voltage Effect. *Revue Économique*, 74(5), 673–680.
- Lusardi, A., & Mitchell, O. S. (2014). The Economic Importance of Financial Literacy: Theory and Evidence. *Journal of Economic Literature*, 52(1), 5–44. <https://doi.org/10.1257/jel.52.1.5>
- Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, 341(6149), 976–980. <https://doi.org/10.1126/science.1238041>
- McConnell, J. J., & Muscarella, C. J. (1985). Corporate capital expenditure decisions and the market value of the firm. *Journal of Financial Economics*, 14(3), 399–422. [https://doi.org/10.1016/0304-405X\(85\)90006-6](https://doi.org/10.1016/0304-405X(85)90006-6)
- McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, 99(2), 210–221. <https://doi.org/10.1016/j.jdeveco.2012.01.002>

- McKenzie, D. (2021). Small business training to improve management practices in developing countries: Re-assessing the evidence for ‘training doesn’t work.’ *Oxford Review of Economic Policy*, 37(2), 276–301. <https://doi.org/10.1093/oxrep/grab002>
- McKenzie, D., & Woodruff, C. (2014). What Are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World? *The World Bank Research Observer*, 29(1), 48–82. <https://doi.org/10.1093/wbro/lkt007>
- McKenzie, D., Woodruff, C., Bruhn, M., Cai, J., Gonzalez Uribe, J., Quinn, S., Sonobe, T., & Valdivia, M. (2021). Training Entrepreneurs: Issue 2. *VoxDevLits*.
- Media Landscapes. (2017). Media Landscapes. <https://medialandscapes.org/country/honduras/media/social-networks>
- Mehmood, Z. (2024). *Essays on Development and Political Economy—ProQuest*. <https://www.proquest.com/docview/3111156031?pq-origsite=gscholar&fromopenview=true>
- Mollick, E. R., & Mollick, L. (2022). *New Modes of Learning Enabled by AI Chatbots: Three Methods and Assignments* (SSRN Scholarly Paper No. 4300783). Social Science Research Network. <https://doi.org/10.2139/ssrn.4300783>
- Mollick, E. R., & Mollick, L. (2023). *Using AI to Implement Effective Teaching Strategies in Classrooms: Five Strategies, Including Prompts* (SSRN Scholarly Paper No. 4391243). Social Science Research Network. <https://doi.org/10.2139/ssrn.4391243>
- Muralidharan, K., & Singh, A. (2025). *Adapting for scale: Experimental Evidence on Technology-aided Instruction in India* (SSRN Scholarly Paper No. 5455845). Social Science Research Network. <https://papers.ssrn.com/abstract=5455845>
- Myers, S. C. (1977). Determinants of corporate borrowing. *Journal of Financial Economics*, 5(2), 147–175. [https://doi.org/10.1016/0304-405X\(77\)90015-0](https://doi.org/10.1016/0304-405X(77)90015-0)
- Otis, N., Clarke, R., Delecourt, S., Holtz, D., & Koning, R. (2024). *The Uneven Impact of Generative AI on Entrepreneurial Performance* (SSRN Scholarly Paper No. 4671369). <https://doi.org/10.2139/ssrn.4671369>
- Ramdas, K., & Sungu, A. (2022). *The Digital Lives of the Poor: Entertainment Traps and Information Isolation* (SSRN Scholarly Paper No. 4012172). Social Science Research Network. <https://doi.org/10.2139/ssrn.4012172>
- Ramesh, K. (2025). *Chapter 2*.
- Roodman, D. (2011). Fitting Fully Observed Recursive Mixed-process Models with cmp. *The Stata Journal*, 11(2), 159–206. <https://doi.org/10.1177/1536867X1101100202>
- Roscoe, R. D., & Chi, M. T. H. (2007). Understanding Tutor Learning: Knowledge-Building and Knowledge-Telling in Peer Tutors’ Explanations and Questions. *Review of Educational Research*, 77(4), 534–574. <https://doi.org/10.3102/0034654307309920>
- Small Firm Diaries. (2024). *Global | Small Firm Diaries Global Publications*. Small Firm Diaries. <https://www.smallfirmdiaries.org/global>
- Tomy, R. E., & Wittenberg-Moerman, R. (2024). Community membership and reciprocity in lending: Evidence from informal markets. *Journal of Accounting and Economics*, 101697. <https://doi.org/10.1016/j.jacceco.2024.101697>
- Tucker, J. W., Wang, J., & Zhao, R. (2024). *Corporate use of artificial intelligence and subsequent product market performance*.
- Van Lieshout, S., & Mehtha, P. (2017, June 26). *THE NEXT 15 MILLION Start and Improve Your Business Global Tracer Study 2011-15 | International Labour Organization*.

<https://www.ilo.org/publications/next-15-million-start-and-improve-your-business-global-tracer-study-2011-15>  
World Bank. (2021). [Text/HTML]. World Bank. <https://www.worldbank.org/en/news/press-release/2021/05/11/widespread-informality-likely-to-slow-recovery-from-covid-19-in-developing-economies>

Table 1: Summary Statistics

Panel A: Baseline Characteristics by Treatment Group						
Variable	G1. AI Standard		G2. AI Accounting		G3. Human	
	Obs	Mean	Obs	Mean	Obs	Mean
<i>Female</i>	104	0.663	106	0.689	101	0.683
<i>Low Education</i>	104	0.471	106	0.453	101	0.426
<i>Age</i>	98	36.908	104	37.731	100	37.120
<i>Entrepreneurial Motivation</i>	104	0.096	106	-0.034	101	-0.058
<i>Internet Access</i>	104	0.644	106	0.689	101	0.634
<i>Young Firm: Less than 3 years</i>	104	0.538	106	0.509	101	0.535
<i>Number of Employees</i>	104	1.788	106	1.708	101	2.129
<i>Monthly Sales (\$)</i>	104	428.113	106	387.517	101	438.933
<i>Formal Firm</i>	104	0.327	106	0.245	101	0.376
<i>Post Survey Completion</i>	104	0.438	106	0.453	101	0.342
<i>Participation Intensity</i>	104	0.396	106	0.426	101	0.286

Panel B: Summary Statistics by AI Groups from Conversations						
Variable	G1. AI Standard			G2. AI Accounting		
	Count	Sum	Mean	Count	Sum	Mean
<i>Messages sent by Entrepreneur</i>	104	7,394	71.096	106	9,500	89.623
<i>Accounting Demand</i>	104	1,744	16.769	106	2,671	25.198
<i>Profitability</i>	104	602	5.788	106	1,146	10.811
<i>Cost &amp; Pricing</i>	104	370	3.558	106	598	5.642
<i>Inventory Management</i>	104	96	0.923	106	142	1.340
<i>Accounting Practices</i>	104	540	5.192	106	530	5.000
<i>Financial Management</i>	104	134	1.288	106	255	2.406
<i>Accounting Integration</i>	104	1,499	14.413	106	2,620	24.717
<i>Marketing</i>	104	1,359	13.067	106	1,545	14.575

Panel A reports demographic and firm characteristics collected from the baseline survey administered prior to treatment assignment. Panel B summarizes conversational data extracted from AI mentor-entrepreneur interactions over the intervention period.

Table 2: Participation Intensity and Post Survey Completion

	Survey Completion		Unconditional on Survey Completion		Conditional on Survey Non-Completion	
			Participation Intensity	Participation Intensity	Participation Intensity	Participation Intensity
	(1)	(2)	(3)	(4)	(5)	(6)
<i>G1. AI Standard</i>	0.223 (1.25)	0.220 (1.19)	0.109** (2.20)	0.090* (1.80)	0.174*** (4.23)	0.167*** (3.98)
<i>G2. AI Accounting</i>	0.273 (1.54)	0.273 (1.49)	0.140*** (2.86)	0.125** (2.56)	0.181*** (4.24)	0.186*** (4.34)
<i>Age</i>		0.712** (2.46)		0.132* (1.70)		0.007 (0.10)
<i>Female</i>		0.127 (0.78)		0.012 (0.28)		0.031 (0.87)
<i>Low Education</i>		0.004 (0.03)		0.021 (0.54)		0.038 (1.05)
<i>Internet Access</i>		0.248 (1.54)		0.085** (2.14)		0.029 (0.82)
<i>Entrepreneurial Motivation</i>		0.056 (1.32)		0.020* (1.84)		0.019** (2.06)
<i>Number of Employees (Excluding the Owner)</i>		0.015 (0.42)		-0.008 (-0.89)		-0.002 (-0.37)
<i>Firm Age</i>		0.059 (0.61)		0.026 (1.01)		0.013 (0.55)
Lincom Test: G2 – G1	0.050 (0.29)	0.053 (0.30)	0.031 (0.68)	0.036 (0.80)	0.007 (0.15)	0.020 (0.39)
Observations	311	302	311	302	180	177

Survey completion is defined as whether an entrepreneur completed both the business training program and the exit survey. Participation intensity reflects the engagement during the program, measured as the within-group percentile rank of participation. Columns (1)–(2) report probit estimates for survey completion; columns (3)–(6) report OLS estimates for participation. Columns (5)–(6) restrict the sample to non-completers of the exit survey. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table 3: Do the Entrepreneurs Receive the Treatment as Intended? A Treatment Manipulation Check

	Accounting Demand		Seemingly Unrelated Regressions	
	(1)	(2)	Accounting	Marketing
<i>Treatment</i>	8.429*	10.265**	10.192**	1.570
	(1.67)	(2.17)	(2.16)	(0.45)
<i>Female</i>		5.438	5.829	8.778***
		(0.99)	(1.24)	(2.99)
<i>Age</i>		0.169	0.219	0.258
		(0.82)	(1.00)	(1.35)
<i>Low Education</i>		16.948***	16.757***	6.125
		(2.97)	(3.12)	(1.63)
<i>Entrepreneurial Motivation</i>		2.772**	2.887**	1.379
		(2.19)	(2.31)	(1.58)
<i>Internet Access</i>		11.113**	10.661**	7.331**
		(2.00)	(2.34)	(2.10)
Lincom Test: Treatment (Acc – Mkt)				8.623*
				(1.82)
Observations	210	202	202	202

The treatment group includes entrepreneurs in G2 (AI Accounting), while the benchmark control group corresponds to G1 (AI Standard). Columns (1) and (2) report marginal effects from Poisson models, and Columns (4) and (5) present linear SUR estimates. Estimates are obtained from a joint CMP system for accounting and marketing information demand at the mentee level, with robust standard errors. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table 4: What Did Participants Choose from the Accounting Bento Box? A Validation Analysis

Panel A: Entrepreneurs' Level				
	Accounting Demand	Profitability	Accounting Practices	Cash Management
	(1)	(2)	(3)	(4)
<i>Treatment</i>	10.265** (2.17)	8.134** (2.44)	0.455 (0.27)	1.399 (1.60)
<i>Female</i>	5.438 (0.99)	4.655 (1.27)	1.505 (0.93)	-1.114 (-0.99)
<i>Age</i>	0.169 (0.82)	0.058 (0.35)	0.104 (1.38)	-0.027 (-0.80)
<i>Low Education</i>	16.948*** (2.97)	9.130** (2.45)	4.255** (2.04)	1.730 (1.41)
<i>Entrepreneurial Motivation</i>	2.772** (2.19)	2.357** (2.41)	0.196 (0.52)	0.379* (1.91)
<i>Internet Access</i>	11.113** (2.00)	8.135** (2.00)	1.684 (1.02)	0.422 (0.39)
Observations	202	202	202	202
Panel B: Daily Level				
	Accounting Demand	Profitability	Accounting Practices	Cash Management
	(1)	(2)	(3)	(4)
<i>Treatment</i>	0.342** (2.17)	0.271** (2.44)	0.015 (0.27)	0.047 (1.60)
<i>Female</i>	0.181 (0.99)	0.155 (1.27)	0.050 (0.93)	-0.037 (-0.99)
<i>Age</i>	0.006 (0.82)	0.002 (0.35)	0.003 (1.38)	-0.001 (-0.80)
<i>Low Education</i>	0.565*** (2.97)	0.304** (2.45)	0.142** (2.04)	0.058 (1.41)
<i>Entrepreneurial Motivation</i>	0.092** (2.19)	0.079** (2.41)	0.007 (0.52)	0.013* (1.91)
<i>Internet Access</i>	0.370** (2.00)	0.271** (2.00)	0.056 (1.02)	0.014 (0.39)
Observations	6,060	6,060	6,060	6,060
Day of Week Fixed Effects	Yes	Yes	Yes	Yes

The treatment group includes entrepreneurs in G2 (AI Accounting), while the benchmark control group corresponds to G1 (AI Standard). All columns report marginal effects from Poisson regressions at the entrepreneur level with robust standard errors. Accounting demand (Column 1) is further disaggregated into five conceptual subcategories: (i) Profitability (Column 2), covering break-even analysis, profit and loss, profit margins and markup, cost structures, and interpretation of financial statements; (ii) Inventory (omitted due to data sparsity), covering stock tracking methods, stock levels and replenishment, inventory turnover, and loss prevention; (iii) Accounting Practices (Column 3), covering recordkeeping methods, basic transaction recording, and separation of personal and business finances; and (iv) Cash Management (Column 4), covering cash flow management, budgeting and financial planning, credit and loan management, and taxation and compliance. Inventory-related outcomes are excluded from the reported regressions owing to insufficient observations. Joint estimation of the accounting subcategories using a linear CMP system produces consistent inferences. Panel B reports marginal effects from Poisson regressions at the entrepreneur level with cluster standard errors. Day of Week Fixed Effects are included. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table 5: A Test of the “Language of Business” through an AI Mentor

	Accounting Integration		Linear Seemingly Unrelated Regressions	
	(1)	(2)	Accounting	Participation
<i>Treatment</i>	10.304 <sup>*</sup> (1.80)	11.100 <sup>**</sup> (2.10)	11.061 <sup>**</sup> (2.07)	20.495 (1.20)
<i>Female</i>		4.265 (0.72)	4.701 (0.90)	29.550 <sup>*</sup> (1.90)
<i>Age</i>		0.279 (1.17)	0.329 (1.30)	1.459 <sup>*</sup> (1.71)
<i>Low Education</i>		11.351 <sup>*</sup> (1.81)	11.415 <sup>*</sup> (1.85)	38.147 <sup>*</sup> (1.95)
<i>Entrepreneurial Motivation</i>		2.815 <sup>**</sup> (2.13)	2.950 <sup>**</sup> (2.26)	8.957 <sup>**</sup> (2.20)
<i>Internet Access</i>		9.782 (1.59)	9.412 <sup>*</sup> (1.87)	35.121 <sup>**</sup> (2.13)
Observations	210	202	202	202

The treatment group includes entrepreneurs in G2 (AI Accounting), while the benchmark control group corresponds to G1 (AI Standard). Columns (1) and (2) report marginal effects from Poisson regressions at the entrepreneur level. Columns (3) and (4) present linear SUR estimates jointly modeling participation with the AI mentor and accounting integration. In untabulated specifications, I address potential endogeneity of participation with the AI mentor; the estimated correlation coefficient  $\rho$  is not statistically different from zero. <sup>\*</sup> $p < 0.10$ , <sup>\*\*</sup> $p < 0.05$ , <sup>\*\*\*</sup> $p < 0.01$ .

Table 6: Integrating Accounting into Functional Area Decisions

Panel A: Entrepreneur Level					
	All Acc. Integration	Acc and Marketing	Acc and Strategy	Acc and Business Plan	Accounting Silo
	(1)	(2)	(3)	(4)	(5)
<i>Treatment</i>	11.100**	2.360**	3.713**	7.041	4.283
	(2.10)	(1.99)	(2.06)	(1.64)	(1.35)
<i>Female</i>	4.265	2.705	-1.430	1.557	3.194
	(0.72)	(1.57)	(-0.73)	(0.34)	(0.96)
<i>Age</i>	0.279	0.060	0.132	0.140	0.072
	(1.17)	(0.97)	(1.63)	(0.74)	(0.56)
<i>Low Education</i>	11.351*	1.675	2.774	8.149	10.108***
	(1.81)	(1.22)	(1.37)	(1.59)	(2.66)
<i>Entrepreneurial Motivation</i>	2.815**	-0.036	0.794	2.190**	1.098
	(2.13)	(-0.13)	(1.59)	(2.10)	(1.36)
<i>Internet Access</i>	9.782	4.756**	-1.509	9.455*	6.977*
	(1.59)	(2.43)	(-0.78)	(1.80)	(1.95)
Observations	202	202	202	202	202
Panel B: Firm-Day Level					
	Accounting Integration	Acc and Marketing	Acc and Strategy	Acc and Business Plan	Accounting
	(1)	(2)	(3)	(4)	(5)
<i>Treatment</i>	0.370**	0.079**	0.124**	0.235	0.143
	(2.10)	(1.99)	(2.06)	(1.64)	(1.35)
Observations	6,060	6,060	6,060	6,060	6,060
Day of Week Fixed Effects	Yes	Yes	Yes	Yes	Yes

The treatment group includes entrepreneurs in G2 (AI Accounting), while the benchmark control group corresponds to G1 (AI Standard). All columns report marginal effects from Poisson regressions at the entrepreneur level with robust standard errors. Accounting integration is defined at the message level, where up to three thematic tags can be assigned to each message. A message is classified as accounting integration if at least one tag corresponds to accounting and at least one other tag corresponds to a different functional area. This measure is further decomposed into specific co-occurrence categories: (i) Accounting–Marketing: accounting appears jointly with marketing in the three-tag set; (ii) Accounting–Business Plan: accounting appears jointly with business planning; (iii) Accounting–Strategy: accounting appears jointly with strategic or enterprise planning; and (iv) Full Accounting (this is not part of the definition of Accounting Integration): only accounting mentions across the three tags, with no other functional category present. In Panel B, all columns report marginal effects from Poisson regressions at the entrepreneur level with cluster standard errors, and the same controls than Panel A are included. Day of Week Fixed Effects are included \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table 7: Can Accounting as the Language of Business Promote Entrepreneurial Growth Actions?

Panel A: General Business Actions				
	Sales Expansion	New Branch	Operational Efficiency	Workers Incentives
	(1)	(2)	(3)	(4)
<i>Treatment</i>	1.187*** (2.64)	0.177* (1.68)	2.478* (1.79)	0.442* (1.76)
Observations	210	210	210	210
Panel B: Longer Term Business Actions				
	Sales Expansion	New Branch	Operational Efficiency	Workers Incentives
	(1)	(2)	(3)	(4)
<i>Treatment</i>	1.873*** (2.94)	0.291* (1.70)	3.944** (2.04)	0.673* (1.72)
Observations	126	126	126	126

The dependent variable in each column is a count of distinct actions identified through LLM-based classification of daily mentor-entrepreneur conversations. Actions are grouped into four domains: sales expansion, new branch establishment, operational efficiency improvements, and worker incentive schemes. The treatment group includes entrepreneurs in G2 (AI Accounting), while the benchmark control group corresponds to G1 (AI Standard). Panel A includes all identified business actions. Panel B presents a subsample of Panel A, focusing on long-term actions, those with a planning horizon of three months or more. All specifications estimate OLS models at the entrepreneur level. The treatment variable equals 1 for entrepreneurs assigned to the accounting-framed AI mentorship (G2) and 0 for those in the generic AI mentorship (G1). \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table 8: How Does AI Mentorship Compare to Human-Led Training in Driving Business Growth?

	Perceived Growth	Monthly Sales	Monthly Profits
	(1)	(2)	(3)
<i>G1. AI Standard</i>	0.057 (0.69)	-1119.658 (-0.45)	-500.000 (-0.37)
<i>G2. AI Accounting</i>	0.163** (2.12)	-1324.786 (-0.54)	0.000 (0.00)
Observations	131	121	116
Panel B: Below Mean of Business Practices Pre-Treatment			
	Perceived Growth	Monthly Sales	Monthly Profits
	(1)	(2)	(3)
<i>G1. AI Standard</i>	0.122 (0.83)	-1100.000 (-0.31)	-591.667 (-0.16)
<i>G2. AI Accounting</i>	0.408*** (3.28)	-1500.000 (-0.52)	-258.333 (-0.07)
Observations	70	61	59

Treatment corresponds to the two AI mentorship modalities, while the comparison (control) group consists of participants in the human-led training program. Panel A reports results for the full sample, whereas Panel B includes only entrepreneurs whose business practice scores were below the mean in the pre-treatment period. Column (1) measures perceived business growth, estimating treatment effects using OLS while adjusting for baseline outcomes. Columns (2) and (3) report monthly sales and profits, respectively, both winsorized at the 1st and 99th percentiles and expressed in lempiras, estimating treatment effects using quantile regression while adjusting for baseline outcomes. All regressions include the lagged dependent variable to control for baseline levels. \*, \*\*, \*\*\* indicate significance at the 10%, 5%, and 1% levels.

Table 9: Does AI Mentorship Improve the Adoption of Accounting Practices?

Panel A: Full Sample				
	Accounting Practices			
	Cost	Profits	Use of Acc. Records	Other Business Practices
	(1)	(2)	(3)	(4)
<i>G1. AI Standard</i>	0.102 (0.65)	0.023 (0.26)	0.012 (0.11)	-0.064 (-1.54)
<i>G2. AI Accounting</i>	0.275* (1.69)	0.092 (1.07)	0.114 (1.01)	-0.031 (-0.76)
Observations	123	131	108	131
Panel B: Below Mean of Business Practices Pre-Treatment				
	Accounting Practices			
	Cost	Profits	Use of Acc. Records	Other Business Practices
	(1)	(2)	(3)	(4)
<i>G1. AI Standard</i>	0.307 (1.12)	0.156 (1.12)	0.009 (0.07)	-0.064 (-1.05)
<i>G2. AI Accounting</i>	0.767*** (2.80)	0.272* (1.90)	0.263* (1.77)	-0.013 (-0.22)
Observations	62	70	55	70

Treatment corresponds to the two AI mentorship modalities, while the comparison (control) group consists of participants in the human-led training program. Panel A reports results for the full sample, while Panel B includes only entrepreneurs whose business practice scores were below the mean in the pre-treatment period. Columns (1)–(3) report accounting-practice indices by topic: cost analysis, profit tracking, and use of accounting records, respectively. Column (4) reports the index of other business practices (covering marketing, suppliers, planning, and human resources). Coefficients are estimated using OLS adjusting for baseline outcomes by including the lagged dependent variable. \*, \*\*, \*\*\* indicate significance at the 10%, 5%, and 1% levels.

Table 10: Do Entrepreneurs With Weaker Baseline Practices Benefit More From AI Accounting Mentorship?

	Total Score (1)	Marketing (2)	Cost (3)	Sales (4)	Profits (5)	Legal (6)	Human Resources (7)
<i>G1. AI Standard</i>	10.969*** (2.92)	4.859*** (3.73)	1.328*** (2.87)	0.306 (0.76)	0.320*** (4.95)	1.965*** (3.82)	3.301*** (3.07)
<i>G2. AI Accounting</i>	3.808** (2.07)	2.139* (1.74)	0.703*** (2.76)	0.694** (2.02)	0.127* (1.88)	1.502*** (3.55)	1.313 (1.18)
<i>Age</i>	-0.997 (-0.42)	0.124 (0.09)	0.677** (2.22)	0.518 (0.82)	0.380*** (3.56)	-1.207** (-2.01)	-0.817 (-0.44)
<i>Female</i>	0.345 (0.18)	-0.136 (-0.14)	-0.364** (-2.32)	-0.196 (-0.64)	0.134** (2.36)	-0.011 (-0.03)	0.591 (0.70)
<i>Low Education</i>	-1.478 (-0.56)	-0.925 (-0.79)	0.135* (1.81)	0.233 (0.83)	0.266*** (4.45)	0.275 (0.75)	0.337 (0.38)
<i>Internet Access</i>	9.825*** (3.29)	4.175*** (3.80)	0.545*** (5.47)	0.376 (1.08)	0.012 (0.21)	1.069** (2.57)	2.226** (2.29)
<i>Entrepreneurial Motivation</i>	-0.797 (-0.90)	-0.424 (-0.79)	-0.060 (-1.63)	-0.006 (-0.06)	0.054* (1.83)	-0.040 (-0.29)	-0.198 (-0.54)
<i>Number of Employees (Excl. Owner)</i>	1.584 (1.05)	0.331 (1.29)	0.285*** (2.74)	0.063 (0.72)	0.115*** (5.19)	0.241* (1.82)	0.551* (1.83)
<i>Firm Age</i>	1.873 (0.57)	0.750 (0.93)	-0.047 (-0.75)	-0.026 (-0.13)	0.044 (1.04)	-0.318 (-0.97)	0.264 (0.37)
Observations	75	75	75	75	75	75	75

Treatment corresponds to the two AI mentorship modalities, while the comparison (control) group consists of participants in the human-led training program. All columns represent the business plan score evaluation. The Business Plan has a total of 120 points, and each topic section has a maximum of 5 points. Marketing (column 2) includes five sections related to the topic, Cost (column 3) is the sum of three sections, and Human Resources (column 7) is the sum of four sections. Other sections, such as organizational structure, executive summary, and capital, are not shown in the table but are included in the total score (column 1). All specifications account for selection into business plan submission using a Conditional Mixed Process, with an exclusion restriction based on whether an entrepreneur was referred to the program by another chamber of commerce. \*, \*\*, \*\*\* indicate significance at the 10%, 5%, and 1% levels.

Table 11: How Do AI Mentors Shape What Entrepreneurs Value in Training?

	Support	Learning Content	Flexibility	Networking
	(1)	(2)	(3)	(4)
<i>G1. AI Standard</i>	0.268*** (2.78)	-0.364*** (-3.81)	0.059 (1.18)	-0.173** (-2.48)
<i>G2. AI Accounting</i>	0.160* (1.74)	-0.207** (-1.97)	0.115** (2.01)	-0.174** (-2.51)
Observations	131	131	131	131

Columns (1)–(4) reports entrepreneurs’ preferences regarding the program: *Support* (1) reflects perceived support in addressing business questions; *Learning Content* (2) evaluates the program’s instructional content; *Flexibility* (3) refers to schedule and convenience; and *Networking* (4) captures interactions with peers. \*, \*\*, \*\*\* indicate significance at the 10%, 5%, and 1% levels.

## **Appendix A. Related Literature**

Despite decades of investment in entrepreneurship development, academic evaluations consistently show that business training programs for microenterprises yield, at best, modest and heterogeneous effects. Across studies, randomized controlled trials (RCTs) show substantial variation in design, sample, content, quality, context, and outcomes, making it difficult to draw clear conclusions about the effectiveness of business training programs and identify best practices (McKenzie et al., 2021). The average treatment effect on firm performance, measured in profits, revenues, or adoption of business practices, is often small and frequently not statistically significant (McKenzie et al., 2021). These findings are not due to a lack of experimentation: training interventions have taken a wide range of forms, including classroom-based group courses (e.g., Giné & Mansuri, 2021; Karlan & Valdivia, 2011), peer interactions (e.g., Cai & Szeidl, 2018; Chatterji et al., 2019), role model mentoring (e.g., Bakhtiar et al., 2022; Brooks et al., 2018), simplified financial literacy modules (e.g., Anderson et al., 2018; Arraiz et al., 2019; Drexler et al., 2014), personalized consulting (e.g., Anderson & McKenzie, 2022; Bruhn et al., 2018), digital delivery mechanisms (Davies et al., 2024; Dell’Acqua et al., 2023), structured decision-making processes (Camuffo et al., 2019, 2024), and assessments of business viability (Lee et al., 2024).

This paper builds on this growing body of work but shifts the focus toward understanding why many training programs fall short of transformative impact. Specifically, it identifies four persistent structural barriers: challenges in scalability, inconsistent training quality, lack of knowledge-on-demand, and the marginalization of financial acumen. These limit not only program effectiveness but also the mechanisms through which entrepreneurs learn, adapt, and apply business knowledge.

The first structural barrier concerns the difficulty of scaling high-quality business training. Most programs rely on skilled human trainers, are delivered in small cohorts, and are costly to replicate. As List (2023) notes, “human capital is difficult, if not impossible, to scale.” Most evaluations take place in small, controlled settings, often with intensive delivery formats that are difficult to sustain at scale. The challenge is not merely logistical but pedagogical: as programs scale, instructional quality often declines, trainer expertise becomes inconsistent, and participant engagement weakens (Al-Ubaydli et al., 2017). A program that yields strong results in one setting may fail elsewhere, not due to a flaw in the concept, but because the enabling conditions (trainer quality, contextual relevance, participant motivation) cannot be reproduced. For instance, attempts to replicate the personal initiative training tested initially by Campos et al. (2017) in Togo found much weaker results in Ethiopia, where implementation relied on less experienced trainers (Alibhai et al., 2019; McKenzie et al., 2021). In economic terms, the supply curve for trainer quality is upward sloping, as programs expand, attracting additional high-quality trainers requires progressively higher costs, and beyond a certain point organizations must either pay a premium to secure top talent or settle for less experienced instructors (Al-Ubaydli et al., 2017). More broadly, traditional models, typically limited to cohorts of 15–40 participants per trainer (McKenzie & Woodruff, 2021), have yet to resolve a central question: how can training reach thousands of entrepreneurs while maintaining instructional quality and achieving meaningful learning outcomes?

Second, training quality remains difficult to define and to standardize. Even if business training programs could be scaled, there is no clear consensus on what constitutes a high-quality program. Studies vary widely in content, structure, and intensity, with some covering dozens of practices in brief sessions, while others focus narrowly on a single skill. Many evaluations combine

multiple pedagogical elements, including mentoring, heuristic-based guidance, and structured curriculum, which limits causal attribution and impedes theoretical development (McKenzie et al., 2021). This is partly because most programs focus on outcomes rather than the learning process itself, leaving fundamental questions about how entrepreneurs absorb, apply, and retain knowledge largely unanswered. As a result, the learning process remains a black box, making efforts to improve or replicate training designs inherently ad hoc.

The evidence in the extant literature increasingly suggests that “more” content is not necessarily “better.” Dense or overly theoretical curricula, especially for low-literacy entrepreneurs, tend to overwhelm rather than empower (Drexler et al., 2014; Francis & Chakravarty, 2025). Simplified and actionable rules-of-thumb often outperform formal instruction. However, even successful simplifications can backfire when they are overly generic. For example, while Drexler et al. (2014) and Arraiz et al. (2019) find that heuristic-based financial training improves business outcomes, particularly for low-literacy or lower-cognitive-score entrepreneurs; similar approaches show limited effects in other settings, such as India and the Philippines (Cole et al., 2024) or Mozambique (Batista et al., 2022). These mixed results suggest that the key is not simplification alone, but delivering small, actionable steps that are aligned with the entrepreneur’s specific needs and constraints. One-size-fits-all rules are insufficient. What’s missing is targeted simplicity, guidance that is responsive to the entrepreneur’s current context and decision-making challenges.

Third, traditional models fail to deliver knowledge-on-demand, i.e., timely, personalized guidance aligned with entrepreneurs’ real-time challenges. Although educational research emphasizes the beneficial role of active knowledge construction vs. passive information reception (Mollick & Mollick, 2023; Muralidharan & Singh, 2025; Roscoe & Chi, 2007), most programs

continue to rely on static formats, such as classroom sessions, printed manuals, or scheduled meetings, that assume learning is linear and can be front-loaded (McKenzie & Woodruff, 2021). However, attending a class is not enough to complete the learning process. For knowledge to be fully absorbed, it must be applied to context-specific decisions, for example, when a supplier renegotiates terms, a client defaults on payment, or cash flow becomes constrained. This is precisely where traditional training programs fall short: limited interaction and time-constrained sessions make it difficult to support the kind of real-time application that is critical for adult learning. This reflects what education scholars call the problem of *transfer*, i.e., the challenge of applying learned concepts across contexts (Mollick & Mollick, 2022). Despite short-term gains, many business training programs fail to produce lasting change, as entrepreneurs often revert to previous habits (e.g., Karlan & Valdivia, 2011). This suggests that static, front-loaded training formats may not support the deep understanding required for entrepreneurs to experience sustained behavioral shifts and transfer the knowledge they acquired in the training. The inability to provide just-in-time assistance limits both learning and the application of knowledge, as well as the understanding of the learning process.

Fourth, a critical, but underexamined, limitation is the marginal role of accounting in business training. Despite being widely referred to as the "language of business," accounting is rarely positioned as a strategic tool for decision-making during business training programs. Instead, it is often taught as a separate module of the program, isolated from key functional areas such as marketing, operations, and human resources. This pedagogical disconnect reduces its perceived managerial relevance and undermines adoption. Applying knowledge in new contexts requires a deep understanding of the underlying structure of a concept, as Mollick and Mollick (2022) explain that "[t]o use what they previously learned, students need to recognize that the

former problem (from math class) is the same problem (ah, just like in math class!) in a new context.”

Even literacy programs that are tangentially related to financial acumen report mixed results and fail to integrate accounting into the cognitive frame entrepreneurs use to navigate business decisions (Drexler et al., 2014; Anderson et al., 2018; and Arraiz et al., 2019). The findings of Drexler et al. (2014) and Arraiz et al. (2019) show a positive effect of rule-of-thumb training for certain types of entrepreneurs may stem not only from its simplicity but also from its direct, context-specific application to daily decisions, in contrast to abstract accounting instruction. However, even this approach stops short of fostering deeper financial reasoning across business functions. Similarly, Anderson et al. (2018) highlight the efficiency gains of accounting training but treat it as a standalone module, separate from marketing. Thus, they miss the opportunity to use accounting as a decision-making lens to evaluate marketing investments and manage trade-offs. As a result, accounting remains peripheral to the cognitive frameworks emphasized in the training programs microentrepreneurs rely on, i.e., viewed more as a tool for recordkeeping than as a foundation for informed, strategic decision making.

This paper addresses these limitations by designing and testing a scalable, AI-based mentorship model that delivers financial acumen through personalized and real-time interaction. Recent research in developed economies and large firms shows that large language models (LLMs) can outperform human analysts in financial forecasting using standardized statements, even without detailed context (Kim et al., 2025). Similarly, Tucker, Wang, and Zhao (2024) find that AI adoption boosts firms' sales, efficiency, and profitability, though gains depend on internal capabilities. Dell'Acqua et al. (2023) further show that LLMs can significantly improve task performance and output quality in professional settings, but only when tasks lie within the AI's

frontier of competence. This suggests strong potential for AI mentorship in microenterprises with limited financial expertise, while also underscoring the importance of guided use. However, general-purpose models still fall short on domain-specific accounting tasks (Wood et al., 2023), highlighting the need for tailored tools.

In designing this intervention, I also draw on emerging research in the field of education and AI, which mainly focuses on formal education but offers valuable insights into how individuals learn and engage with complex material when supported by interactive AI tools. Studies on the integration of AI in educational contexts show that AI can foster personalized and differentiated learning by playing multiple pedagogical roles as a tutor, coach, or mentor through personalized explanations, reflective questions, and support for student understanding across contexts (Mollick & Mollick, 2024; Wu & Yu, 2024; see also Bettayeb et al., 2024, for a systematic review). These educational principles are highly relevant for entrepreneurs, who often face cognitive overload, limited access to feedback, and difficulty transferring abstract concepts into practice. My research builds on evidence from cognitive science and instructional design that shows how feedback-rich, interactive AI systems can improve learning and decision-making (Mollick & Mollick, 2022, 2023).

The most directly related study to this project is Otis et al. (2024), who conduct a field experiment with Kenyan entrepreneurs to evaluate whether generative AI could improve business outcomes relative to no support and find mixed results. The study's intervention relies on generic ChatGPT output, lacks contextual adaptation, user profiling, or memory retention, and is not anchored in a structured training methodology. While Otis et al. (2024) focus on establishing whether AI assistance has any impact, my research compares delivery modes and investigates learning mechanisms, with particular attention to how financial acumen, especially accounting, is

framed and applied in decision-making.<sup>1</sup> Additionally, my study integrates the AI mentor into an established business training curriculum, embeds accounting frameworks across functional areas, and uses real-time behavioral data to analyze engagement and application.

Taken together, the previous literature has a clear pattern: traditional training programs have not failed due to lack of innovation, but because innovations have not addressed the core frictions entrepreneurs face: limited time, cognitive load, evolving needs, and decision complexity. Attempts to layer consulting, mentoring, or capital on top of training have produced some gains (Campos et al., 2017; Giné & Mansuri, 2021), but even bundled approaches often suffer from high costs, low engagement, or short-lived effects. As the literature increasingly recognizes, “one-size-fits-all” training is inadequate. What entrepreneurs need is adaptive, demand-driven support that meets them where they are. In sum, this study contributes to four intersecting literatures: entrepreneurship training, learning processes, accounting adoption, and AI-enabled education, by addressing persistent structural frictions in scale, instructional design, contextual relevance, and financial reasoning.

---

<sup>1</sup> Prior studies using digital tools with limited interaction and predefined rules, such as chatbots and SMS-based training, show limited or fading impact due to low engagement and lack of personalization (de Barros et al., 2024; Mehmood, 2024).

## **Appendix B. The AI-based Mentor Design**

The AI mentor deployed in this study is a WhatsApp-based virtual assistant powered by GPT-4, designed to deliver personalized, scalable business training grounded in the ILO's Start and Improve Your Business (SIYB) curriculum. Its architecture supports a dynamic and iterative learning process that adapts to each entrepreneur's evolving needs in real time. When an entrepreneur sends a message, the system classifies its intent and tags it to one or more functional domains, such as marketing, costing and pricing, or human resources (see Figure A1 below). It then retrieves relevant material from a vectorized document base that includes SIYB manuals, legal formalization guidelines, and targeted educational content, including short video modules. Drawing on this content and using retrieval-augmented generation, the AI constructs a tailored response that reflects the specific question posed, the entrepreneur's profile, and the ongoing conversation context.

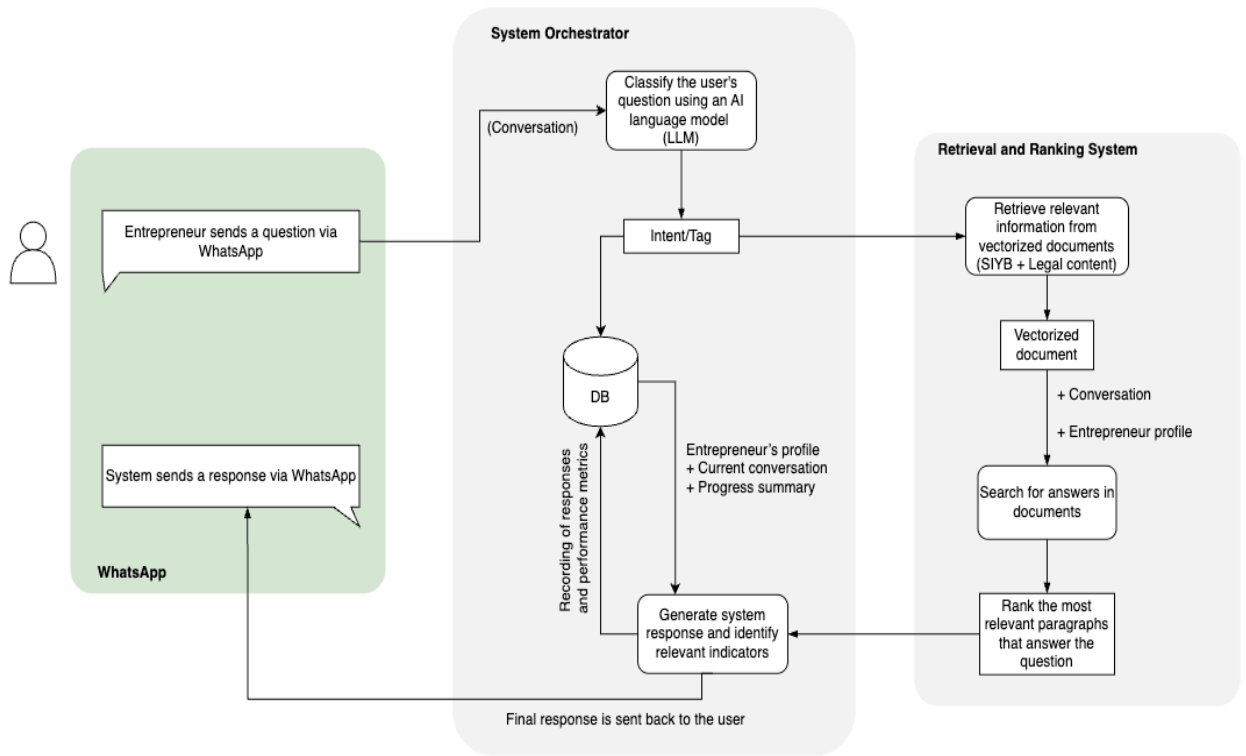
The mentor maintains a memory of all previous interactions and continuously updates each user's profile using metadata from those exchanges. This profile includes information collected at baseline, such as business characteristics and initial knowledge levels, as well as dynamic indicators like inquiry types, engagement with different topics, and inferred gaps in understanding. These elements allow the mentor to personalize both the content and delivery style of future responses, adapting the tone, depth, and complexity of the information provided. In addition to on-demand guidance, the AI mentor periodically reinforces key lessons through follow-up messages, offering reminders, clarification, and optional nudges to revisit previously discussed topics or to explore complementary areas of the curriculum.

To support learning evaluation and ensure alignment with real business challenges, the system automatically generates daily summaries of each user's conversations. These summaries

capture the core business problem raised, the AI's proposed solutions, and the user's level of interest or follow-through. These data are stored in the user's mentorship history and used to inform future interactions. On a weekly basis, the system also prompts entrepreneurs to reflect on whether they implemented the advice, how effective it was, and what challenges remain. This feedback loop enables both accountability and adaptive learning, while also creating a high-frequency behavioral record of the learning process.

Unlike static training models that front-load information without follow-up, this AI mentor builds a progressive and personalized learning arc. It responds in real time to user needs, embeds core business logic, particularly financial reasoning, into everyday decision-making, and creates an intelligent mentorship experience that is both cost-effective and pedagogically grounded. By combining real-time responsiveness, profile-based adaptation, and structured reinforcement, the AI mentor provides a flexible, context-sensitive, and scalable alternative to traditional business training and advisory services.

Figure A1. AI-based Mentor Design



## Appendix C. Example of Accounting Integration (Accounting with Planning)

**Context:** Entrepreneur consults the AI mentor about when to sell a batch of piglets and whether the current age is optimal.

**Entrepreneur:**

*Mia, I sold the batch of piglets at 40 days; I'm not sure if next time I should sell them at that age.*

**AI Mentor:**

*Congratulations on that sale! 🎉 Let's use this opportunity to record how much you spent and earned with this real batch. I just need:*

- *Total sale price of the batch*
- *Total costs for the piglets (feed, medicine, etc.)*

**Entrepreneur:**

*L 1,500 each for 5 piglets.*

**AI Mentor:**

*Perfect. For the costs, should we use the estimate we had of L 1,000 per piglet (feed + medicine) to make a provisional calculation?*

**Entrepreneur:**

*Yes, let's use that estimate.*

**AI Mentor (calculation):**

- **Revenue:**  $5 \times 1,500 = L 7,500$
- **Costs (estimated):**  $5 \times 1,000 = L 5,000$
- **Net profit:**  $7,500 - 5,000 = L 2,500$
- **Profit per piglet:**  $2,500/5 = L 500$
- **Profit margin (on sales):**  $2,500/7,500 = 33.3\%$
- **Markup (on cost):**  $2,500/5,000 = 50\%$

*If you keep them longer, costs will rise. For example, if the **extra cost** were **L 200 per piglet**, total cost becomes **L 1,200 per piglet**. To keep the same margin (~33%), the **target price** should be  $\approx L 1,800 = 1,200/(1 - 0.333)$ .*

**Entrepreneur (reflection):**

*So, for what we had planned as fattening, it's cheaper to raise our own piglets, because it lowers costs.*

## **Appendix D. Entrepreneurs' conversation and business actions**

In the following example from an entrepreneur's conversation, the system identified multiple business actions that reflect both strategic and operational decision-making. The entrepreneur discussed the introduction and profitability analysis of several new products, including rotisserie chicken, chicharron combos, and prepared soups, each supported by detailed cost breakdowns and margin calculations, indicating an intermediate level of accounting use. The conversation also revealed a medium-term staffing plan, as the entrepreneur mentioned hiring a new employee and allocating labor costs to product pricing. In addition, the entrepreneur showed an interest in marketing initiatives, requesting a list of products and promotional text for a storefront sign to attract more customers in the short term. To boost immediate sales, the entrepreneur designed combo offers and price promotions, demonstrating short-term actions to increase revenue. Finally, the conversation highlighted a financial management objective, as the entrepreneur aimed to generate sufficient sales to pay rent by a specific date, reflecting short-term cash flow planning based on simple accounting reasoning. Together, these actions illustrate how the system captures structured evidence of entrepreneurial planning.

Conversation:

### **Entrepreneur**

*Analyze the margin on rotisserie chicken*

### **Tags:**

["financial\_records:cash\_flow\_management","business\_planning:return\_analysis","costs\_and\_pricing:cost\_allocation"]

### **Mentor**

*"Perfect, Mario. To calculate the margin on rotisserie chicken, I need these data:*

- 1 Purchase cost of the chicken (per unit or per pound)
- 2 Cost of seasonings and gas per unit
- 3 Customer sale price per unit
- 4 Quantity you plan to sell per week
- 5 Any packaging or labor cost allocated per unit?

With these numbers I'll estimate the true margin and see whether it's as attractive as chicharrón  
💡

### Entrepreneur

"128 unit cost

Seasonings and gas 12 per unit

Packaging cost 2 lempiras; I plan to sell 40 chickens per week

1500 new employee per week"

#### Tags:

["costs\_and\_pricing:cost\_allocation","business\_planning:return\_analysis","financial\_records:cash\_flow\_management"]

### Mentor

"Excellent, Mario. Here's the full analysis for rotisserie chicken:

📦 Cost per unit:

– Chicken: L 128

– Seasonings + gas: L 12

– Packaging: L 2

– Labor (L 1,500 ÷ 40): L 37.50

👉 Total cost = L 179.50 per chicken

📦 If you sell at L 190 per unit (is that your price?), your profit per chicken would be:

$L 190 - L 179.50 = L 10.50$

📊 Estimated weekly profit (40 chickens):

$L 10.50 \times 40 = L 420$

📌 Margin per unit:

$(L 10.50 \div L 190) \times 100 = 5.5\%$ , much lower than the margin on chicharrón (almost 60%).

💡 This profit is low for the required investment. If chicken helps you attract more customers or lets you sell other higher-margin items, it might work as a loss leader.

Shall we explore slightly raising the price or reducing costs to improve that margin?"

### Entrepreneur

How could I increase the profit margin?

#### Tags:

["costs\_and\_pricing:cost\_allocation","business\_planning:return\_analysis","financial\_records:cash\_flow\_management"]

### **Mentor**

*"To raise that margin, you could apply one or more of these strategies:*

**1** *Increase the price to L 200 if customers perceive more value (e.g., nicer packaging, a side or sauce).*

**2** *Reduce labor cost per chicken: if the employee supports other areas and not only roasting, allocate their cost across products, not just chicken.*

**3** *Buy chicken in bulk or from a cheaper supplier. Is your L 128 price from a fixed supplier or the daily market?*

*Want to try a combo version with rotisserie chicken + another high-margin product like chicharrón?"*

### **Entrepreneur**

*Let's do the calculation without the 1500 cost for the new employee*

#### **Tags:**

["business\_planning:return\_analysis","costs\_and\_pricing:cost\_allocation","financial\_records:cash\_flow\_management"]

### **Mentor**

"Here's the new calculation excluding the employee:

#### **📦 Total cost per chicken:**

L 128 (chicken) + L 12 (gas & seasonings) + L 2 (packaging) = **L 142**

#### **💰 Profit per unit:**

L 190 – L 142 = **L 48**

#### **📊 Margin per unit:**

$(L 48 \div L 190) \times 100 = \mathbf{25.3\%}$

#### **📈 Weekly profit (40 chickens):**

L 48 × 40 = **L 1,920**

This margin is much stronger. If you can sell this product without dedicated staff, it's viable to keep it in the lineup.

Shall we build a combo that includes this chicken with chicharrón to further raise the average ticket per sale? 🍷🐷"

## Entrepreneur

Build a combo with a soft drink

### Tags:

["costs\_and\_pricing:cost\_allocation","business\_planning:return\_analysis","financial\_records:cash\_flow\_management"]

### mentor


"Here's a simple proposal:


#### **“Rotisserie with Drink” Combo:**

- 1 Rotisserie chicken (cost: L 142)
- 1 Soda 500 ml (estimated cost: L 12)

 **Total combo cost:** L 142 + L 12 = **L 154**

 **Suggested sale price:** L 210

 **Net profit:** L 210 – L 154 = **L 56**

 **Combo margin:**  $(L 56 \div L 210) \times 100 = 26.7\%$

This combo helps you move more product without hurting margin and with good value perception for the customer.

Would you also like to try a “family” version with 2 drinks and adjust the price?"

## Entrepreneur

"Let's adjust the size and price of the drink.

1.5-liter drink costs 30 lempiras"

### Tags:

["costs\_and\_pricing:cost\_allocation","business\_planning:return\_analysis","financial\_records:cash\_flow\_management"]

### mentor


"Perfect. Using the 1.5 L soda (cost L 30):

#### **New “Family Rotisserie” Combo**

- 1 Rotisserie chicken (cost: L 142)
- 1 Soda 1.5 L (cost: L 30)

 **Total cost:** L 172

 **Suggested sale price:** L 230

 **Profit:** L 230 – L 172 = **L 58**

 **Margin:**  $(L 58 \div L 230) \times 100 = 25.2\%$

It's attractive for families and keeps a reasonable margin. Want me to calculate weekly profit if you sell 30 combos?"

## Entrepreneur

Let's calculate a chicharrón + tortilla + lime combo

### Tags:

["costs\_and\_pricing:cost\_allocation","financial\_records:cash\_flow\_management","business\_planning:return\_analysis"]

### mentor

"Let's do the chicharrón combo 🍌🐷:

#### 📦 Components per unit:

– 1/4 lb chicharrón → Cost: L  $76.11 \div 4 = 19.03$

– 3 tortillas → estimated L 3

– 1/2 lime → estimated L 2

👉 **Estimated total cost:** L 24.03 (round to L 24 for ease)

💰 **Suggested sale price:** L 45

💰 **Profit per combo:** L 45 – L 24 = L 21

📈 **Estimated margin:**  $(L 21 \div L 45) \times 100 \approx 46.6\%$

It's tasty, profitable, and easy to sell. Do you want to calculate how many of these you need to sell per day to hit your daily target?"

## Entrepreneur

Help me do a calculation with prepared soups

### Tags:

["costs\_and\_pricing:cost\_allocation","business\_planning:return\_analysis","financial\_records:cash\_flow\_management"]

### mentor

"Perfect, Mario 💡. To see whether prepared soups are a good product, I need:

- 1 Cost per unit of the main ingredients (meat, vegetables, soup base, etc.)
- 2 Per-unit costs of gas, seasonings, and packaging
- 3 Customer sale price per unit
- 4 How many soups you plan to sell per week

With that I'll compute total cost per unit, profit, and margin 👍"

## Entrepreneur

"Cost 13.33

Gas per unit 3

Sale price 25

I estimate selling about 15 per week"

### Tags:

["costs\_and\_pricing:cost\_allocation","business\_planning:return\_analysis","financial\_records:cash\_flow\_management"]

**mentor**

"Thanks for the data, Mario. Here's the analysis:

 **Total cost per soup:**

L 13.33 (ingredients) + L 3 (gas) = **L 16.33**

 **Profit per soup:**

L 25 – L 16.33 = **L 8.67**

 **Weekly profit (15 soups):**

L 8.67 × 15 = **L 130.05**

 **Margin per unit:**

(L 8.67 ÷ L 25) × 100 = **34.7%**

That's a decent margin. If sales are steady, this product can help your daily cash flow. Want us to build a combo or a marketing suggestion to boost soup sales? 🍲"

**Entrepreneur**

Let's build a combo with a soda

**Tags:**

["costs\_and\_pricing:cost\_allocation","business\_planning:return\_analysis","financial\_records:cash\_flow\_management"]

**mentor**


"Here's the combo, Mario 🍲🥤 :


**“Express Lunch” Combo**


– 1 Prepared soup (total cost: L 16.33)

– 1 Soda 500 ml (estimated cost: L 12)

👉 Total cost: L 28.33 (round to L 28)

 Suggested sale price: L 40

 Profit per combo: L 40 – L 28 = **L 12**

 Margin: (L 12 ÷ L 40) × 100 = **30%**

Good acceptance at lunch time with a reasonable margin. Want to calculate how many of these you should sell per day to hit your daily target?"

**Entrepreneur**

Give me a list of 6 products to use on a storefront sign

**Tags:**

["costs\_and\_pricing:cost\_allocation","business\_planning:return\_analysis","financial\_records:cash\_flow\_management"]

**mentor**

"Sure, Mario. Here are 6 ideal products to highlight on the entrance sign:

- 1 Rotisserie chicken with 1.5 L soda
- 2 Chicharrón combo + tortillas + lime
- 3 Prepared soup with drink
- 4 Special butter (high demand)
- 5 Sweet bread (new and fast-moving)
- 6 Bulk groceries combo

This attracts both lunch seekers and weekly shoppers 🛒.

Want help with a catchy phrase or layout for the sign?"

**Entrepreneur**

Give me specific, high-rotation, attention-grabbing items

**Tags:**

["marketing:traditional\_advertising","marketing:customer\_preferences","marketing:sales\_techniques"]

**mentor**

"Here are 6 ideal products for the sign, Mario, because they're fast-moving or good hooks to attract customers:

- 1 Whole rotisserie chicken
- 2 Prepared soup
- 3 Chicharrón combo + tortillas + lime
- 4 Fresh-baked sweet bread
- 5 Special butter
- 6 Rice or sugar by the pound (price visible)

You can highlight them with lines like "Ready lunches every day!" or "Fresh offers from L 10."  
Want help drafting the sign text?"

**Entrepreneur**

Help me write the text

**Tags:**

["marketing:customer\_preferences","marketing:traditional\_advertising","marketing:sales\_techniques"]

## mentor

"Here's sample wording for your main sign, Mario:

- ◆ Come for your favorites of the day!
- 🍗 Whole rotisserie chicken
- 🍲 Prepared soup with soda
- 🐷 Chicharrón combo + tortillas + lime
- 🍞 Fresh-baked sweet bread
- 🔪 Special butter
- ⚖️ Rice by the pound from L 10

📍 Open Monday to Sunday – Tasty and affordable here!

Do you want to include hours or a phone number for orders on the sign?"

## Entrepreneur

I need to focus on paying rent by the 10th of this month

### Tags:

```
["marketing:traditional_advertising","marketing:customer_preferences","marketing:sales_techniques"]
```

## Mentor

I understand, Mario. If your goal is to pay rent before the 10th, the best move is to start today by pushing your most profitable combos (chicken, soup, chicharrón). I suggest a limited-time promo through the 9th, like:

- 🎯 “Quick promo through Saturday:
- 🍗 Chicken + 1.5 L Soda for only L 225
- 🍲 Soup combo + drink for L 38
- 🐷 Chicharrón combo just L 42!”

You can use this text on your WhatsApp status, direct messages, and a sign at the entrance. Want me to help you calculate how much you should sell per day to make sure you cover rent?

## **Appendix E. Business Plan Evaluation**

As part of the program, entrepreneurs developed a structured business plan using a standardized template introduced at the outset of training. Completing this plan was a prerequisite to participate in the Seed Capital Competition, where the most promising projects received some implementation grants. The template covered nine core areas of business design and financial management: (i) Executive Summary, describing what the business offers, its target market, and financing; (ii) Business Idea, defining the product or service, market opportunity, and client need; (iii) Market Research, assessing understanding of customers, competitors, and potential gaps; (iv) Commercialization Plan, detailing product quality, pricing, distribution, promotion, personnel management, and business image; (v) Sales Estimates, ensuring realistic projections; (vi) Organizational Structure, clarifying roles and responsibilities; (vii) Formalization and Legal Aspects, considering compliance and fiscal awareness; (viii) Costing and Financial Planning, verifying accuracy of costs, profits, and cash flow; and (ix) Initial Capital and Financing, justifying funding needs and sources. Each section was scored from 1 (minimal compliance) to 5 (fully coherent and realistic).

Using these criteria, I employed a Large Language Model as an independent “judge” (LLMaJ) to systematically evaluate all submitted business plans. The LLMaJ received each plan along with the structured evaluation form and rubric, producing section-level scores and concise qualitative comments. This approach provided a consistent, transparent, and scalable assessment across treatment groups. To validate the automated evaluation, CCIT coaches independently reviewed a subset of business plans, and their comments were highly consistent with those of the LLM, confirming alignment between human and model-based evaluations. The evaluation form

was jointly designed by the Chamber of Commerce and Industry of Tegucigalpa (CCIT) to ensure conceptual validity and comparability across participants.

The seed capital selection followed four phases: eligibility screening, automated evaluation, technical verification, and final presentation. In the first phase, only participants in at least the 60th percentile of program participation (based on attendance for the in-person group and interaction metrics for AI-mentored groups) who submitted a complete business plan advanced to the next stage. In the second phase, the LLM-as-Judge conducted the automated assessment, producing section-level scores and rankings from which the top fifteen plans per group (forty-five in total) were shortlisted. In the third phase, CCIT advisors verified completeness and confirmed eligibility; no substitutions were necessary, as all shortlisted plans met the required standards. Finally, the forty-five entrepreneurs with the highest verified scores presented a three-minute elevator pitch before the selection committee, which determined the recipients of the seed capital awards.

## Appendix E. Recordkeeping examples

Below are photos shared by entrepreneurs with the AI mentor documenting their recordkeeping practices. These images illustrate the types of businesses in the study and the range of accounting approaches observed. Some notebooks are well organized and structured; others contain only rough, disorganized notes. A subset of entrepreneurs also use Excel spreadsheets instead of paper notebooks.

Exhibit E1.

MCS	can. de frascos	Gastos totales
Junio	ventas total 134 frascos 16,080	En Insumos Botes = 2640 Fruta = 5,250 Azucar = 840 Estiker = 420 Energia = 7,500 pasajes = 1,500 <hr/> 7,550
		Dejando un margen de ganancia 4,530

Exhibit E2.

Fecha	Detalle	Ingreso	Egreso
01/7/25	Compra de cajas (Materia prima)		600.00
02/7/25	Venta a piñatas aso Cu	500.00	
03/7/25	Venta de ropa Interior	450.00	
04/7/25	Regalo de cumpleaños	7,000.00	
05/7/25	Compra de papel crepe		450.00
06/7/25	Venta de una piñata	280.00	
07/7/25	Venta de ropa	380.00	
08/7/25	Venta de un par zapatos	450.00	

Exhibit E3.

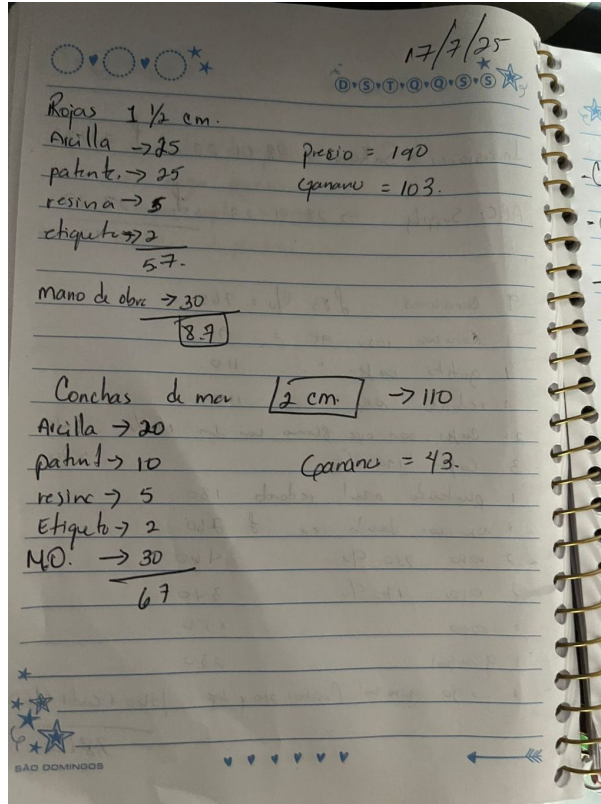


Exhibit E4.

INGRESOS Y EGRESOS DE PIÑATER

Fecha	Descripción	Ingreso	Egreso	#VALOR!
1/7/2025	Compra de Cajas		600	-600
2/7/2025	venta de 2 piñatas	500		-100
3/7/2025	venta de ropa interior	450		350

**Figure 1. Timeline for treatment groups**

