

The Uneven Impact of Generative AI on Entrepreneurial Performance

Nicholas G. Otis*

Berkeley Haas

Rowan Clarke

Harvard Business School

Solène Delecourt

Berkeley Haas

David Holtz

Berkeley Haas

Rembrand Koning

Harvard Business School

[\[Click here for latest version\]](#)

Scalable and low-cost AI assistance has the potential to improve firm decision-making and economic performance. However, running a business involves a myriad of open-ended problems, making it difficult to know whether recent AI advances can help business owners make better decisions in real-world markets. In a field experiment with Kenyan entrepreneurs, we assessed the impact of AI advice on small business revenues and profits by randomizing access to a GPT-4-powered AI business assistant via WhatsApp. While we are unable to reject the null hypothesis that there is no *average* treatment effect, we find the treatment effect for entrepreneurs who were high performing at baseline to be 0.27 standard deviations greater than for low performers. Sub-sample analyses show high performers benefited by just over 15% from the AI assistant, whereas low performers did about 8% worse. This increase in performance inequality does not stem from differences in the questions posed to or advice received from the AI, but from how entrepreneurs selected from and implemented the AI advice they received. More broadly, our findings demonstrate that generative AI is already capable of impacting—though in uneven and unexpected ways—real, open-ended, and unstructured business decisions.

*Otis: notis@berkeley.edu (Corresponding author); Clarke: rclarke@hbs.edu; Delecourt: solened@berkeley.edu; Holtz: dholtz@haas.berkeley.edu; Koning: rem@hbs.edu. All authors after Otis are listed in alphabetical order. The authors gratefully acknowledge financial support from the Agency Fund, the South Park Common Social Impact Fellowship, the Weiss Fund, Microsoft, the Cora Jane Flood Endowment at Berkeley Haas, Berkeley Haas, the Digital Data Design (D³) Institute at Harvard, and Harvard Business School. The authors thank the Busara Center, and especially Channing Jang for their exceptional support in our field operations, and Sara Dahiya, Brian Mutisyo, Elyes Laalai, Kelsey Shipman, and Noa Ben Haim for superb research assistance. The authors are also grateful to M. Hassan Siddique and M. Tariq Sajid for their assistance in developing the AI assistant, and to Alex Pompe for his assistance with this project. We have benefited from seminar and conference feedback at UC Berkeley, UPenn Wharton, D³, MIT CODE, CFXS, Maryland, Strategy Science, INSEAD, HBS, AEIA, Microsoft, Northeastern, the World Bank, and the University of Washington. This project was approved by the U.C. Berkeley Office for the Protection of Human Subjects.

1 Introduction

Since the launch of ChatGPT in November 2022, there has been an explosion of research on generative AI and its potential economic implications (Mollick, 2022; The White House, 2022; Agrawal, Gans, and Goldfarb, 2023; Eloundou et al., 2024; Goldfarb, 2024). Much of this recent work is driven by the belief that conversations with large language models (LLMs) can help people learn and develop new skills (Choi et al., 2023; Mollick and Mollick, 2022), and in work contexts, can improve firm performance and growth (Brynjolfsson, Li, and Raymond, 2023; Dell’Acqua et al., 2023; Noy and Zhang, 2023; Peng et al., 2023; Kumar et al., 2023). Given the substantial variation in worker and firm productivity, both within and across countries, the emergence of nearly zero marginal cost generative AI “assistants” has the potential to radically improve the productivity and performance of everyone, from the thousands of CEOs running companies listed on the New York Stock Exchange to the hundreds of millions of entrepreneurs running small and medium-sized businesses in developing economies (McAfee, Rock, and Brynjolfsson, 2023; Björkegren, 2023).

Consistent with the optimism currently surrounding generative AI (AI hereafter), recent experiments show that conversing with AI and receiving AI assistance causes workers to write better business text more quickly, including press releases, ad copy, consulting memos, and customer support messages (Brynjolfsson, Li, and Raymond, 2023; Dell’Acqua et al., 2023; Doshi and Hauser, 2023; Noy and Zhang, 2023; Chen and Chan, 2023). However, it remains unclear whether the benefits of such AI feedback generalize to the broader set of tasks that firms engage in. Beyond these text-based tasks, firms must also manage employees, raise capital, pilot new initiatives, run advertising strategies, price their services, react to competitors, and decide which of these and myriad other tasks to focus their efforts on (Chandler, 1977). The sheer multitude of tasks involved in running a business greatly increases the complexity of effectively learning how to improve business performance, including for early-stage or small business entrepreneurs (Kim, 2024). Even when they have access to human advisors and mentors, entrepreneurs often struggle to decide which tasks they should ask for help with, how to formulate effective questions to get useful feedback, and how to interpret and act upon the advice they receive (Bryan, Tilcsik, and Zhu, 2017;

Camuffo et al., 2020; Agrawal, Gans, and Stern, 2021; Dimitriadis and Koning, 2022).

These challenges are potentially exacerbated when seeking advice from AI systems. Entrepreneurs may be reluctant to ask questions of an AI rather than another person (Lebovitz, Lifshitz-Assaf, and Levina, 2022). Without good questions, even the most advanced AI system will fail to provide relevant and useful guidance. Even if an entrepreneur’s question is well-crafted, it is unclear how useful advice from AI will be in practice. Given that many business tasks are tacit and may not yet be codified as text—especially tasks in emerging markets—AI tools may lack the relevant training data to provide helpful answers (Autor, 2014; Tao et al., 2023). It is also possible that AI can generate ineffective or overconfident “solutions,” which, if implemented, might worsen rather than improve firm performance (Dell’Acqua et al., 2023; Ji et al., 2023). As a consequence, entrepreneurs may lack the judgment and complementary knowledge, skills, and resources to select and implement helpful AI-generated suggestions over harmful ones (Brynjolfsson and Hitt, 2000; Brynjolfsson, Rock, and Syverson, 2021). Yet, if entrepreneurs can ask good questions, if the AI can suggest helpful pieces of advice, and if entrepreneurs can select and implement this advice, then AI advice has the potential to enable entrepreneurs across the globe.

In this paper, we test the potential of AI in a randomized control trial (RCT) with 640 Kenyan small business entrepreneurs who were randomly assigned to receive access to a GPT-4-powered AI business assistant via WhatsApp. We are unable to reject the null hypothesis that on average, access to generative AI did not have an effect on business performance relative to the control. Although we cannot reject that our average treatment effect on business performance is zero, this estimate masks substantial pre-registered heterogeneity with respect to pre-treatment firm performance. We estimate the treatment effect for entrepreneurs with above-median baseline performance (high performers) to be 0.27 standard deviations greater than for entrepreneurs with below-median baseline performance (low performers). Sub-sample analysis shows that the AI assistant caused a 15% increase in performance, for high performers but that for low performers, the treatment effect led to an 8% decline in performance relative to the control group.

To better understand this effect, we engage in exploratory text analysis of the questions that entrepreneurs asked over WhatsApp, the answers generated by the AI assistant, and

survey responses from the entrepreneurs describing their business changes. This analysis reveals that the performance effects we observe are not rooted in differences in the number or kinds of questions asked, in the types of suggestions generated by the AI, or in the likelihood of making business changes based on the AI’s advice. Instead, we find that low and high performers selected and then implemented different aspects of the nearly two dozen suggestions that the average entrepreneur received from our AI assistant. We find that treated low performers were especially likely to implement “generic” advice focused on lowering prices and investing in advertising, which, if inappropriate for the entrepreneur’s business situation, would decrease revenues and increase costs. Conversely, treated high performers worked with the AI to discover tailored and specific changes that benefited their businesses. These changes ranged from finding alternative ways to get power during blackouts to introducing a new in-demand car wash detergent. Alongside our performance results, these findings highlight that our AI assistant enabled both low- and high-performing entrepreneurs to ask questions and receive more advice. However, only high performers were able to effectively screen the AI-generated advice and implement valuable, as opposed to detrimental, suggestions.

Beyond identifying the impact of generative AI in our specific research setting, our study also makes multiple contributions to the conversation regarding the technology’s broader economic impacts. First, our research sheds light on whether and when AI advice can improve entrepreneurial learning and firm productivity, both of which are core concerns for managers and policymakers. A large and diverse literature shows that business training programs can lead to modest but meaningful improvements in firm productivity ([McKenzie, 2021](#)). For the high performers in our sample, our point estimates suggest that access to generative AI may have driven improvements comparable in magnitude to the impact of the most effective business training programs, suggesting that generative AI offers an opportunity to scale personalized business mentorship and training at a fraction of the cost of existing programs. Second, understanding if AI can assist entrepreneurial decision-making provides more general insights into whether generative AI can provide meaningful benefits in contexts that are more complex and interconnected than those studied in prior work ([Sato et al., 2023](#)). Our findings suggest that while AI can already produce a plethora of potentially useful advice for a multitude of business tasks, understanding which advice users choose to

implement is crucial to determine who will benefit from these new AI tools (Dell’Acqua et al., 2023; Wiles and Horton, 2024). Third, our findings show how generative AI tools are already capable of impacting business decisions—both for better and for worse—across the globe (Hui, Reshef, and Zhou, 2023; Wiles, Munyikwa, and Horton, 2023).

2 Experimental Design

We measure the impact of AI advice on entrepreneurial performance by running a field experiment with 640 Kenyan entrepreneurs running small and medium businesses. We selected this research context for two reasons. First, while recent experiments document that entrepreneurs in developing economies greatly benefit from the context-specific and wide-ranging nature of human feedback (Chatterji et al., 2019; Dimitriadis and Koning, 2022), this context has been almost entirely overlooked in recent research on the potential productivity impacts of generative AI (Björkegren, 2023). Second, obtaining personalized advice is challenging in developing economies, and the benefits of AI advice in these contexts might be especially large (Dimitriadis and Koning, 2024).

Specifically, we developed an AI assistant for Kenyan entrepreneurs using GPT-4, a large language model released by OpenAI in March 2023 (OpenAI, 2023). ChatGPT, a simple Q&A interface for interacting with LLMs released by OpenAI, scaled to an estimated 100 million monthly active users within two months of its launch (Hu, 2023). Building on the popularity of ChatGPT’s chat interface, we developed our own AI tool that can be accessed through WhatsApp. Development of this AI assistant took place over four months and involved extensive user testing by the authors, research assistants, and small business owners in Kenya (see Appendix D for more details on the AI assistant development process). We selected WhatsApp because of the low cost of sending text over WhatsApp relative to SMS texts, and because internet usage via smartphones is widespread in developing economies (Ramdas and Sungu, 2022); in Kenya, WhatsApp is used by nearly 90% of people (Wamuyu, 2020). Figure 1 depicts an example participant interaction with the AI assistant.

Interacting with the AI assistant differs from interacting directly with an LLM such as GPT-4 in three ways. First, we developed a system prompt that provided the AI with

additional context about the Kenyan small and medium businesses that would use our app, which increased the odds that the advice provided to entrepreneurs by the AI assistant was contextually relevant. For example, if asked about how to raise capital, GPT-4 may recommend raising venture capital funds; a financing strategy that is out-of-scope for the small and medium businesses in our study. Instead, our system prompt would lead the AI assistant to focus on contextually relevant alternatives, like approaching family and friends or building a “chama,” an informal cooperative savings group. Second, as illustrated in [Figure 1](#), we instructed our AI assistant to generate three to five practical pieces of advice in response to each question posed by the entrepreneur, with each piece of advice accompanied by an explanation of its benefits and implementation details.

Third and finally, to encourage entrepreneurs to engage further with the advice they received, the system prompt instructed the AI assistant to structure responses in such a way that users could quickly and easily ask for more information by simply typing a number to learn more about the first suggestion, the second suggestion, and so on. The diverse ways that entrepreneurs interacted with the AI assistant and used these numeric followups, as well as the level of personalization provided by the tool are further highlighted in [Appendix I](#), which presents full chat logs for two entrepreneurs in our sample. In one conversation, a restaurant owner is considering changing the menu and asking for assistance thinking through the possibilities and sources of uncertainty involved in making this decision. In another conversation, a business owner selling wholesale and retail milk is asking how to expand their product offerings to increase profits. Other conversation topics across our full sample include how to motivate employees, the best way to deploy capital when expanding a store, tips for hatching and raising healthy chickens, and how to deal with bankruptcy.

Our AI assistant is one of many AI tools that have recently been developed to assist entrepreneurs and firms ([Baxter and Schlesinger, 2023](#)). However, despite the popularity of these tools, there is scant evidence of their causal impact. This could be in part because measuring the causal impact of AI assistance on business performance presents multiple types of challenges. Conceptually, defining the appropriate counterfactual for the AI assistant is non-trivial because AI-based interventions can offer multifaceted benefits. Some of these benefits can also be obtained with non-AI interventions, whereas others cannot. In

our context, the AI assistant provides personalized advice but also an opportunity for the entrepreneur to reflect on their business needs. As a result, the treatment effect of access to the AI assistant relative to an unassisted control might simply reflect the fact that giving entrepreneurs any form of assistance nudges them to think more about how to improve their business. Empirically, measuring business performance in developing economies almost always involves surveying participants, which raises concerns that giving entrepreneurs an AI assistant might lead to “demand effects,” i.e., changes in the behavior of participants due to cues about what is considered appropriate and/or desirable (Zizzo, 2010).

To address these conceptual and empirical issues, we tested our AI tool against a control group that received a placebo intervention: business guides designed specifically for entrepreneurs operating small and medium businesses in developing economies (International Labor Organization, 2015). Despite the ubiquity of business training guides and manuals, prior research shows—and data from our experiment confirms¹—that access to business training texts almost always has no impact, as asynchronous online or offline guides tend to be rarely used (Cusolito, Darova, and McKenzie, 2023; Mehmood, 2023; Davies et al., 2024). As a result, distributing guides to the control allows us to rule out a “placebo effect” that simply intervening can improve performance while also reducing the potential for demand effects. Beyond these benefits to internal validity, providing the control access to a guide increases the policy relevance of our results by directly pitting the value of AI assistance against the current “standard of care” available to most emerging market entrepreneurs.

We recruited entrepreneurs into our study over the Meta ad platform in partnership with the Busara Center, a Kenyan research organization, starting in May 2023. Figure 2 and Appendix C provide additional details on the experiment timeline and sample recruitment process. Our recruitment strategy involved running ads on Facebook to invite entrepreneurs to a short paid survey (Figure A1). All entrepreneurs who responded to our ads were required to pass basic attention checks and take part in three rounds of pre-treatment surveys delivered over the course of three months, which helped us reduce post-treatment attrition and ensure valid causal inference. Appendix E describes our surveys, which asked about

¹As discussed in our Results section below, and in Appendix K, only 1 of the 640 entrepreneurs in our sample directly mention the business training guide when describing how they changed their business in the post-treatment period.

firm profits, revenues, and management practices, among other measures. Our final sample includes 640 Kenyan entrepreneurs who completed all three pre-treatment surveys. The median entrepreneur in this sample was 26 years old, had been running their business for one year, and held a college degree (Table A1). Our sample reflects the heterogeneity present amongst Kenyan small and medium businesses, with entrepreneurs running businesses from fast-food joints to poultry farms to cybershops² across Kenya (Figure A2 and Figure A3). Pre-treatment performance in our sample ranged from monthly profits of 2,650 Kenyan Shillings (5th percentile; about \$15 USD) to over 55,000 Kenyan Shillings (95th percentile; about \$340 USD).

Following the final pre-treatment survey wave, the entrepreneurs in our sample were block-randomized into treatment and control, with entrepreneurs stratified based on their gender and pre-treatment business performance (Appendix C). Treated entrepreneurs received free and unlimited use of the WhatsApp-based AI assistant, along with regular reminders to use the tool. Control participants were provided easy access to the aforementioned business training guides and were also sent regular reminders to use the guides (Appendix D). Control participants were not provided any information about generative AI, nor the existence of the AI assistant. Entrepreneurs in the two groups were comparable in terms of both performance and the other characteristics we measured prior to treatment (Table A1). Of the 640 participants who were randomized into control or treatment, 634 (99%) completed at least one post-treatment survey and 622 (97%) completed all four post-treatment surveys deployed over the two months following treatment (Appendix C).

Because our main outcome, firm performance, exhibits considerable variability in emerging markets (Anderson, Lazicky, and Zia, 2021), we pre-registered several steps to improve the statistical precision and credibility of our experimentally estimated causal effects (see Appendix E for more details on the pre-registration). First, our outcome variable is an index that combines standardized measures of weekly and monthly revenue and profits to reduce noise in our dependent variable. Second, we analyze our data using a variant of simple ordinary least squares (OLS) regression that conditions on the average pre-treatment perfor-

²Cybershops provide a range of services ranging from computer access, internet browsing, printing, scanning, photocopying, and computer repair and technical support.

mance data collected from participants in the three pre-treatment survey waves (McKenzie, 2012). Third, we control for additional pre-treatment variables using a double-LASSO approach and report the covariates selected for each of our main analyses in Table A5 (Belloni, Chernozhukov, and Hansen, 2014). Fourth, we pool data from all four of our post-treatment periods to further increase statistical power. We detail the construction of our outcome variables and our econometric strategy in Appendix E.

3 Results

Our main pre-registered performance results are depicted in Figure 3. Using the regression specification outlined in our pre-analysis plan, we first estimate a null average treatment effect (ATE) of AI access on business performance (Estimate A in Figure 3). This outcome holds whether we winsorize performance at the 99% level ($\beta = 0.05$ standard deviations (s.d.), $p = 0.36$, 95% CI = $[-0.06, 0.16]$) or the 95% level ($\beta = -0.01$ s.d., $p = 0.92$, 95% CI = $[-0.14, 0.07]$). As illustrated in Figure A4, the results also remain the same after excluding 18 participants who did not complete our entire post-treatment survey panel (95% winsorization: $\beta = -0.01$ s.d., $p = 0.86$, 95% CI = $[-0.14, 0.07]$; 99% winsorization: $\beta = 0.05$ s.d., $p = 0.34$, 95% CI = $[-0.06, 0.16]$).

Building on recent studies showing that AI assistance especially helps lower productivity workers (see Table A2), we estimate heterogeneous treatment effects (HTE) by splitting our sample into two groups, based on the pre-treatment version of the same performance index measure used as our primary outcome variable: *initial low performers* (below-median pre-treatment performance) and *initial high performers* (above-median).

Estimates B and C in Figure 3 show the results when we decompose our average treatment effect into these sub-samples of low and high performers. We find that for low performers access to the AI assistant reduced average business performance relative to the control condition ($\beta = -0.09$ s.d., $p = 0.007$, 95% CI = $[-0.15, -0.03]$). Conversely, for pre-treatment high performers, we find an increase in the average business performance compared to control ($\beta = 0.19$ s.d., $p = 0.07$, 95% CI = $[-0.02, 0.38]$), though this effect is noisier given the long right tail of our business performance outcomes. Table A4 shows this pattern—

negative treatment effect estimates for low performers and positive but noisier estimates for high performers—holds across a variety of specifications.³

Figure 4 plots the effect of the AI assistant on the performance distribution of initially low-performing firms (Panel A) and initially high-performing firms (Panel B). In both cases, there are noticeable differences between the treated and control empirical cumulative distribution functions (eCDFs): In Panel A, the control eCDF is shifted to the right of the treated participants’ eCDF, while in Panel B, this pattern is reversed. These results clarify that the difference in treatment effects observed does not reflect the impact of a handful of outliers but instead points to the AI shifting the distribution of profits and revenues for low and high performers differently.

Consistent with Figure 4, estimate D in Figure 3 presents our last pre-registered test for treatment effect heterogeneity. We find initially high-performing entrepreneurs experience a 0.27 s.d. larger causal effect ($p = 0.013$, 95% CI = $[-0.06, 0.48]$) from the AI assistant compared to initial low performers. Figure A5 presents robustness checks of which the most conservative indicate a heterogeneous effect of 0.23 standard deviations ($p = 0.033$, 95% CI = $[0.02, 0.42]$). Appendix H rules out the possibility that our ATE or HTE estimates are due to spillovers.

In Appendix J we present further evidence that these patterns are not simply the result of chance by applying recently developed “Generic Machine Learning” approaches for discovering and testing heterogeneous treatment effects (Chernozhukov et al., 2018; Davies et al., 2024). Using these methods, we strongly reject the null that there is no detectable effect heterogeneity in our sample ($p < 0.0001$). Moreover, this approach allows us to split the sample into entrepreneurs estimated to have below and above median treatment effects. For entrepreneurs estimated to have below median treatment effects, we find the average impact of the AI for this group is -0.269 S.D. ($p < 0.0001$, 95% CI = $[-0.427, -0.116]$). For the group estimated to have above median treatment effects we find a group average

³Table A7, Table A9, Table A8, and Table A10 shows this pattern—negative treatment effect estimates for low performers and positive but noisier estimates for high performers—holds when we analyze revenues and profits directly. Subsection E.4 discusses why our weakest set of performance results in Table A6 are the result of inappropriately applying an inverse hyperbolic sine (IHS) transformation to our index measure. Appendix E details the three other heterogeneous treatment effect models we pre-registered, and that we find no heterogeneity with regard to the other two variables we planned to test (gender or prior ChatGPT use).

effect of 0.347 S.D. ($p < 0.0001$, 95% CI = [0.144, 0.555]). Consistent with our pre-registered heterogeneity results, we find these groups do not differ in terms of the share who are women ($\delta = 0.001$, $p = 0.919$, 95% CI = [-0.058, 0.059]) or who had used ChatGPT prior to our study ($\delta = -0.041$, $p = 0.127$, 95% CI = [-0.101, 0.019]), but that the fraction of high performers is 33.8 percentage points greater ($p < 0.0001$, 95% CI = [0.279, 0.398]) in the above versus below median treatment effect group. These results, along with our pre-registered analyses, are inconsistent with the sharp null hypothesis that our treatment did not have an impact on any individual entrepreneur. Instead, they suggest that using the AI assistant caused both negative and positive treatment effects, with the benefits largely accruing to entrepreneurs who were already performing better.

The treatment effects we observe are especially surprising relative to recent experiments focused on the economic impact of AI (Brynjolfsson, Li, and Raymond, 2023; Dell’Acqua et al., 2023; Doshi and Hauser, 2023; Noy and Zhang, 2023; Chen and Chan, 2023). Papers in this literature have found consistently positive and significant average treatment effects on productivity and performance, with effects up to roughly 0.40 s.d. (See Appendix F). In contrast, the largest average treatment effect we estimate across specifications is 0.05 S.D. ($p = 0.356$, 95% CI = [-0.06, 0.16]), suggesting that our overall null is informative relative to prior estimates. Moreover, as mentioned above, these experiments all find that pre-treatment low performers are the most likely to benefit from AI assistance, leading the impact of AI tools to *reduce* inequality (Again, see Appendix F). Yet, our field experiment produces the opposite results. Table A12 shows that our negative point estimate among low performers is equivalent to an approximately 8% drop in profits or revenue among low performers, while our positive point estimate among high performers is equivalent to an approximately 15% increase in profits or revenue. Given the stark differences between our results and those reported by recent studies of generative AI, we next present a series of exploratory abductive analyses conducted with the aim of identifying mechanisms that can both explain the heterogeneous treatment effects we observe and help rationalize our findings relative to these recent experiments (King, Goldfarb, and Simcoe, 2021).

A major difference between our field experiment and prior research on the economic impacts of generative AI is the level of discretion participants were granted in how and when

they used AI assistance. In previous experiments, experimenters narrowly constrained the tasks participants completed, the extent to which they engaged with the AI tool (Brynjolfsson, Li, and Raymond, 2023), and/or the length of time they were granted use of generative AI (Noy and Zhang, 2023; Dell’Acqua et al., 2023). In contrast, the entrepreneurs in our experiment had much more discretion over whether and how they used AI assistance. For instance, they were able to choose how many questions to ask the AI assistant and the importance of the tasks on which they requested assistance. They could also ask for assistance on a wide range of topics, including, but not limited to, financing, marketing, operations, and farming practices. In turn, differences in the quantity, types, and structure of questions could result in large differences in the types and quality of advice the AI assistant generated, especially when compared to narrower tasks that have been the focus of prior research. Finally, even if two entrepreneurs received the same set of AI-generated recommendations, differences in which suggestions each selects and implements could also result in heterogeneous performance effects.

To understand whether differences in questions asked, advice generated, or suggestions implemented account for our findings we turn to three sets of textual data. First, for the 322 treated entrepreneurs, we have the full text of the 4,810 messages sent by the 275 entrepreneurs who used the AI assistant at least once. On average treated entrepreneurs sent 17.1 messages to the AI assistant, of which 4.96 were numeric follow ups. For this same group, we also have the text of the response generated by the AI assistant. As illustrated in Figure 1, our prompt leads the AI to respond with 3-5 actionable pieces of advice when appropriate. Over the course of our experiment, treated entrepreneurs received just under 24 such suggestions. Table A26 shows the 200 most common pieces of advice from the 4,207 different suggestions the AI generated. These pieces of advice run the gamut and include: “Community Engagement” (suggested 59 times, the most common idea suggested), “Location” (28), “Diversify Your Product Range” (25), “Social Media Marketing” (17), “Promotions & Discounts” (8), “Employee Training” (6), and “Seeds” (5). Finally, for both the treated and control groups, we have text from survey data that asked entrepreneurs to describe the product, services, process, and most impactful changes they made to their business in the last 30 days. We concatenate these four survey questions into a single text

variable that we then use in our analysis. [Appendix K](#) provides further details on these datasets and survey questions.

We start by checking if low- and high-performing entrepreneurs differ in their propensity to ask questions or in the types of questions they ask. [Table A17](#) and [Table A18](#) show low and high performers are equally likely to ask the AI at least one question (88.2% versus 84.2%), send in a similar number of messages (18.6 versus 15.5),⁴ are equally likely to engage in numeric followups (5.34 versus 4.57), and write questions of similar length (13.5 words per question versus 14.4). [Table A20](#) shows none of these differences are statistically significant. To check for differences in the content of the questions, we follow [Choi et al. \(2023\)](#) to first use GPT-4 to build a candidate list of topics present in the corpus of questions. Distilling this list down to 10 topics, we then used GPT-4 to classify each question as most related to one of these topics. [Table 1](#) shows the ten topics, an example question representative of each topic, and the fraction of questions asked by low and high performers that focus on each. For low and high performers, roughly 20% of questions are primarily focused on discovering “entrepreneurial opportunities and business ideas” and focus on how to launch a new type business beyond what the entrepreneur was currently doing. The next most common questions focus on farming (about 15% of the total), then finances, business planning, customer acquisition, marketing, technology, operations, market research, and finally risk management. While the questions in [Table 1](#) illustrate the diversity of topics the AI had to advise on—from how to deal with crop pests to digital payment systems—there are no clear high-level differences between high and low performers in terms of what is being asked about. To test if there are more fine-grained differences in the types of questions asked, we estimate 3,072 dimensional word embedding vectors for each entrepreneur on the full set of messages each entrepreneur sent to the AI assistant. Using these embedding vectors, we use 80% of our sample to train a random forest model to predict if an entrepreneur is a low or high performer. If there are systematic differences in the types of questions asked, the model should pick up on them, enabling us to predict if an entrepreneur in the test set is a high or low performer. [Table A19](#) shows that our model’s accuracy is 49.2% on the test

⁴The median number of messages sent is especially similar, with low performers sending in 9 questions at the median and high performers 8.5.

set ($p = 0.74$, 95% CI = [0.42,0.58]), worse than the no information rate of 51.8%, again suggesting that there are no meaningful differences in the types of questions asked.

Unsurprisingly, given that the only personalized information the AI assistant learns about an entrepreneur comes from the messages it receives, we also find that the AI’s advice and suggestions are also similar across groups. Panel B [Table A17](#) shows low and high performers receive 21.5 and 25.6 pieces of advice on average, though at the median, low performers receive 13 whereas high performers receive 11 pieces of advice. The length of the AI’s responses are also similar, with low performers receiving messages that run for 168 words versus 166 for high performers. None of these differences are statistically significant. In terms of content, [Figure A10](#) shows the answer topic distribution is similar for high and low performers. Finally, [Table A19](#) shows that training a model to predict performance status on the answer embedding vectors results in an accuracy rate of 48.2% ($p = 0.86$, 95% CI = [0.41,0.56]), again worse than the no information rate of 51.8%. To summarize, we find that low and high-performing entrepreneurs are equally willing to seek advice, ask comparable questions, and receive similar advice.

Entrepreneurs from both groups also appear to be influenced by the content of the AI generated advice they receive. As detailed [K](#), when one of the 12,762 non-stop words from the full corpus of answer text is part of an entrepreneur’s own advice text the word is 4.18 percentage points ($p < 0.0001$, 95% CI = [3.39; 4.96]) more likely to also be used in that entrepreneur’s business changes text ([Table A21](#), Model A1). Crucially, because the AI’s answers are only connected to the entrepreneur through the content of the questions they ask, we can use the “backdoor criterion” to isolate the causal effect of a word appearing in the answer text simply by controlling for whether the word appears in the entrepreneur’s questions ([Pearl \(2009\)](#); [Figure A11](#)). When doing so we estimate that words appearing in the AI advice are 2.84% points ($p < 0.0001$, 95% CI = [2.345; 3.327]) more likely to be used in the entrepreneur’s business changes text ([Table A21](#) Model A2). While this approach has the benefits of simplicity, these word-level estimates implicitly rest on a strong “bag of words” model ([Grimmer and Stewart, 2013](#)). To relax this assumption, we swap the 12,762 words for the 3,072 elements of the embedding vectors estimated on each entrepreneur’s questions, answers, and business changes text. [Table A21](#), Panel B shows that even after

controlling for the embedding value from the questions text, a standard deviation increase in an element’s embedding value in the answer text leads to a 0.127 s.d. increase for that element in the business changes text ($p < 0.0001$, 95% CI = [0.1073; 0.1475]). Whether we use words or embeddings to test for the influence of advice on the answer text, [Figure A12](#) shows that both low and high performers appear to be influenced by the content of the AI’s advice. Given that the changes made by the entrepreneur are responsive to the content of the advice, it does not appear that entrepreneurs are merely ignoring what the AI suggests.

Given that an entrepreneur’s business changes are influenced by the AI’s suggestions, the business changes of treated entrepreneurs—irrespective of whether they are low or high performers—should differ from the control. To test for the presence of such differences, we again train a model but this time to predict treatment status using word embedding representing each entrepreneur’s business changes text. Model 4 in [Table A19](#) shows our accuracy in the test set is 61.7%, significantly larger than the no information rate of 50.5% ($p = 0.013$, 95% CI = [0.52, 0.71]). Moreover, these treatment differences appear to be different for treated versus control low and high performers (Model 5 [Table A19](#)). Training a model to predict if an entrepreneur is a treated low performer, treated high performer, control high performer, or a treated high performer yields an accuracy of 45.3%, substantially greater than the no information rate of 25.5% ($p < 0.0001$, 95% CI = [0.36, 0.55]).⁵ The model predicts equally well across the four classes (McNemar’s p -value= 0.49). While low and high-performing entrepreneurs both send the AI assistant 18 messages and receive roughly 25 pieces of advice focused on similar topics, they appear to be converting this AI-generated advice into different types of business changes relative to their respective control groups.

To understand how our treatment impacted the kinds of changes made by low and high performers, we find example pairs of actual business changes text whose differences are similar to each group’s average treatment effect embedding vector. To calculate the average treatment effect embedding vector we simply difference the average of the treated and control embedding vectors in each group. To illustrate the types of semantic changes these vectors corresponds to, we then find pairs of control and treated text whose differences are cosine similar to the average treatment effect embedding vectors in each group. [Appendix K](#)

⁵The no information rate is lower because we are now predicting four instead of two classes.

provides additional details for how we leveraged word embedding methods to generate *qualitative* evidence on how our treatment impacted the types of business changes made.

Table 2 shows text from pairs that are in the top percentile of cosine similarity to the treatment effect embedding vector for high performers. The left column shows an abridged example from the control, the right column abridged from the treated, and the qualitative textual differences between the columns reflect the average shift in semantic embedding space captured by the average treatment effect embedding (Table A24 shows the full text). Scanning the table reveals no obvious differences in broad topic, length, or writing style. However, while the control text never mentions the AI mentor and rarely discusses it, the treated group examples mention how they worked with the AI to discover tailored and specific improvements to their businesses (highlighted text). Crucially, as discussed in Appendix K, when describing their business changes entrepreneurs were *not* reminded or prompted in anyway to mention the AI assistant. However, unprompted, treated high performers describe how they worked with the AI to identify and implement tailored changes to their businesses, with examples ranging from getting “advice from AI mentor on other ways of getting power when there is electricity power black out.” to “renting out gaming accessories in my cyber cafe following advice from your AI.” As these two examples show, there appears to be no single mechanism (e.g., bookkeeping), a pattern consistent with Figure A8 which shows that there are no clear shifts along any of our pre-registered mechanism checks. Instead, reviewing Table 2 suggests that high-performing entrepreneurs are not simply implementing generic undifferentiated changes, but are instead working with the AI to identify specific and more differentiated improvements that can tend to be correlated with stronger business performance (Guzman and Li, 2023; Carlson, 2023).

By way of comparison, Table 3 shows text from pairs in the top percentile of cosine similarity to the treatment effect embedding vector for low performers (Table A25 shows the full text). We see no mention of working with or learning from the AI assistant. Instead, for low performers, the treated text appears much more likely to mention lowering prices, offering discounts, and investing in advertising (highlighted text). Given that high and low performers ask similar questions and receive similar types of advice from the AI, treated low performers appear to be selecting and implementing relatively undifferentiated pricing and

advertising advice that fails to help their specific business needs (Carlson, 2023). In fact, if the AI leads entrepreneurs to lower prices, offer discounts, or invest in advertising—but these efforts fail to increase quantities sold—revenues and profits will fall relative to the control.

To quantitatively test the qualitative insights, we turn to basic word matching and counting techniques. To test if the treatment shifted whether entrepreneurs are more likely to work with the AI, we flag each business changes text for whether it matches keywords that are likely to refer to the AI assistant. To test for the impact on the specificity of business changes, we count the number of words in each entrepreneur’s business changes text that are used by five or fewer entrepreneurs. These uncommon words appear to be much more tailored to the specific of the business with examples including “150cc,” “arcade,” “biodegradable,” “detergent,” “headphones,” “peanut,” “rabbit” and “uganda.” Common words are much more generic and include tokens like “advertising,” “customers,” “market,” “price,” and “vegetables.” To test for differences in whether the business changes focus on lowering prices or investing in advertising, we flag each business change text based on whether it matches to common words used to describe discounting and advertising. Both sets of keywords, and our word-counting approach, are detailed in [Appendix K](#).

[Figure 5](#) Panel A shows that the treated entrepreneurs are 6.3 percentage points more likely to describe working with, learning from, and thanking the AI assistant when describing business changes ($p < 0.0001$, 95% CI = [3.336; 9.188]). For high performers, the effect is 10.8 percentage points ($p < 0.001$, 95% CI = [5.623; 15.90]), whereas for low performers, it is 1.8 ($p = 0.19$, 95% CI = [-0.8866; 4.573]) with the difference being statistically significant ($p < 0.01$). Relative to the base rate of 0.64 percentage points, the effect for high performers is roughly 15 times as large.⁶ Consistent with the qualitative observation that high performers use the AI to make tailored and specific changes to their businesses, [Figure B](#) shows that treated high performers use 20% more uncommon words relative to high performers in the control group ($p < 0.05$, 95% CI = [1.04; 1.36]). If anything, low performers use somewhat fewer uncommon words relative to the control, though the estimate is noisy. Finally, Panel

⁶Consistent with our arguments that the PDF serves as a placebo, we find that only one of the entrepreneurs in our study ever mention the business guide when describing the changes they made. [Table A22](#) Panel B shows regressions illustrating that essentially no entrepreneurs thank or mention learning from the ILO guide.

C shows that treated low performers appear to lower their prices or invest in advertising at higher rates. While control low performers mention discounting or advertising 39% of the time, treated low performers are 11.8 percentage points likely to do so ($p < 0.05$ 95% CI = [0.8775; 22.78]). Relative to the base rate, low performers are 30% more likely to discount products or invest in ads. Despite receiving similar suggestions from the AI, high performers appear better able to work with the AI to select and implement ideas more tailored to their needs, whereas low performers select more generic suggestions like discounting and advertising that tend to harm business performance.

The qualitative patterns in [Table 2](#) and [Table 3](#), along with the quantitative results in [Figure 5](#), also help to reconcile our inequality-increasing heterogeneous treatment effects with recent results that suggest generative AI reduces differences in productivity (See [Table A2](#)). In prior work, the task space is tightly constrained resulting in the AI-produced advice that is on either mostly useful or largely detrimental ([Dell’Acqua et al., 2023](#)). In our case, the entrepreneurs’ questions result in a large set of advice focused on a wide array of tasks from which the entrepreneurs must select a few to implement. Low performers appear to struggle to determine which of the suggested tasks would benefit them, instead implementing off-the-shelf advice like discounting prices or investing in advertising that leads to worse performance. High performers appear better at selecting advice and implementing tasks that can benefit their specific business needs. [Figure 6](#) illustrates this logic: while prior studies have focused on domains where the AI’s suggestions are either all bad (red boxes) or all good (green boxes), the open-ended nature of the problem we study likely implies a much wider distribution of potential impact, from which entrepreneurs must have the judgement and skills to select the good from bad advice ([Agrawal, Gans, and Stern, 2021](#)).

4 Discussion

Our study is—to our knowledge—the first randomized control trial testing the impact of generative AI focused on people and firms in developing economies. Many people in developing economies face barriers to getting helpful personalized feedback, mentorship, and training ([Chatterji et al., 2019](#); [Dimitriadis and Koning, 2024](#); [Björkegren, 2023](#)). While researchers

have raised concerns that LLMs may be less effective when used to address problems in developing economies due to the dominance of AI training data from the U.S. and other developed economies (Tao et al., 2023), our findings show that generative AI-based tools can serve as an efficient source of personalized feedback for many entrepreneurs in low- and middle-income countries. In fact, the treatment effect we find for high performers is similar to the 10% to 15% performance improvements driven by human-to-human training programs (Brooks, Donovan, and Johnson, 2018; McKenzie, 2021), but are likely delivered at a fraction of the cost.⁷ That being said, we also find evidence that generative AI can widen the gap between low- and high-performing businesses in these contexts. More broadly, our findings are consistent with the idea that generative AI has the potential—if designed and deployed with care—to benefit the billions of people and millions of firms in developing countries (Björkegren, 2023; Lou, Sun, and Sun, 2023).

Our experiment is not without limitations. First, the AI assistant that we evaluate is built on top of GPT-4 with only a relatively short system prompt and was distributed to entrepreneurs along with only a 5-minute-long online training survey. It is possible that by fine-tuning the language model using a corpus of relevant information, by iterating further on the system prompt provided to the model, and/or by providing more training to the entrepreneurs interacting with the AI assistant, we might be able to further improve the positive impacts we observe among high performers and reduce (or reverse) the adverse effects that we observe among low performers. Given that interactions with the AI assistant are fully observed, qualitative and quantitative analysis of the questions asked, topics discussed, and answers provided should enable researchers and managers to quickly iterate and improve on these AI systems (Choi et al., 2023).

Second, although the five-month-long duration of our experiment (three months pre-treatment, two months post-treatment) is considerably longer than the overwhelming majority of prior experiments measuring the economic impacts of generative AI (cf. Brynjolfsson, Li, and Raymond, 2023), there may still be longer-term impacts of exposure to the AI assistant that we do not observe. Finally, our experiment took place from May to November 2023,

⁷Similar to recent work highlighting the cost-effectiveness of text message interventions (Fabregas et al., 2024), our AI assistant costs a few dollars per participant whereas most training programs cost hundreds of dollars (USD) per participant (McKenzie, 2021).

a period of time during which exposure to and experience with generative AI tools such as ChatGPT was still low, particularly in developing market contexts such as Kenya. As both business owners and workers become more familiar with generative AI tools, experiments such as ours may yield different results because treated participants, control participants, and even competing businesses outside the experiment may all be using AI tools (Raymond, 2023, e.g.).

Our findings also suggest various opportunities for further research. As previously mentioned, the fact that our AI assistant may have had positive impacts for even a subset of entrepreneurs suggests that generative AI-based tools merit further study as a low-cost alternative to other interventions that have typically been studied in developing economies. One important piece of this research agenda is better understanding the mechanisms that drive entrepreneurs, firms, and workers to select and implement different types of advice; while we find that low-performing entrepreneurs implement advice that harms performance, we do not know if we can improve their choices or if the negative effects stem from more fundamental differences in skills or capabilities. Similarly, we find that some entrepreneurs use AI despite the presence of negative returns, yet our analysis sheds little light on why this is the case. Relatedly, it’s possible that the AI assistant’s heterogeneous impacts are at least in part attributable to biases in GPT-4’s training data, and that a model trained on better data could yield positive AI returns for even the most challenging business tasks (Tao et al., 2023). Finally, the AI assistant in our experiment does not learn, nor have we explored how the AI assistant impacted entrepreneurial learning. In theory, the data from our experiment could be used to improve the AI assistant’s recommendations, opening up the possibility not only of better mentors, but also of strategic concerns around asking for advice from an AI tool that may also be used by a competitor (Brynjolfsson, Li, and Raymond, 2023). How entrepreneurs will change their ideas, skills, and judgment in the presence of generative AI tools also remains an open question (Lebovitz, Lifshitz-Assaf, and Levina, 2022; Boussioux et al., 2023; Girotra et al., 2023; Hui, Reshef, and Zhou, 2023).

While there is still much to learn, our field experiment demonstrates the value of studying the impact of AI within the actual open-ended day-to-day decisions and tasks that take place within businesses (Harrison and List, 2004). That our treatment effects differ from

prior estimates on the impact of AI provides further evidence of the value of measuring the impact of interventions through real-world field experiments, above and beyond measuring the impact of those same interventions in controlled settings ([Aral and Walker, 2012](#)). As our findings show, generative AI is already capable of impacting real, open-ended, and unstructured economic outcomes, though in uneven and unexpected ways.

References

- Agrawal, Ajay, Joshua S Gans, and Avi Goldfarb. 2023. “Do we want less automation?” *Science* 381 (6654):155–158.
- Agrawal, Ajay, Joshua S Gans, and Scott Stern. 2021. “Enabling entrepreneurial choice.” *Management Science* 67 (9):5510–5524.
- Anderson, Stephen J, Christy Lazicky, and Bilal Zia. 2021. “Measuring the unmeasured: Aggregating, anchoring, and adjusting to estimate small business performance.” *Journal of Development Economics* 151:102655.
- Anderson, Stephen J and David McKenzie. 2022. “Improving business practices and the boundary of the entrepreneur: A randomized experiment comparing training, consulting, insourcing, and outsourcing.” *Journal of Political Economy* 130 (1):157–209.
- Aral, Sinan and Dylan Walker. 2012. “Identifying influential and susceptible members of social networks.” *Science* 337 (6092):337–341.
- Asiedu, Edward, Dean Karlan, Monica Lambon-Quayefio, and Christopher Udry. 2021. “A call for structured ethics appendices in social science papers.” *Proceedings of the National Academy of Sciences* 118 (29):e2024570118.
- Autor, David. 2014. “Polanyi’s paradox and the shape of employment growth.” NBER Working Paper 20485.
- Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. “A multifaceted program causes lasting progress for the very poor: Evidence from six countries.” *Science* 348 (6236):1260799.
- Baxter, Kathy M and Yoav Schlesinger. 2023. *How Companies Can Build Trustworthy AI Assistants*. Harvard Business Review Press.
- Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. “High-dimensional methods and inference on structural and treatment effects.” *Journal of Economic Perspectives* 28 (2):29–50.
- Björkegren, Daniel. 2023. “Artificial Intelligence for the Poor: How to Harness the Power of AI in the Developing World.” Foreign Affairs.
- Boussioux, Leonard, Jacqueline N Lane, Miaomiao Zhang, Vladimir Jacimovic, and Karim R Lakhani. 2023. “The Crowdless Future? How Generative AI Is Shaping the Future of Human Crowdsourcing.” *The Crowdless Future* .
- Brooks, Wyatt, Kevin Donovan, and Terence R Johnson. 2018. “Mentors or teachers? Microenterprise training in Kenya.” *American Economic Journal: Applied Economics* 10 (4):196–221.
- Bryan, Kevin A, András Tilcsik, and Brooklynn Zhu. 2017. “Which entrepreneurs are coachable and why?” *American Economic Review* 107 (5):312–316.

- Brynjolfsson, Erik and Lorin M Hitt. 2000. “Beyond computation: Information technology, organizational transformation and business performance.” *Journal of Economic Perspectives* 14 (4):23–48.
- Brynjolfsson, Erik, Danielle Li, and Lindsey R Raymond. 2023. “Generative AI at work.” NBER Working Paper 31161.
- Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2021. “The productivity J-curve: How intangibles complement general purpose technologies.” *American Economic Journal: Macroeconomics* 13 (1):333–372.
- Camuffo, Arnaldo, Alessandro Cordova, Alfonso Gambardella, and Chiara Spina. 2020. “A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial.” *Management Science* 66 (2):564–586.
- Carlson, Natalie A. 2023. “Differentiation in microenterprises.” *Strategic Management Journal* 44 (5):1141–1167.
- Chandler, Alfred D. 1977. *The visible hand*. Harvard university press.
- Chatterji, Aaron, Solène Delecourt, Sharique Hasan, and Rembrand Koning. 2019. “When does advice impact startup performance?” *Strategic Management Journal* 40 (3):331–356.
- Chen, Jiafeng and Jonathan Roth. 2023. “Logs with zeros? Some problems and solutions.” ArXiv: 2212.06080.
- Chen, Zenan and Jason Chan. 2023. “Large language model in creative work: The role of collaboration modality and user expertise.” *Available at SSRN 4575598* .
- Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Ivan Fernandez-Val. 2018. “Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India.” Tech. rep., National Bureau of Economic Research.
- Choi, Jun Ho, Oliver Garrod, Paul Atherton, Andrew Joyce-Gibbons, Miriam Mason-Sesay, and Daniel Björkegren. 2023. “Are LLMs Useful in the Poorest Schools? theTeacherAI in Sierra Leone.” *arXiv preprint arXiv:2310.02982* .
- Cusolito, Ana P, Ornella Darova, and David McKenzie. 2023. “Capacity building as a route to export market expansion: A six-country experiment in the Western Balkans.” *Journal of International Economics* 144:103794.
- Davies, Elwyn, Peter Deffebach, Leonardo Iacovone, and David Mckenzie. 2024. “Training microentrepreneurs over Zoom: Experimental evidence from Mexico.” *Journal of Development Economics* 167:103244.
- Dell’Acqua, Fabrizio, Edward McFowland, Ethan R Mollick, Hila Lifshitz-Assaf, Katherine Kellogg, Saran Rajendran, Lisa Kraymer, François Candelon, and Karim R Lakhani. 2023. “Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality.” SSRN 4573321.

- Dimitriadis, Stefan and Rembrand Koning. 2022. “Social skills improve business performance: Evidence from a randomized control trial with entrepreneurs in Togo.” *Management Science* 68 (12):8635–8657.
- . 2024. “Networking Frictions: Evidence from entrepreneurial networking events in Lomé.” *Management Science* (forthcoming).
- Doshi, Anil R and Oliver Hauser. 2023. “Generative artificial intelligence enhances creativity.” Available at SSRN 4535536.
- Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock. 2024. “GPTs are GPTs: Labor market impact potential of LLMs.” *Science*.
- Fabregas, Raissa, Michael Kremer, Matthew Lowes, Robert On, and Giulia Zane. 2024. “Digital information provision and behavior change: Lessons from six experiments in East Africa.” Tech. rep., National Bureau of Economic Research.
- Girotra, Karan, Lennart Meincke, Christian Terwiesch, and Karl T Ulrich. 2023. “Ideas are dimes a dozen: Large language models for idea generation in innovation.” Available at SSRN 4526071.
- Glennerster, Rachel and Kudzai Takavarasha. 2014. *Running randomized evaluations: A practical guide*. Princeton University Press.
- Goldfarb, Avi. 2024. “Pause artificial intelligence research? Understanding AI policy challenges.” *Canadian Journal of Economics/Revue canadienne d’économie* .
- Grimmer, Justin and Brandon M Stewart. 2013. “Text as data: The promise and pitfalls of automatic content analysis methods for political texts.” *Political analysis* 21 (3):267–297.
- Guzman, Jorge and Aishen Li. 2023. “Measuring founding strategy.” *Management Science* 69 (1):101–118.
- Harrison, Glenn W and John A List. 2004. “Field experiments.” *Journal of Economic Literature* 42 (4):1009–1055.
- Holtz, David, Felipe Lobel, Ruben Lobel, Inessa Liskovich, and Sinan Aral. 2023. “Reducing interference bias in online marketplace experiments using cluster randomization: Evidence from a pricing meta-experiment on airbnb.” *Management Science* .
- Hu, Krystal. 2023. “ChatGPT sets record for fastest-growing user base - analyst note — reuters.com.” URL <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>. [Accessed 21-11-2023].
- Hui, Xiang, Oren Reshef, and Luofeng Zhou. 2023. “The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market.” SSRN 4527336.
- International Labor Organization. 2015. “SIYB Entrepreneur Level Materials.” URL https://www.ilo.org/empent/areas/start-and-improve-your-business/WCMS_192062/lang--en/index.htm.

- Ji, Ziwei, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. 2023. “Survey of hallucination in natural language generation.” *ACM Computing Surveys* 55 (12):1–38.
- Karlan, Dean, Ryan Knight, and Christopher Udry. 2015. “Consulting and capital experiments with microenterprise tailors in Ghana.” *Journal of Economic Behavior & Organization* 118:281–302.
- Kim, Hyunjin. 2024. “The value of competitor information: Evidence from a field experiment.” *Management Science* (forthcoming).
- King, Andrew, Brent Goldfarb, and Timothy Simcoe. 2021. “Learning from testimony on quantitative research in management.” *Academy of Management Review* 46 (3):465–488.
- Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2007. “Experimental analysis of neighborhood effects.” *Econometrica* 75 (1):83–119.
- Kumar, Harsh, David M Rothschild, Daniel G Goldstein, and Jake M Hofman. 2023. “Math Education with Large Language Models: Peril or Promise?” Available at SSRN 4641653.
- Lakens, Daniël. 2017. “Equivalence tests: A practical primer for t tests, correlations, and meta-analyses.” *Social psychological and personality science* 8 (4):355–362.
- Lebovitz, Sarah, Hila Lifshitz-Assaf, and Natalia Levina. 2022. “To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis.” *Organization science* 33 (1):126–148.
- Lou, Bowen, Hongshen Sun, and Tianshu Sun. 2023. “GPTs and labor markets in the developing economy: Evidence from China.” Available at SSRN 4426461 .
- McAfee, A, D Rock, and E Brynjolfsson. 2023. “How to Capitalize on Generative AI.” *Harvard Business Review* 101 (6):42–48.
- McKenzie, David. 2012. “Beyond baseline and follow-up: The case for more T in experiments.” *Journal of Development Economics* 99 (2):210–221.
- . 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.
- . 2023. “Interpreting treatment effects on an inverse hyperbolic sine outcome variable and alternatives.” *Development Impact* Accessed: 2023-11-29.
- McKenzie, David and Christopher Woodruff. 2017. “Business practices in small firms in developing countries.” *Management Science* 63 (9):2967–2981.
- Mehmood, Muhammad Zia. 2023. “Short Messages Fall Short for Micro-Entrepreneurs: Experimental Evidence from Kenya.” .
- Mollick, Ethan. 2022. “ChatGPT is a tipping point for AI.” *Harvard Business Review* .
- Mollick, Ethan R and Lilach Mollick. 2022. “New modes of learning enabled by ai chatbots: Three methods and assignments.” Available at SSRN 4300783 .

- Noy, Shaked and Whitney Zhang. 2023. “Experimental evidence on the productivity effects of generative artificial intelligence.” *Science* 381:187–192.
- OpenAI. 2023. “GPT-4 Technical Report.”
- Pearl, Judea. 2009. *Causality*. Cambridge university press.
- Peng, Sida, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer. 2023. “The Impact of AI on Developer Productivity: Evidence from Github Copilot.” *arXiv preprint:2302.06590* .
- Ramdas, Kamalini and Alp Sungu. 2022. “The digital lives of the poor: Entertainment traps and information isolation.” *Management Science*, *forthcoming* .
- Raymond, Lindsey. 2023. “The Market Effects of Algorithms.” Working Paper.
- Sato, Megan Kinniment Lucas Jun Koba, Haoxing Du, Brian Goodrich, Max Hasin, Lawrence Chan, Luke Harold Miles, Tao R Lin, Hjalmar Wijk, Joel Burget, Aaron Ho, Elizabeth Barnes, and Paul Christiano. 2023. “Evaluating Language-Model Agents on Realistic Autonomous Tasks.” arXiv: 2312.11671.
- Simonsohn, Uri. 2015. “Small telescopes: Detectability and the evaluation of replication results.” *Psychological science* 26 (5):559–569.
- Tao, Yan, Olga Viberg, Ryan S. Baker, and Rene F. Kizilcec. 2023. “Auditing and Mitigating Cultural Bias in LLMs.” arXiv: 2311.14096.
- The White House. 2022. ““The Impact of Artificial Intelligence on the Future of Workforces in the European Union and the United States of America.” Technical Report.
- The World Bank Group. 2018. “Enterprise Surveys: What Businesses Experience - Kenya 2018 Country Profile.” URL <https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/country-profiles/Kenya-2018.pdf>.
- Wamuyu, PK. 2020. “The Kenyan Social Media Landscape: Trends and Emerging Narratives, 2020.” SIMElab Working Paper.
- Wiles, Emma and John J Horton. 2024. “More, but Worse: The Impact of AI Writing Assistance on the Supply and Quality of Job Posts.” .
- Wiles, Emma, Zanele T Munyikwa, and John J Horton. 2023. “Algorithmic writing assistance on jobseekers’ resumes increases hires.” Tech. rep., National Bureau of Economic Research.
- Zizzo, Daniel John. 2010. “Experimenter demand effects in economic experiments.” *Experimental Economics* 13:75–98.

Figures

Figure 1: Example entrepreneur-AI interaction over WhatsApp.

I have fast food joint that I run just adjacent to a matatu terminus. There are several such food joints at that particular place and it's fair to say competition is stiff. How can I overcome the competitive business environment and be ahead of the rest of the pack?

Thank you for sharing about your business. Operating a fast food joint in a highly competitive environment like a matatu terminus can indeed be challenging, but there are several strategies you could consider to stand out:

1. **Differentiate Your Menu:** Offer unique or special dishes that aren't available at the other joints. This could be traditional Kenyan foods with a twist, or it could involve using locally sourced ingredients that others don't use. The benefit is that customers will associate your place with these special offerings and choose you over competitors.
2. **Improve Customer Service:** Prioritize excellent customer service to make people feel welcome and appreciated when they visit your joint. This might involve training staff on how to interact with customers positively, quickly serving orders, and handling any complaints promptly and professionally.
3. **Loyalty Program:** Implement a loyalty program where customers earn points for every purchase which they can later redeem for free meals or discounts. It encourages repeat business as customers will want to accumulate more points.
4. **Community Involvement:** Engage in local community events or sponsor local initiatives such as sports teams or school activities. This raises your profile within the community and makes people more likely to support your business.
5. **Cleanliness & Hygiene:** Ensure high standards of cleanliness both inside the kitchen and in the dining area because hygiene matters greatly in food businesses.

Enter a number to learn more about one of these points on "improving competitiveness" or ask a new question.

3

Great, let's delve into how you can implement a **Loyalty Program** in your fast food joint:

1. **Point-Based System:** Implement a simple point-based system where customers earn points for every shilling spent. Once they accumulate a certain number of points, they can redeem them for free meals or discounts. You could use simple tools like punch cards or digital apps that are inexpensive to manage.
2. **Special Offers for Regular Customers:** Identify your regular customers and offer them exclusive deals or discounts. This not only rewards their loyalty but also encourages them to keep coming back.
3. **Birthday/Anniversary Rewards:** Offer special treats or discounts to customers on their birthdays or anniversaries if they share this information with you. This personalized touch can make customers feel valued and appreciated.
4. **Referral Program:** Encourage your loyal customers to bring in new clients by giving them incentives such as a discount or free meal when someone they refer makes their first purchase.
5. **Feedback Reward:** Encourage customer feedback by offering small rewards for completing surveys about the quality of food and service at your joint.

Remember, the goal of any loyalty program is to make your customers feel valued so that they continue patronizing your business over others. Enter a number to learn more about one of these points on "loyalty programs" or ask a new question.

Notes: This figure presents a rendering of an entrepreneur-AI interaction in WhatsApp. The text is original (including the typos). A "matatu" is a word commonly used to describe a minibus in Kenya. In [Appendix I](#), we display the full conversation histories for two treated entrepreneurs.

Figure 2: Overview of the experimental timeline from the date of the first survey

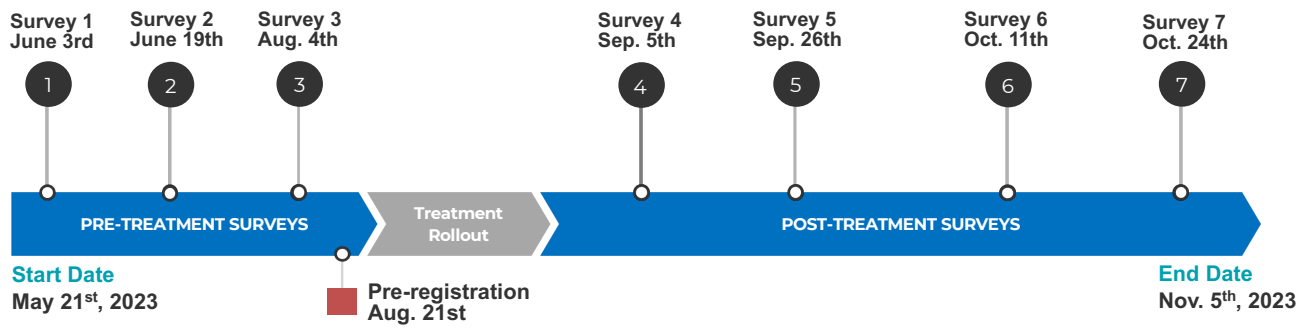
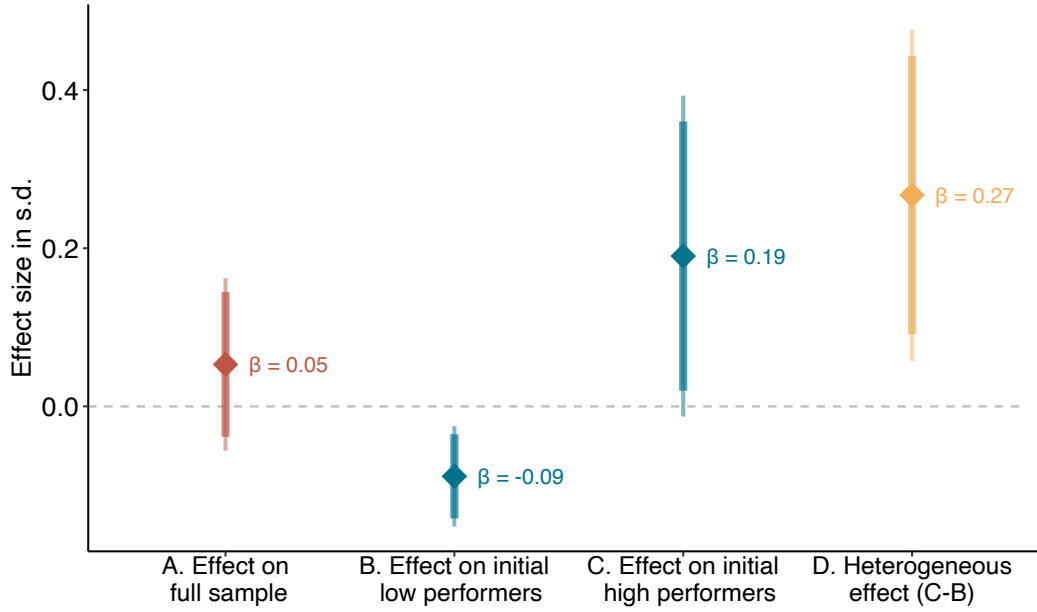
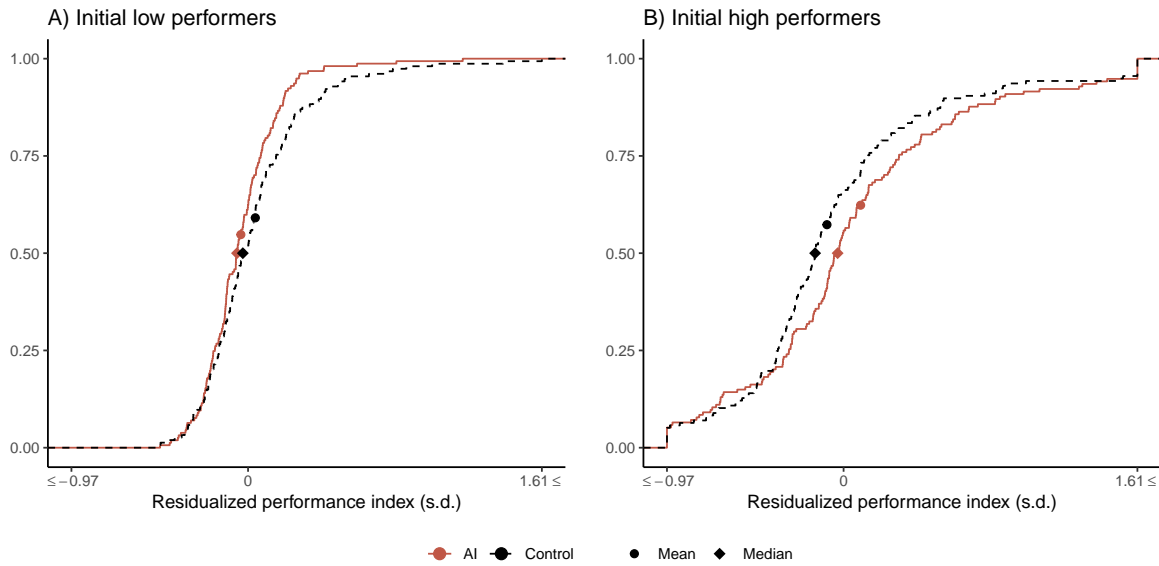


Figure 3: Treatment effects on business performance. Estimates reveal an overall null effect for our full sample of entrepreneurs, but a negative treatment effect for low performers (defined as having below-median pre-treatment performance), a positive treatment effect for high performers (above-median pre-treatment performance), and a significant heterogeneous treatment effect.



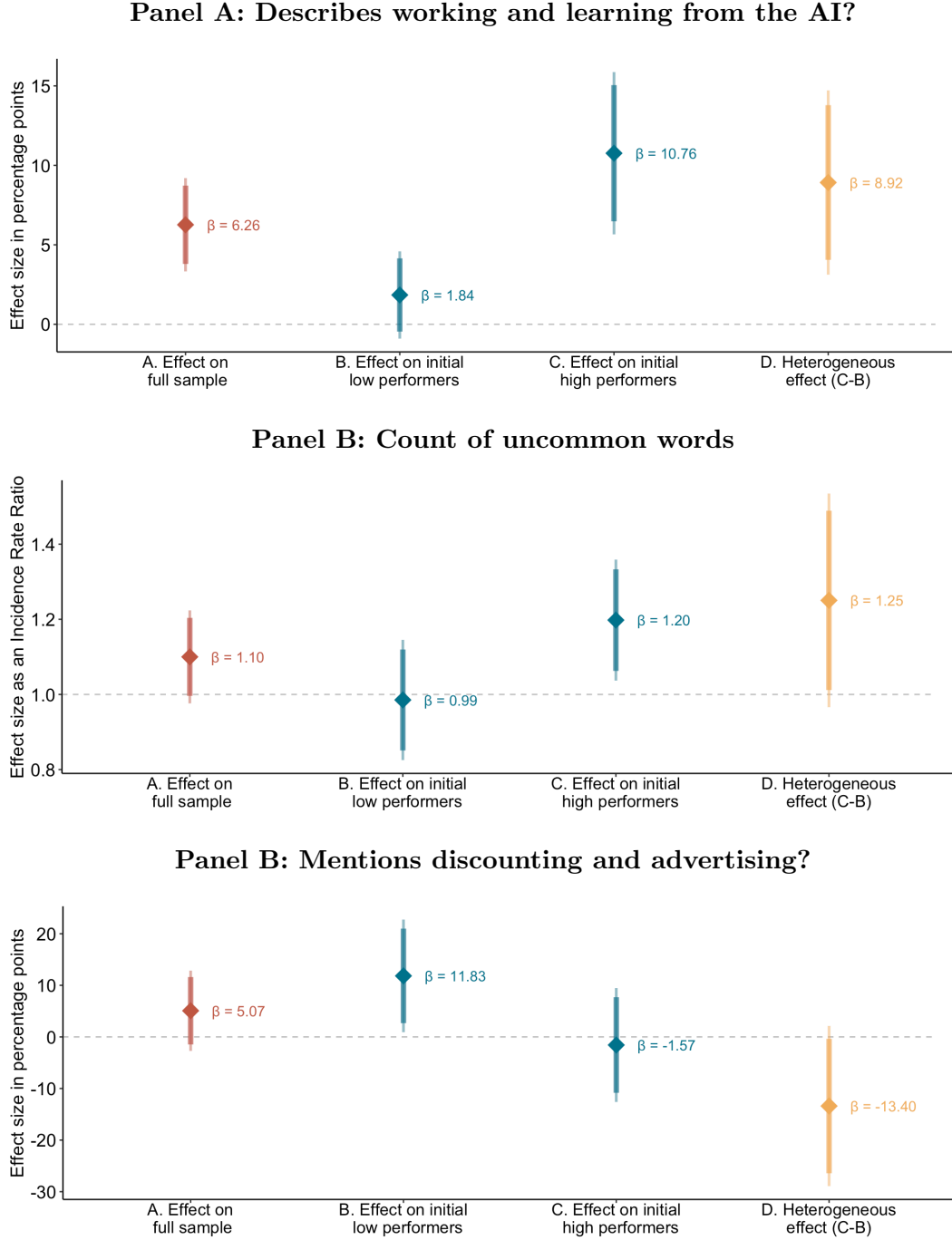
Notes: The effect of access to the AI assistant on our index of business performance winsorized at the 99% level, restricting to participants who have completed the panel. Outcomes are presented in standard deviations. Effects are estimated using a regression model that controls for pre-treatment outcome levels and baseline covariates selected using a double-LASSO procedure. Controls selected by the double-LASSO are report in [Table A5](#). Error bars present 90 and 95% confidence intervals. The presented p -value is from a two-tailed test of the difference in treatment effects between entrepreneurs with below versus above median pre-treatment performance.

Figure 4: The distribution of changes in performance shifts to the left for treated low performers, whereas it shifts to the right for treated high performers



Notes: Panels A and B present cumulative density plots of the standardized index of business performance winsorized at the 99% level. It compares those with initial low performance (below-median pre-treatment performance; left panel) to those with initial high performance (above-median pre-treatment performance; right panel), across both treatment and control groups. The performance index is residualized using the double-LASSO-selected covariates reported in [Table A5](#) and for stratification block and time period. All results are restricted to participants who have completed the panel.

Figure 5: Estimates showing the heterogeneous impact of the AI treatment on the types of business changes made by entrepreneurs. These panels quantitatively test the qualitative results in Table 2 and Table 3. Panel A shows that high performers are more likely to describe working and learning from the AI assistant. Panel B shows that high performers use more uncommon words that are tailored to their specific business needs. Panel C shows that low performers are more likely to implement generic and potentially costly strategies like lowering prices or investing in ads.



Notes: Error bars present 90 and 95% confidence intervals. Robust standard errors. The estimates in Panel A are from the linear probability models in Table A22, Panel B are the exponentiated coefficients from the Poisson regressions in Table A23, and Panel C from the linear probability models in??.

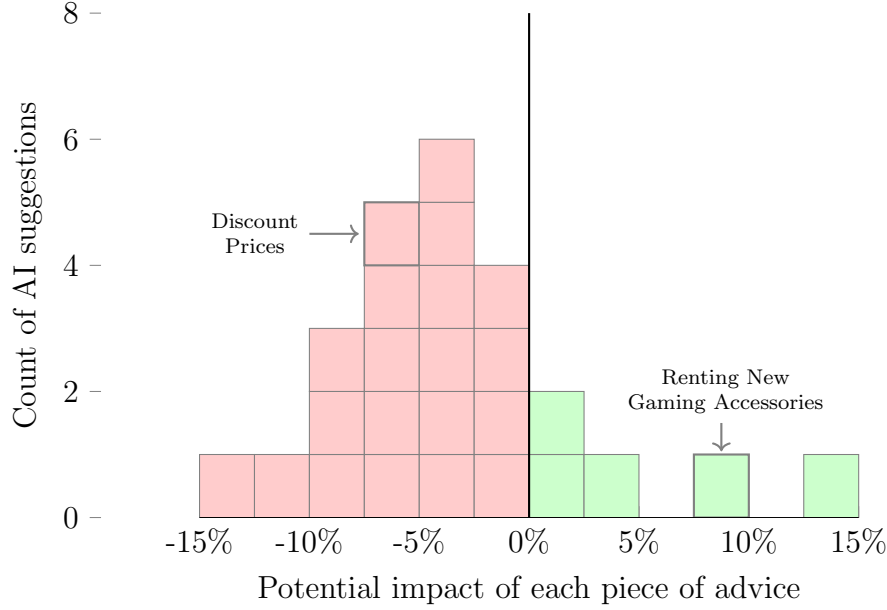


Figure 6: Stylized figure illustrating how heterogeneous treatment effects can stem from differences in the ability of entrepreneurs to select and implement suggestions even if they are drawing from a similar pool of AI generated advice. If most AI suggestions harm performance (e.g. dropping prices), but some advice can lead to meaningful improvements (e.g. a suggestion to try renting new gaming accessories for a cyber shop) then differences in the pieces of advice different types of entrepreneurs select and implement can lead to markedly different treatment effects. Prior experiments evaluating the impact of Generative AI on performance outcomes have largely truncated the advice distribution by focusing on well-defined tasks where the distribution of AI advice lies clearly to the left or the right of zero (e.g., [Brynjolfsson, Li, and Raymond, 2023](#); [Dell’Acqua et al., 2023](#)), thus limiting the need for user judgement and skill in selecting and implementing from AI generated advice.

Tables

Table 1: Distribution of 10 topics across questions sent to the AI assistant. The red bars show the percentage of questions from each topic for below median low performers. The blue bars show the percentage for above median high performers. [Appendix K](#) discusses the creation of these topics and question classification.







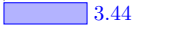



Topic with Example	Fraction from Topic
Entrepreneurial Opportunities & Business Ideas: “I want to start a cyber, printing and photocopy shop in Kericho town. What are some of the things I should know before I set up that kind of a business? I own a desktop and a reliable printer for start up, however I’ll upgrade the desktop later on according to the needs.”	
Livestock & Farm Management: “I’m planning to be a maize farmer in Kiambu county with a 2-acre plot. We’ve been having few problems from pests and disease attacks which are hurting other farmers’ crops and profits. How can I make more money without spending more?”	
Financial Management: “I’m a coffee vendor around Oyugis, I do walk with my coffee from place to place but the number of debtors is alarming. Am I being too lenient or how should I politely refuse to give one coffee on credit if they don’t pay after”	
Business Planning & Growth: “I am a new business starter planning to supply t-shirts in my school yet there are other suppliers of t-shirts in the school. How can I make my business successful with all the competition around?”	
Customer Acquisition & Retention: “I have a small auto motorcycle spares shop in a busy city. It is working well but lately due to the economy my customers are shying away from my business my competitors although prices and quality of the spares is the same. What do I need to do to get my groove back and bring my customers back?”	
Marketing Strategies: “I own a small hotel selling tea and food of about 10-15 customers a day, I am trying to find ways to promote my business I already displayed posters outside my kiosk, what are some other ways I can gain new customers?”	
Technology Integration & Digital Marketing: “I have a website for business. I want to integrate a payment system that automatically detects payments made and generates receipts. How and where can I get this service?”	
Inventory & Operations Management: “I sell household goods in Nakuru. I have lots of different items and it’s hard to keep good track of them. How can I manage my stock better?”	
Market Research & Analysis: “I have more customers on my shopping ground and they need more things to buy but I don’t know which things I can bring for them so that they can love to use them more, how can I get more profits?”	
Risk Management: “How do I deal with the issue of some of my hotel items getting lost in the course of the business? My staff claim they never take anything yet when I do my auditing, some of the items are nowhere to be traced.”	

Table 2: Example treated and control business changes text whose differences in embedding space are in the top percentile of cosine similarity to the **high** performer’s treatment effect embedding vector. Table A24 displays the full non-abridged text. Highlights corresponds to the qualitative differences we observe between the control and treated text.

Control Example	AI Treated Example
[...] I’ve started baking honey filled muffins. Honey baked muffins. Home delivery of my products. [...]	Use of new technology of AI mentor. By doing more benchmarking on the competitors. Cooking oil, cereals. [...]
Adding more products into the business. [...] Advertising what am selling. [...] Delivering door to door services.	Improved on customer care skills after getting some tips and advice from Almentor and introduction of discounts to customers who purchased the products in bulk. [...]
[...] I have started selling hair bonnets. [...] Delivery of products to customers has really made a difference in my business. Making myself available at any time for my customers. [...]	[...] I have inserted a small salon beside my cyber [...] I got advice from AI mentor on other ways of getting power when there is electricity power black out. [...]
[...] Cooking quality food. [...] Delivering my products to customers. Advertising my business through online. [...] Giving my customers carrier bags. [...] Selling coffee	Improving the quality of chicks in my chicken roster has brought high quality chicken to my folder... I ma grateful for the Chat AI. [...] I have bought new chicks that were recommended by whe WhatsApp AI. [...] Reducing slightly the prices of my commodities [...]
The use of new computers which has made my work really easy. I’ve been making my own products like soaps and detergents which is making profits for my business. I tried some new body oils. I’ve tried door to door delivery. [...]	I employed someone [...] I gave my customers special offers [...] I expanded my car wash, added a cyber a place for soft drinks thanks to the AI tool you gave me. Trying a new washing detergent. [...] customers can pay using mobile money. [...] deliveries of the washed cars to their owners.
[...] i introduced other kitchen ware sets like cups and plates. Delivery of goods to customers. Engaging clients. Use of media in advertising. Male clothes. [...]	[...] With the help of AI, I have been able to do more market research and discovered new ways of advertising my products. I have added cellphones, speakers and electric wires [...] I have started electronics repairs [...] I have started free delivery for Items worth 5000 and above.
Having own website which people can order and have goods delivered in their door steps. [...] Selling electronics. [...]	Buying in bulk saves input in resources. [...] Provide entertainment services for adults and children. Negotiating with suppliers [...] renting out gaming accessories in my cyber cafe’ following advice from your AI. I have started offering cashbacks to loyal customers. [...]
I have changed location [...] Advertising on social media. I started the perfule refill business. Providing education consultancy. Introducing a new product [...]	[...] I change the location [...] Technology. Marketing policy. Value addition in help of chat bot advice. [...] Familiarizing with customers and looking for demand curves.
Home delivery has made it easier for customers with a busy schedule. Interact with clients to know the trend that is in the market that they like. Started selling clothing. [...] Online business has really impacted my business positively [...] Installed CCTV in my business premise to help with security. [...] I offer cash/cashless money transfer. [...]	[...] I’ve brought ChatGPT on board cause of it’s immense importance especially with regards to awide array of information it offers. [...] Onboarding government digital services [...] increased customer traffic in my business premises [...] Offering first time customers products on discounted prices has really proved worthwhile. [...]

Table 3: Example treated and control business changes text whose differences in embedding space are in the top percentile of cosine similarity to the **low** performer’s treatment effect embedding vector. Table A24 displays the full non-abridged text. Highlights corresponds to the qualitative differences we observe between the control and treated text.

Control Example	AI Treated Example
Increase in profit in the business. Sales of eggs. Marketing my goods online. Selling of clothes. Transport service. [...]	Giving so offers by reducing the selling price. [...] Advertising on social media platforms. [...]
The new ways of watering has enabled most of my crops not to get dried up and hence selling more seedlings thus increasing the profits. [...] My new activities are by cooking and selling some of the products of my businesses. [...] Selling by means of digital or online. [...]	[...] I have advertising my products through WhatsApp and Facebook. Giving some extra discount on my customers and also giving some products offer to my customers increase the number of my customers. [...] Door delivery products for my customers who gives orders through phone. [...]
[...] Getting to know what customers wants, precise record keeping and online marketing. [...] Reduced owing to reduce debts. [...] Providing loan at an interest.	Advertisement and discounts. Offering after sales service. Training of staff. [...]
Budgeting and data keeping. [...] Increased profits because I can now manage cashflow appropriately without possible losses. Maize. Mpesa services. Fruit selling. [...]	Posting online has brought more customers. [...] Discount for items above 1000. Offers for goods above 1500. [...] Selling shoes in my mtumba warehouse. [...] Cereal outside my shop
Consulting my customers. grocery. [...] using of technology. selling of cosmetics products. [...]	Issuing discounts on some products. Painting my shop as advertising strategy. Discount. Packaging.
Collection nof used oil and reselling it. [...] Advertising the product. [...] Middleman in buying and selling of secondhand. Farm implements and tractors. Stocking the products. [...]	Introducing discounts. [...] Aftersales services. To satisfy customers and offer best services to them. [...] Opening of more branches. Wireless CCTV cameras. Free transportation [...]
[...] I have started selling materials for making vitenge. [...] Bodaboda service. [...] Including mitumbas in shoes selling. Selling credit. Started layering poultry	The free delivery for my products is really working out for me. [...] New Imperial leather for men. Hair dressing. [...] Rearing and selling chicken
Selling of milk is highly profitable and increasesing of products in general has made my business grow. [...]	Advertising the new products in my shop. Selling my products at a relatively lower prices and offering discounts to my loyal customers. [...] Considered customer feedback [...]
Increase in profit in the business. Sales of eggs. Marketing my goods online. Selling of clothes. Transport service. [...]	Giving customer offers. I’ve installed neon lights to advertise my business. [...] I have included Mpesa to my business to help customers do mobile phone transactions. [...]
Including some of the potential customers and suppliers in decision making, also acquiring a loan to expand the business. [...] Involving workers in decision making and also being consultant. [...]	The use of billboard for advertising my business [...] Offering dicounts to my customers. [...] I have tried using different hair foods and am good with my customers. [...]

Online Appendix

Table of Contents

A Appendix Figures	39
B Appendix Tables	51
C Experimental Design and Timeline	81
D Detailed Intervention Overview	85
E Pre-analysis Plan	92
E.1 Overview of Pre-registration	92
E.2 Econometric Strategy	93
E.3 Overview of Measures	94
E.4 Explanation of Preferred Functional Forms	98
F Comparison to existing literature	99
G Ethical Considerations	101
H Accounting for Spillovers	103
I Detailed Case Examples	105
J Generic Machine Learning for Heterogeneous Treatment Effects	119
K Text Analysis	121
K.1 Overview and Summary Statistics	122
K.2 Question and Answer Topic Analysis	123

K.3	Predicting Treatment and Performance Heterogeneity	132
K.4	Impact of Advice Text on Business Changes Text	133
K.5	Qualitative Analysis using Treatment Effect Embeddings	134
K.6	Quantitative tests of qualitative insights	135

A Appendix Figures

Figure A1: Recruitment ad



Figure A2: The distribution of study participants across Kenya

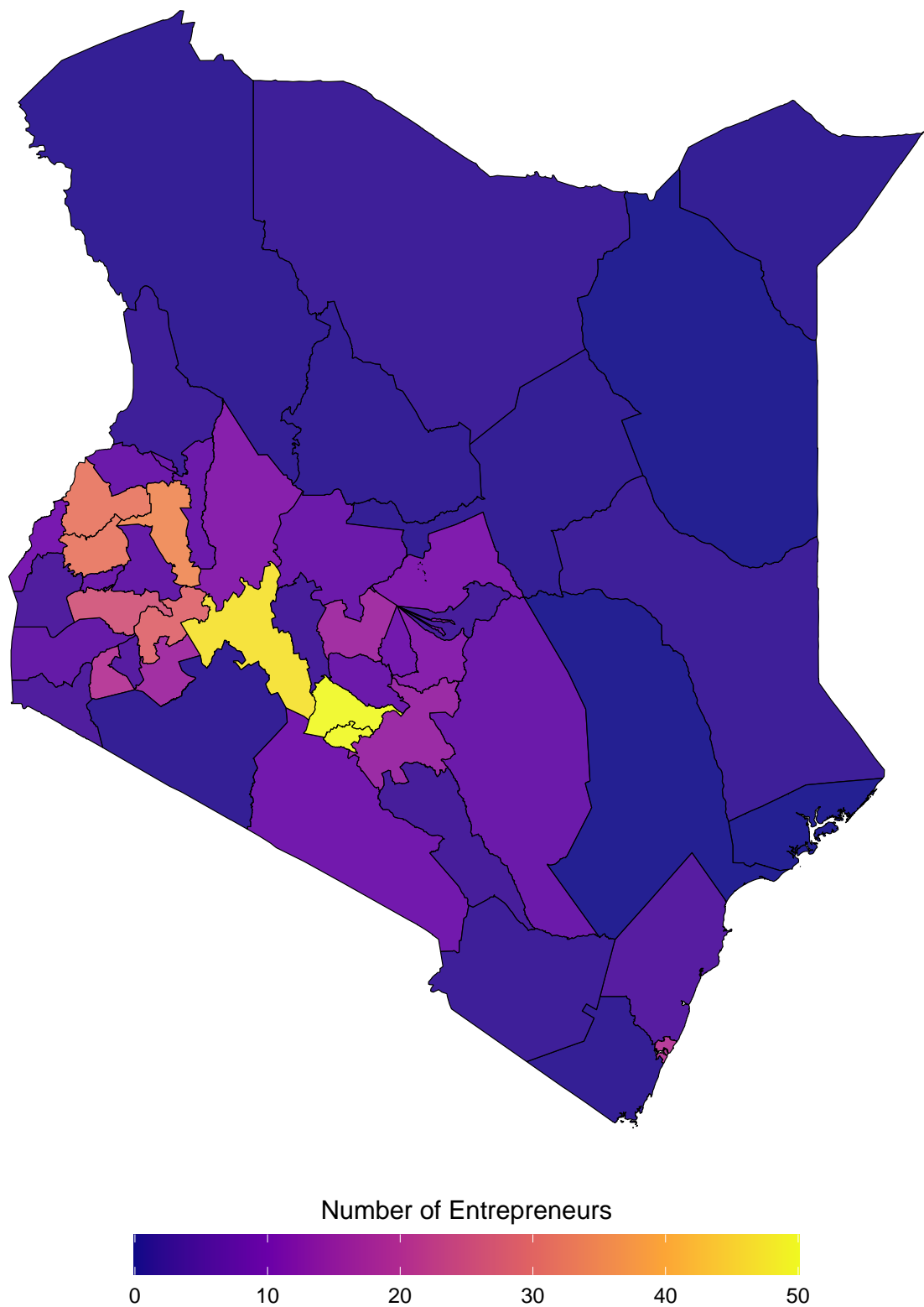
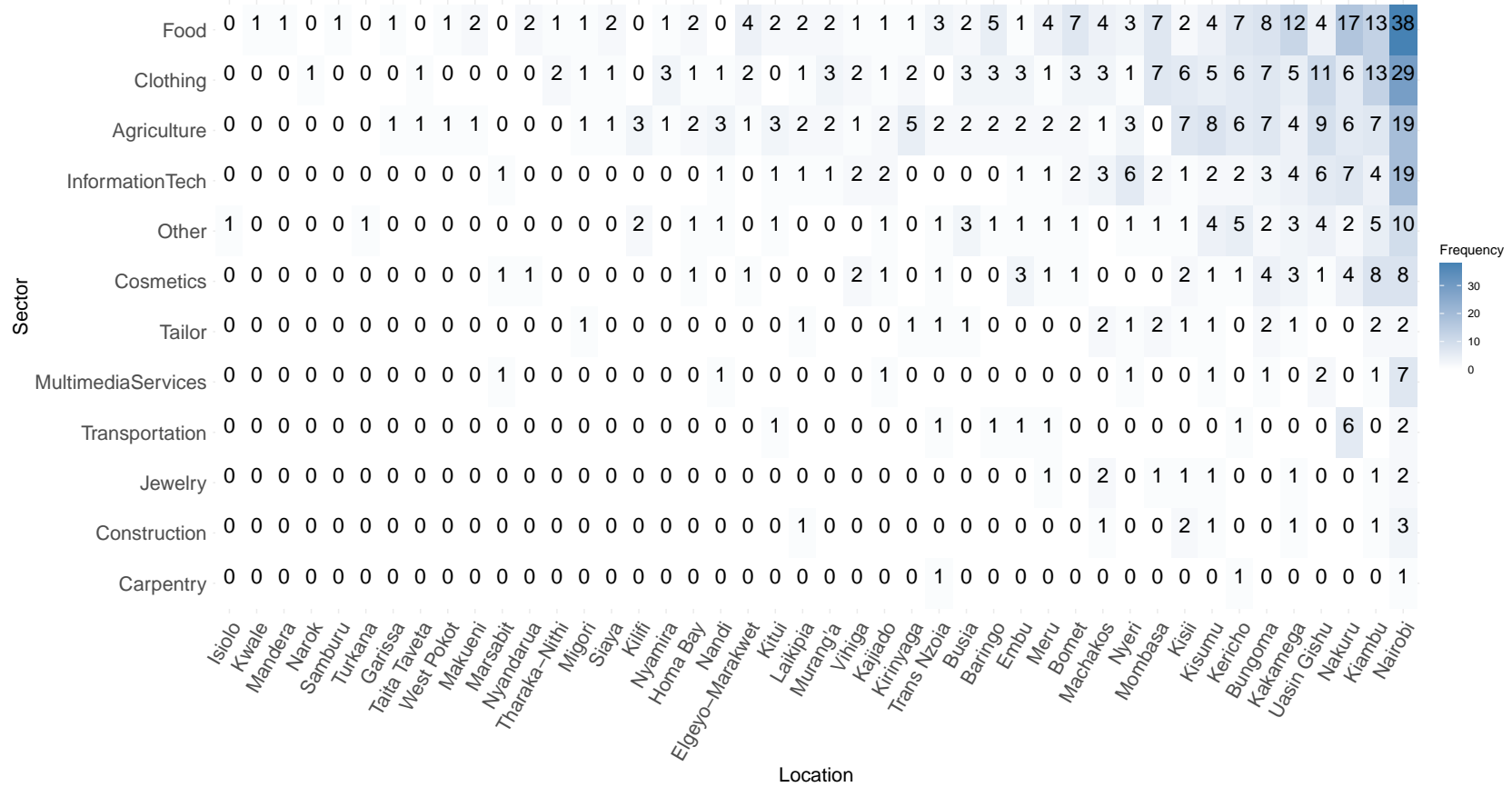


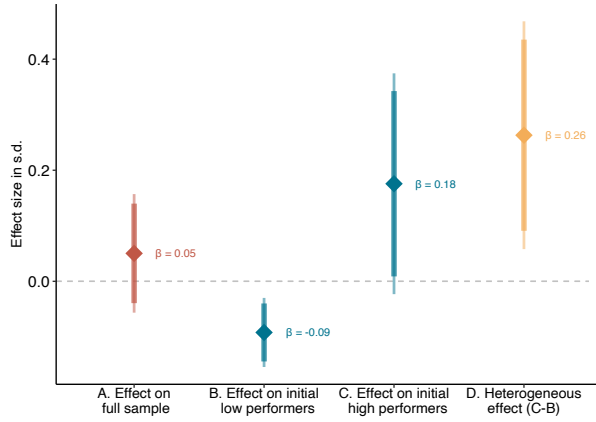
Figure A3: Distribution of Sectors and Kenyan Counties



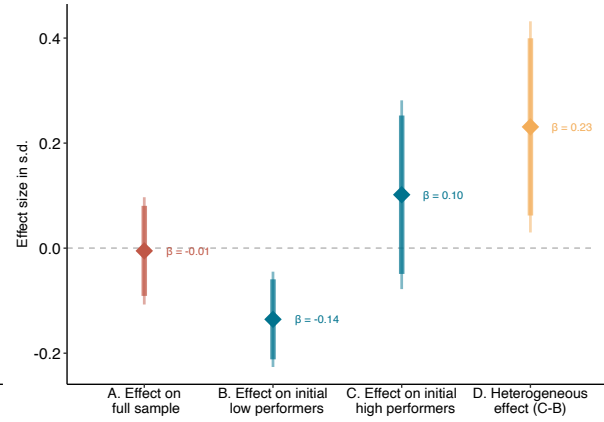
Notes: This figure depicts the distribution of Kenyan counties and firm sectors. Note that each participant may be involved in multiple sectors.

Figure A4: Performance effects (Robustness checks)

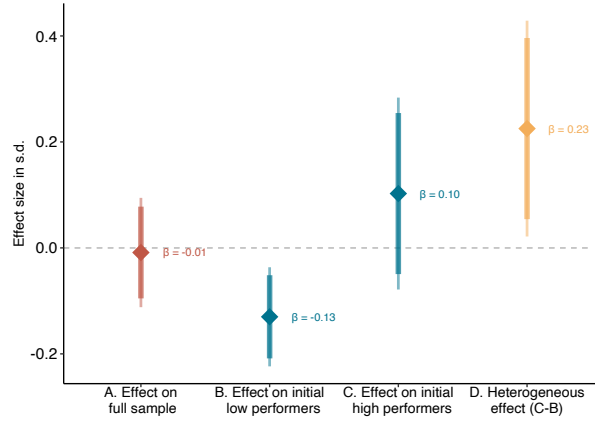
(A) Winsorized at the 99% level, all observations



(B) Winsorized at the 95% level, all observations

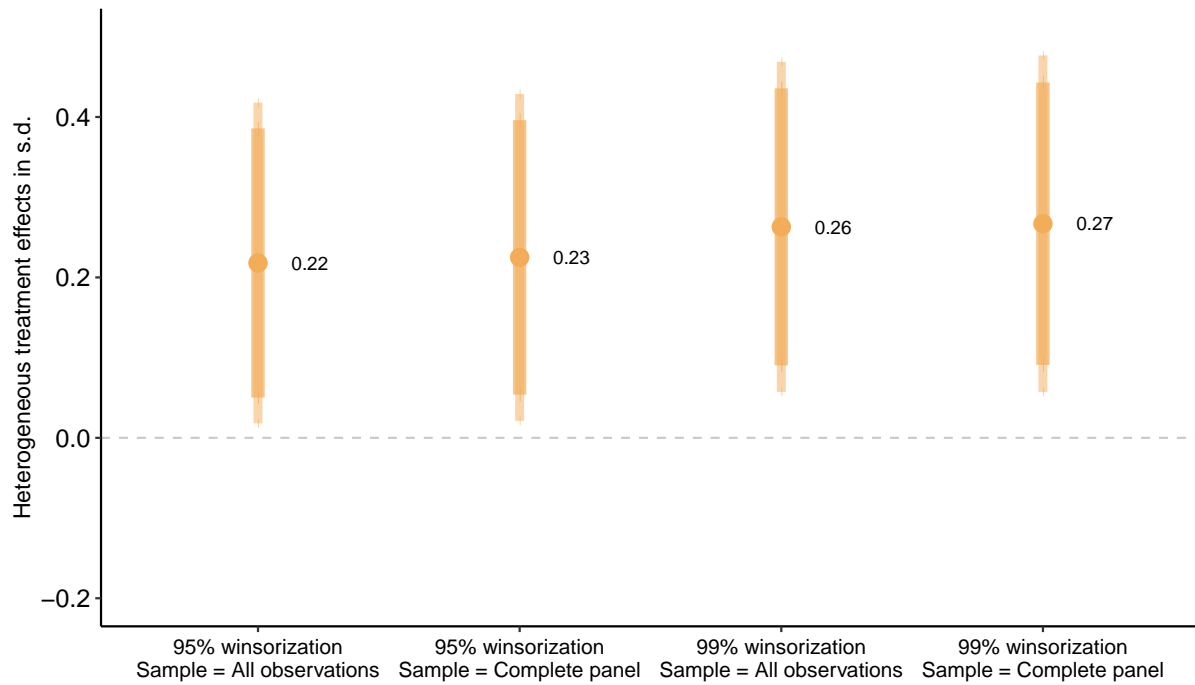


(C) Winsorized at the 95% level, completed panel



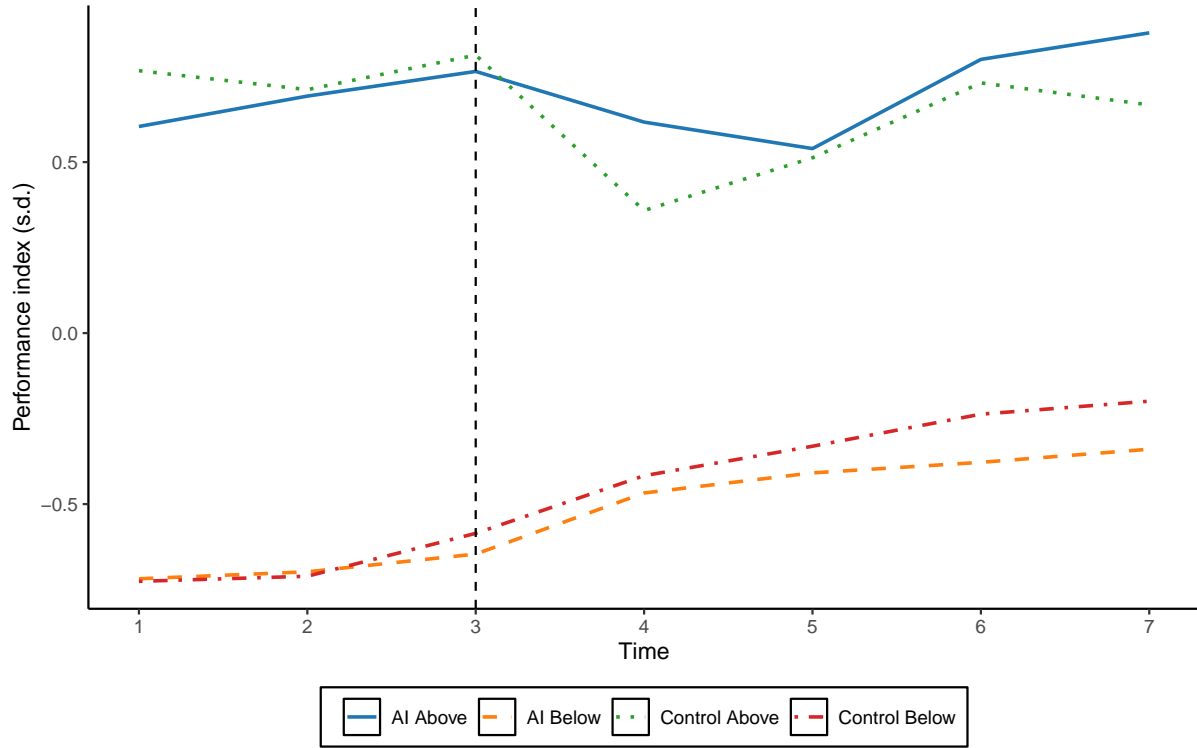
Notes: This figure presents the causal effect of access to the AI assistant on business performance (profits and revenues). Outcomes are presented in standard deviations. Panel (A) and (B) present results winsorized at the 99% and 95% respectively for all participants' observations. Panel (C) presents results winsorized at the 95% level restricting to individuals who have completed the panel. Effects are estimated using an intent-to-treat specification that controls for pre-treatment outcome levels and baseline covariates selected using a double LASSO procedure. The presented p -values are from a two-tailed test of the difference in treatment effects between entrepreneurs with below versus above median pre-treatment performance. All tests are two-tailed. Error bars present 95 and 90% confidence intervals.

Figure A5: Heterogeneous treatment effect estimates for performance specifications



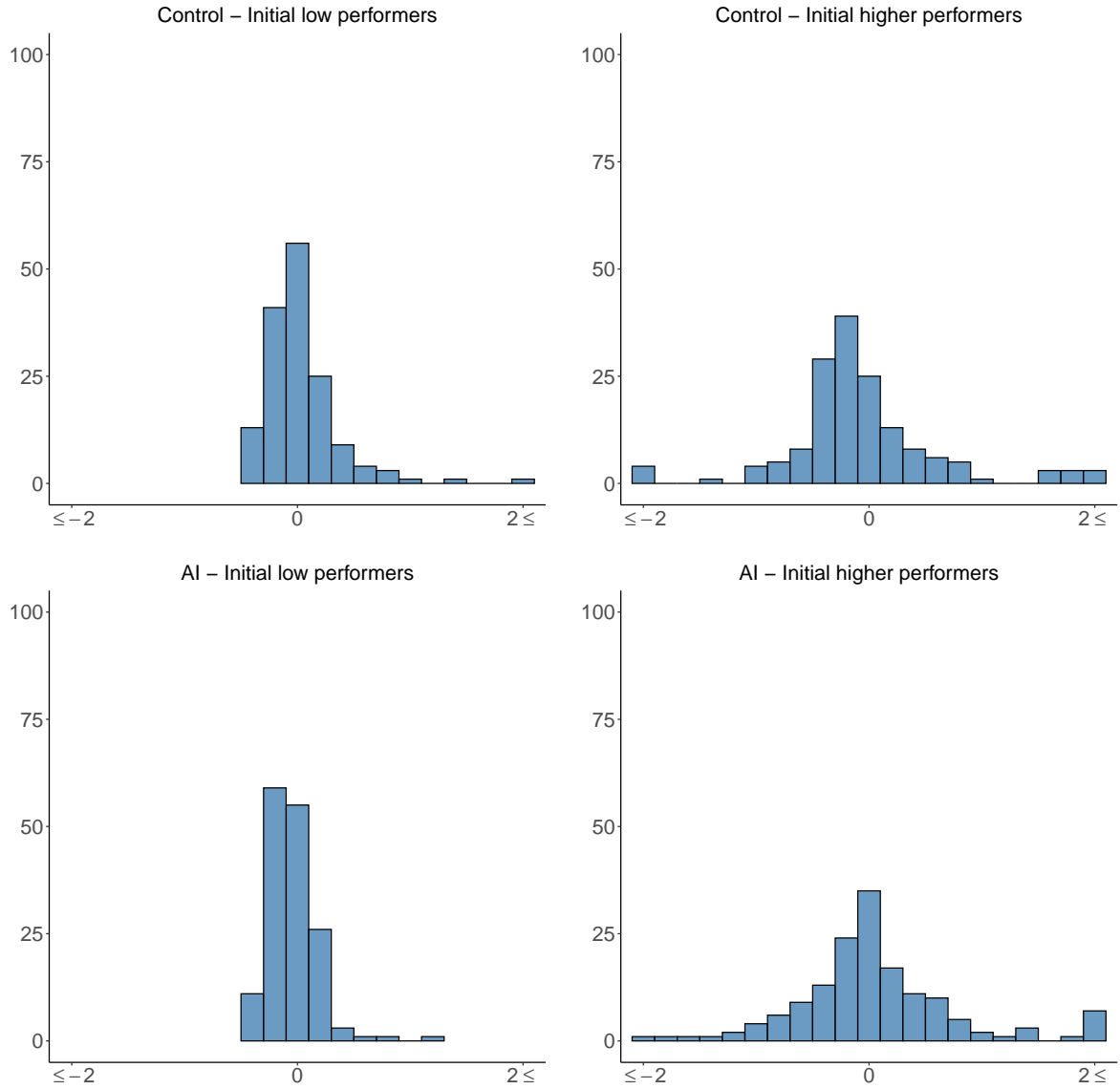
Notes: This figure presents heterogeneous treatment effect estimates for entrepreneurs with above (versus below) median pre-treatment performance. Outcomes are presented in standard deviations. Each point represents a different winsorization/sample robustness check. Effects are estimated using an intent-to-treat specification that controls for pre-treatment outcome levels and baseline covariates selected using a double LASSO procedure. All tests are two-tailed. Error bars present 95 and 90% confidence intervals clustered at the individual level.

Figure A6: Performance time trends



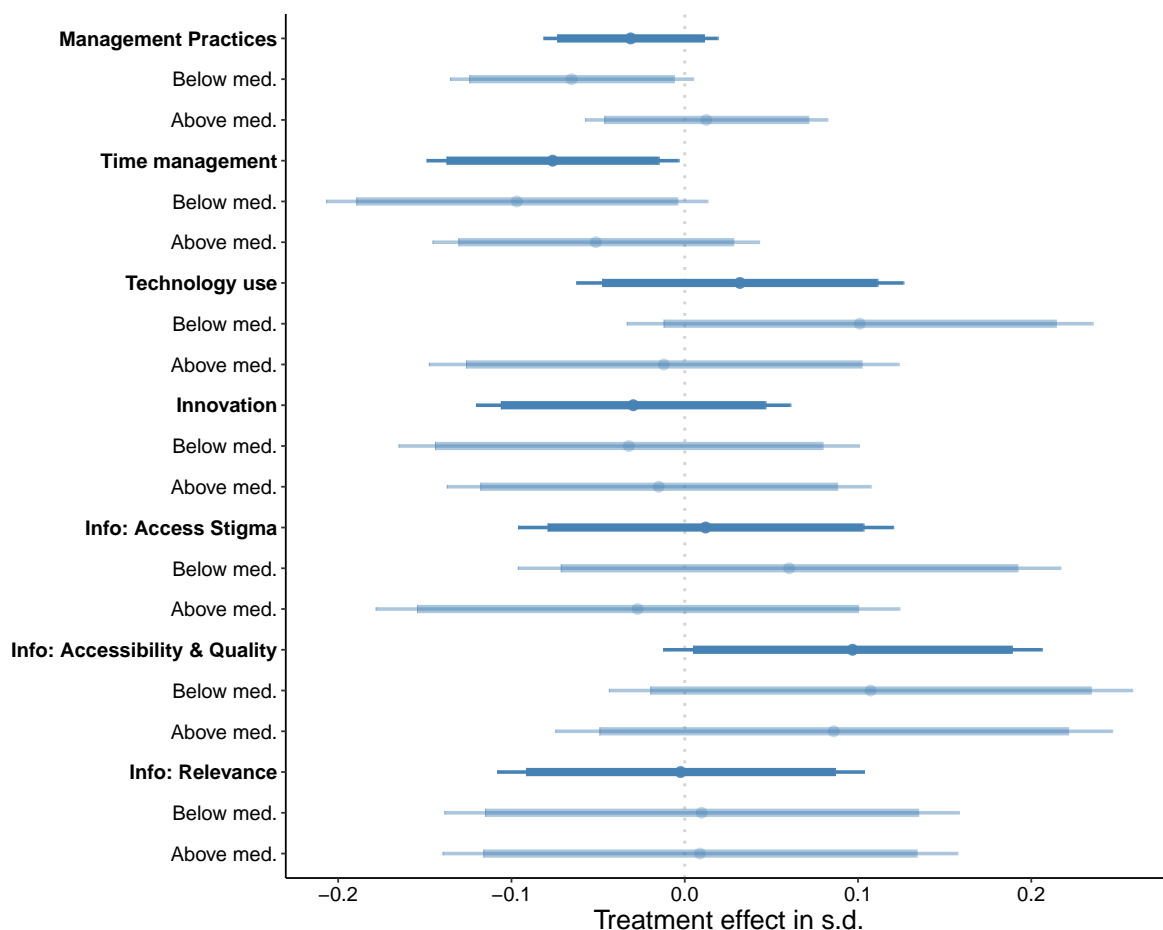
Notes: This figure presents time trends for the average performance index for above and below-median pre-treatment entrepreneurs who were randomly assigned to receive access to the AI assistant. Treatment was administered following the third time-point.

Figure A7: Residualized Performance Indices
by Condition and Pre-treatment Performance



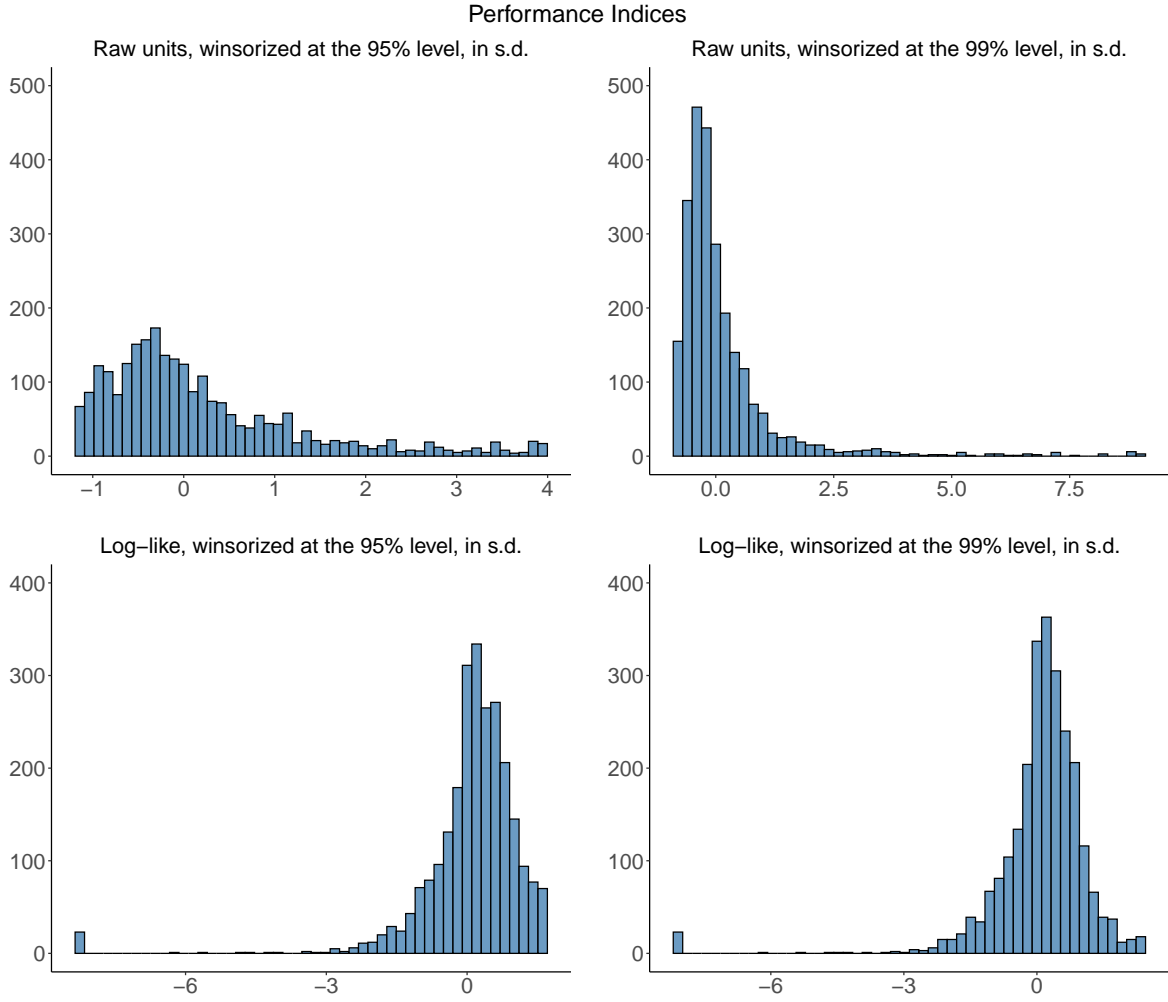
Notes: This figure presents histograms of the standardized performance index by experimental group and whether the entrepreneur was above or below median for pre-treatment performance. The index is residualized using the double-LASSO-selected covariates reported in [Table A5](#) and for stratification block and time period. Results are restricted to participants who have completed the panel, and observations are averaged at the entrepreneur.

Figure A8: Effects on mechanism questions. It is worth noting that if only 10% of our sample asked about management, 10% about time use, 10% about technology, and so on, then we would be severely under-powered to detect treatment effects along any of these individual dimensions.



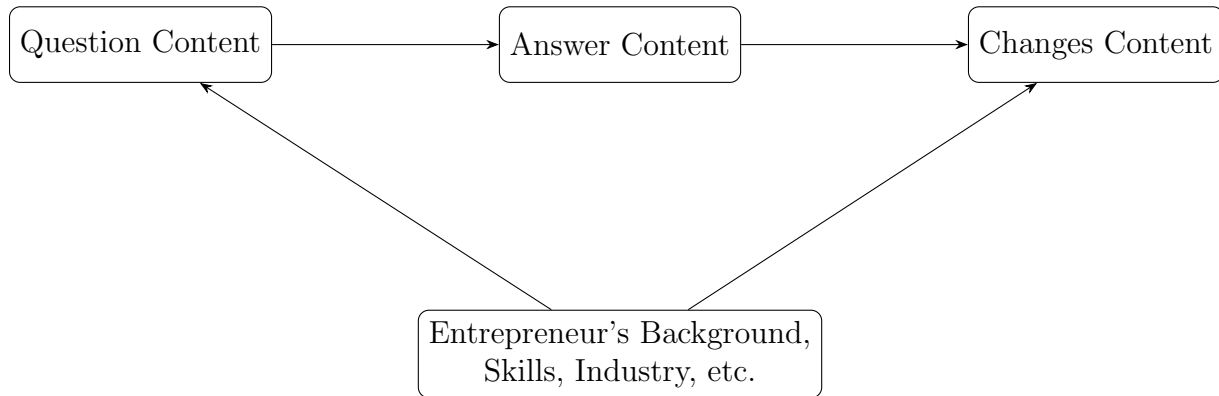
Notes: This figure depicts treatment effect estimates for our seven pre-registered mechanism indices. Dark bars present average treatment effect estimates with lasso covariates selection among all responding participants. Light bars present treatment effect estimates for above- and below-median initially performing firms using the 99% winsorized pre-treatment performance index to generate the median split. Error bars present 95 and 90% confidence intervals.

Figure A9: Distribution of outcomes



Notes: These figures present the distribution of four functional forms of our performance index. The top two figures depict results measured in raw units (KSH), and the bottom two depict results measured following an inverse hyperbolic sine transformation. In the first column results are winsorized at the 95% level and in the second column they are winsorized 99% level. In all cases, these results are presented in standard deviations for comparability.

Figure A11: DAG illustrating the backdoor criterion. Business changes depend on the AI generated answers and the entrepreneur's background, skills, industry, and much much more. The only variation linked to the AI's answer linked to the entrepreneur stems from the content of the questions because the AI has no way of knowing or communicating with the entrepreneur other than through the WhatsApp chat, which we fully observe. As a result, conditional on a given model of the text data generating process, controlling for the question content closes all possible backdoor causal paths between the answer content and the changes which allows us to estimate the causal impact of an entrepreneur's answer content on the entrepreneur's business changes.



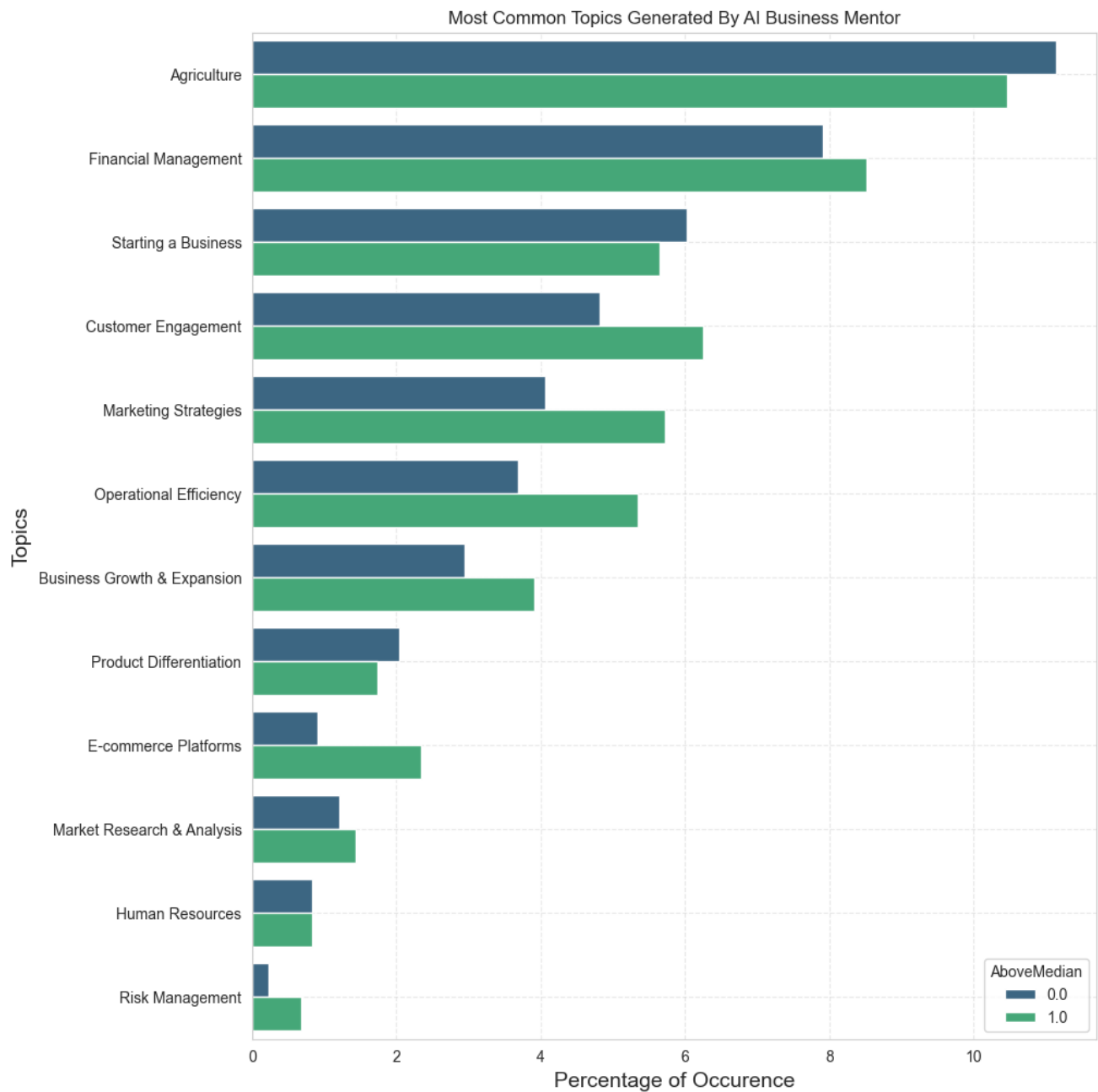
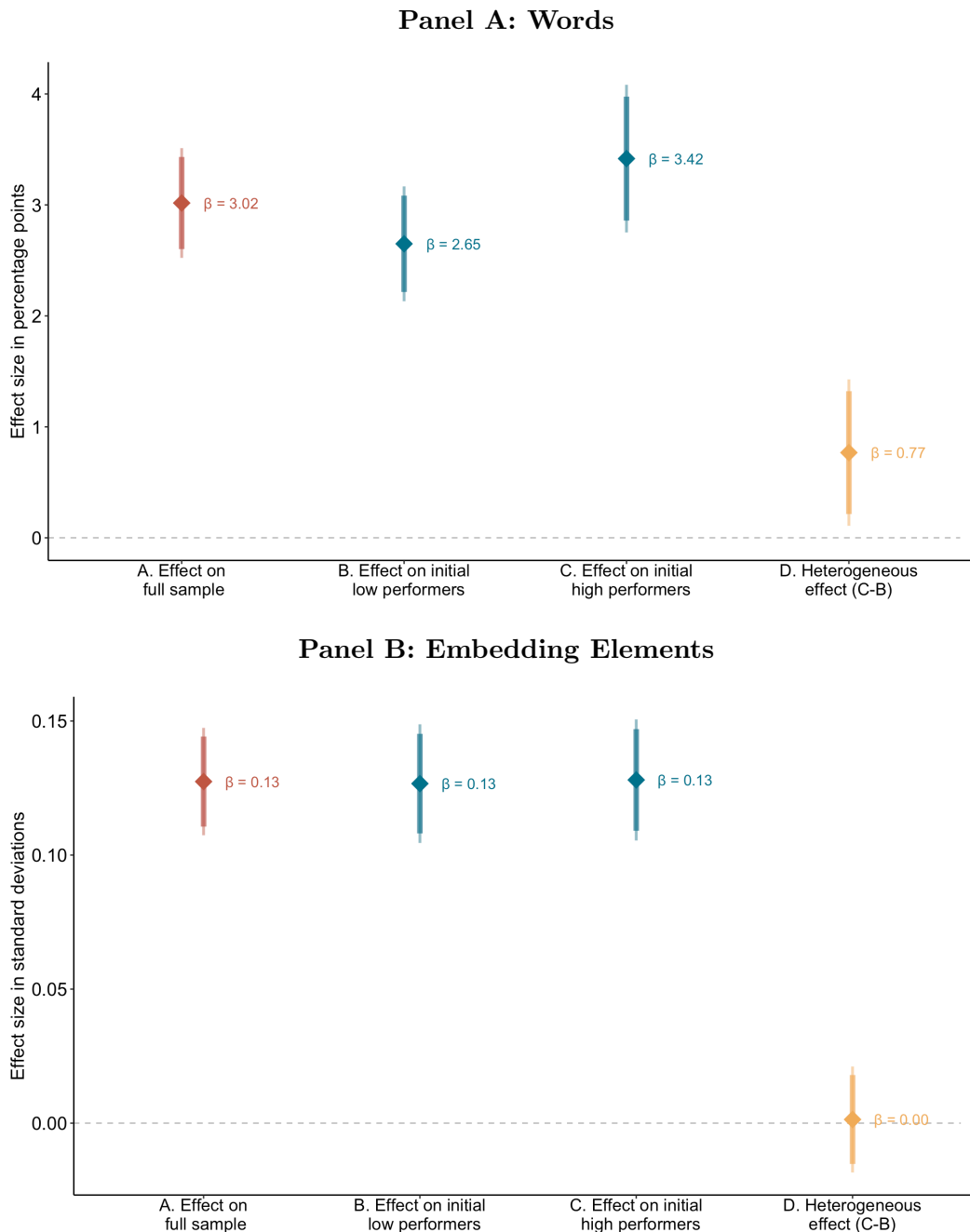


Figure A10: Distribution of the most common topics for the AI generated responses. As with the topic distribution for messages shown in [Table 1](#) there are few differences between low and high performers. See [Subsection K.2](#) for additional details.

Figure A12: Estimates showing that the content of the AI advice impacted the types of business changes made by an entrepreneur. Panel A shows that words that appear in the AI's advice text increase the chance the word appears in the entrepreneur's description of their business changes. Panel B shows that when the value of a word embedding element increases in the AI's advice text the value of the same element increases for the entrepreneur's business changes text.



Notes: Error bars present 90 and 95% confidence intervals. Standard errors are clustered at the entrepreneur and word level in Panel A. At the entrepreneur and embedding element level in Panel B. The estimates in Panel A (B) are generated from the estimates in Panel A (B) Models 3-6 in [Table A21](#). All estimates account for entrepreneur fixed effects and for whether the presence of the word or value of the dimension in the entrepreneur's question text. As discussed in [Figure A11](#) by controlling for the question the backdoor criterion lets us recover the causal impact of the advice text on the business changes text.

B Appendix Tables

Table A1: Balance

	Full analytic sample				Any post-treatment				Completed Panel			
	All	AI	Control	<i>p</i>	All	AI	Control	<i>p</i>	All	AI	Control	<i>p</i>
Business features												
Pre-treatment performance												
99pct winsorize (s.d.)	-0.00	-0.02	0.02	0.58	0.00	-0.02	0.02	0.61	0.01	-0.01	0.03	0.64
95pct winsorize (s.d.)	0.00	-0.03	0.04	0.34	0.00	-0.03	0.03	0.38	0.01	-0.02	0.04	0.39
Log-like 99pct winsorize (s.d.)	-0.00	-0.02	0.02	0.53	-0.00	-0.02	0.02	0.55	0.01	-0.02	0.03	0.55
Log-like 95pct winsorize (s.d.)	-0.00	-0.02	0.02	0.53	-0.00	-0.02	0.02	0.56	0.01	-0.02	0.03	0.56
Other business features												
Management practices count	13.10	13.08	13.12	0.88	13.11	13.08	13.13	0.84	13.12	13.11	13.14	0.87
Oldest bus.: less than 1 year	0.20	0.20	0.21	0.86	0.21	0.20	0.21	0.78	0.20	0.20	0.21	0.84
Oldest bus.: 1 year	0.41	0.41	0.40	0.79	0.41	0.42	0.40	0.73	0.41	0.42	0.40	0.68
Oldest bus.: 2 years	0.23	0.22	0.24	0.71	0.23	0.22	0.23	0.64	0.23	0.22	0.24	0.57
Oldest bus.: 3 or more years	0.16	0.16	0.15	0.80	0.16	0.16	0.15	0.71	0.16	0.16	0.15	0.74
Sector												
Agriculture	0.19	0.18	0.19	0.78	0.19	0.18	0.19	0.78	0.19	0.19	0.19	0.92
Cosmetics	0.07	0.06	0.08	0.26	0.07	0.06	0.08	0.26	0.07	0.05	0.08	0.16
Carpentry	0.00	0.01	0.00	0.57	0.00	0.01	0.00	0.57	0.00	0.01	0.00	0.56
Information technology	0.11	0.14	0.09	0.05	0.11	0.13	0.09	0.09	0.11	0.14	0.09	0.08
Jewelry	0.02	0.02	0.01	0.21	0.02	0.02	0.01	0.21	0.02	0.02	0.01	0.20
Multimedia services	0.03	0.03	0.02	0.13	0.03	0.03	0.02	0.13	0.03	0.04	0.02	0.13
Clothing	0.21	0.22	0.20	0.49	0.21	0.22	0.19	0.37	0.21	0.23	0.19	0.32
Food	0.27	0.24	0.30	0.09	0.27	0.24	0.30	0.11	0.27	0.24	0.30	0.09
Tailor	0.03	0.03	0.03	0.84	0.03	0.03	0.03	0.98	0.03	0.03	0.03	0.81
Transportation	0.02	0.02	0.02	0.98	0.02	0.02	0.02	0.98	0.02	0.02	0.02	1.00
Other	0.08	0.09	0.08	0.60	0.09	0.09	0.08	0.60	0.09	0.09	0.08	0.67
Participant features												
Female	0.32	0.31	0.33	0.48	0.32	0.31	0.33	0.48	0.31	0.30	0.33	0.39
Children	0.73	0.71	0.76	0.48	0.73	0.70	0.76	0.38	0.73	0.69	0.76	0.35
Age	25.86	25.73	25.99	0.54	25.85	25.71	25.99	0.51	25.86	25.72	26.01	0.50
Complete college	0.53	0.53	0.53	0.82	0.53	0.54	0.52	0.76	0.53	0.53	0.53	0.87
Started college	0.37	0.37	0.37	0.90	0.37	0.37	0.37	0.90	0.37	0.38	0.37	0.87
No college	0.10	0.09	0.11	0.56	0.10	0.09	0.11	0.47	0.10	0.09	0.10	0.59
ChatGPT at least once a week	0.36	0.34	0.38	0.31	0.36	0.34	0.38	0.27	0.36	0.34	0.38	0.28
Agreeableness	9.57	9.46	9.68	0.09	9.58	9.47	9.69	0.10	9.59	9.50	9.69	0.15
Conscientiousness	10.21	10.15	10.27	0.28	10.22	10.15	10.29	0.21	10.22	10.14	10.29	0.20
Extraversion	8.45	8.31	8.60	0.07	8.46	8.32	8.59	0.09	8.46	8.31	8.60	0.07
Openness	10.20	10.13	10.27	0.23	10.19	10.13	10.26	0.24	10.21	10.15	10.27	0.29
Neuroticism	7.26	7.31	7.22	0.51	7.26	7.31	7.22	0.55	7.27	7.31	7.22	0.55
Location (county)												
Machakos	0.02	0.02	0.02	0.81	0.02	0.03	0.02	0.81	0.02	0.03	0.02	0.79
Bomet	0.03	0.03	0.02	0.63	0.03	0.03	0.02	0.63	0.03	0.03	0.02	0.61
Nyeri	0.03	0.02	0.03	0.98	0.02	0.02	0.03	0.78	0.02	0.02	0.02	1.00
Kisii	0.03	0.02	0.04	0.07	0.03	0.02	0.04	0.07	0.03	0.02	0.05	0.07
Mombasa	0.03	0.03	0.03	0.67	0.03	0.03	0.03	0.67	0.03	0.04	0.03	0.65
Kisumu	0.04	0.03	0.05	0.22	0.04	0.03	0.05	0.22	0.04	0.03	0.05	0.23
Kericho	0.05	0.04	0.05	0.82	0.04	0.04	0.04	0.97	0.05	0.05	0.05	1.00
Bungoma	0.05	0.06	0.04	0.49	0.05	0.06	0.04	0.49	0.05	0.05	0.05	0.58
Kakamega	0.05	0.06	0.04	0.49	0.05	0.06	0.04	0.49	0.05	0.06	0.05	0.47
Uasin Gishu	0.05	0.05	0.06	0.83	0.06	0.05	0.06	0.83	0.05	0.05	0.06	0.72
Nakuru	0.07	0.07	0.08	0.62	0.07	0.07	0.08	0.73	0.07	0.07	0.08	0.76
Kiambu	0.08	0.07	0.09	0.45	0.08	0.08	0.09	0.45	0.08	0.07	0.09	0.47
Nairobi	0.21	0.21	0.21	0.99	0.21	0.21	0.21	0.99	0.21	0.20	0.21	0.84
Other	0.26	0.28	0.24	0.24	0.26	0.28	0.24	0.28	0.26	0.28	0.24	0.23
Completed Panel	0.97	0.97	0.98	0.35	0.98	0.97	0.99	0.25	1.00	1.00	1.00	0.35
Sample size	640	322	318		634	319	315		622	311	311	

Mean values and balance statistics for all participants all in the analytic sample (i.e., all treated participants), all participants who completed at least one post-treatment outcome survey, and all participants who completed the entire panel.

Table A2: Literature Review: Generative AI, Productivity, Performance

Authors (1)	Main outcome (2)	Can select task (3)	Average treatment effect (4)	Inequality (5)
This paper	Profits, revenue	Yes	Null	Increase
Noy and Zhang (2023)	Writing quality	No	18% increase in writing quality	Decrease
Peng et al. (2023)	Coding efficiency	No	55.8% improvement in coding speed.	Decrease
Brynjolfsson, Li, and Raymond (2023)	Customer chat efficiency	No	14% increase in productivity.	Decrease
Dell’Acqua et al. (2023)	Product design & data analysis quality	No	Increase*	Decrease
Doshi and Hauser (2023)	Creative writing	No	7% increase in novelty.	Decrease

This table reviews literature on the experimental impacts of generative AI on productivity and performance. “Can select task” (Col. 3) refers to whether the user has broad scope of control over what task they use generative AI for (versus being assigned to use AI for a specific task). Col. 4 presents the average treatment effect in the units provided by the author. Col. 5 provides evidence on whether the AI tool studied in the experiment led to an increase or decrease in inequality. Because inequality was measured in different ways across the studies (e.g., based on experience versus pre-treatment performance), we report only qualitative conclusions about the effect on inequality. Col. 4 indicates whether AI led to an increase or decrease in productivity/performance (* [Dell’Acqua et al. \(2023\)](#) provide on product design task and a data analysis task, and find performance improvements only for the product design task). Col. 5 describes whether AI use led to an increase or a decrease in inequality.

Table A3: Overview of measures

Outcomes (1)	Waves (2)	pre-registered (3)
Performance Index	Surveys 1-7	✓
Management Practices	Surveys 3-7	✓
Time Management	Surveys 4-7	✓
Technology Use	Surveys 4-7	✓
Innovation Activities	Surveys 4-7	✓
Info: Access Stigma	Surveys 4-7	✓
Info: Relevance	Surveys 4-7	✓
Info: Accessibility and Quality	Surveys 4-7	✓
Text Analysis	Post-treatment (~4-7)	✗

This table reports which outcomes were collected in each survey wave. Surveys 1-3 are pre-treatment waves and surveys 4-7 are post-treatment waves. The final column indicates whether the outcome was pre-registered.

Table A4: Effects on performance

	Full sample				Below median		Above median	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A) All observations, winsorized at the 99% level								
AI mentor	0.035 (0.072)	0.050 (0.054)	-0.098** (0.034)	-0.082* (0.034)	-0.098** (0.034)	-0.092** (0.032)	0.163 (0.141)	0.176† (0.101)
AI mentor*(\geq median)			0.263† (0.145)	0.263* (0.105)				
Post-treatment obs	2,514	2,514	2,514	2,514	1,251	1,251	1,263	1,263
B) All observations, winsorized at the 95% level								
AI mentor	-0.041 (0.058)	-0.005 (0.052)	-0.147** (0.050)	-0.126** (0.047)	-0.147** (0.050)	-0.136** (0.046)	0.052 (0.104)	0.102 (0.092)
AI mentor*(\geq median)			0.198† (0.115)	0.231* (0.103)				
Post-treatment obs	2,514	2,514	2,514	2,514	1,251	1,251	1,263	1,263
C) Completed panel, winsorized at the 99% level								
AI mentor	0.033 (0.073)	0.053 (0.056)	-0.095** (0.035)	-0.081* (0.034)	-0.094** (0.035)	-0.088** (0.032)	0.162 (0.143)	0.190† (0.104)
AI mentor*(\geq median)			0.256† (0.147)	0.267* (0.107)				
Post-treatment obs	2,488	2,488	2,488	2,488	1,244	1,244	1,244	1,244
D) Completed panel, winsorized at the 95% level								
AI mentor	-0.046 (0.058)	-0.009 (0.053)	-0.141** (0.051)	-0.122* (0.049)	-0.140** (0.051)	-0.130** (0.048)	0.050 (0.105)	0.103 (0.092)
AI mentor*(\geq median)			0.190 (0.116)	0.225* (0.104)				
Post-treatment obs	2,488	2,488	2,488	2,488	1,244	1,244	1,244	1,244
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
Rand. block FE	Y	Y	Y	Y	Y	Y	Y	Y
Pre-treat. covariates	N	Y	N	Y	N	Y	N	Y

This table reports the performance effects of the AI assistant. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All tests are two-tailed. Cols. 1-2 report average treatment effects. Cols. 3-4 report heterogeneous treatment effects based on whether the entrepreneur has pre-treatment performance that was above median. Cols 5-6 report average effects on below-median entrepreneurs, and 7-8 report effects on above-median entrepreneurs. Panels A and B report effects using all post-treatment observations, and panels C and D report effects for participants who complete the entire panel. Panels A and C report effects winsorized at the 99% level, and panels B and D report effects winsorized at the 95% level. Covariates (in even-numbered columns) were selected using a double-LASSO procedure. Covariates selected for each model are reported in [Table A5](#). Standard errors are clustered at the individual level.

Table A5: Double-post-LASSO selected covariates

	Table A3, Panel [X]			
	[A]	[B]	[C]	[D]
Pre-treatment performance index	Y	Y	Y	Y
Pre-treatment management practices	Y	Y	Y	Y
Sector: IT Revenue and services	Y	N	N	N
Completed college	Y	Y	Y	Y
Extraversion (Big 5 personality)	Y	Y	Y	Y
Gender	N	Y	N	Y

This table reports the double-LASSO-selected covariates for [Table A4](#). Panel C corresponds to the results reported in [Figure 3](#). The following variables are included in our covariate selection: pre-treatment performance, management practices (sum), gender, oldest business age, sector, children (count), age, education, GPT use, Big 5 personality (agreeableness, conscientiousness, extroversion, openness, neuroticism), and county. Additionally, we always control for randomization strata blocks and time period.

Table A6: Effects on performance
(Log-like transformation)

	Full sample				Below median		Above median	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A) All observations, winsorized at the 99% level								
AI mentor	-0.015 (0.065)	-0.004 (0.060)	-0.088 (0.096)	-0.071 (0.092)	-0.091 (0.097)	-0.099 (0.091)	0.055 (0.088)	0.049 (0.075)
AI mentor*(\geq median)			0.144 (0.131)	0.136 (0.118)				
Post-treatment obs	2,514	2,514	2,514	2,514	1,251	1,251	1,263	1,263
B) All observations, winsorized at the 95% level								
AI mentor	-0.024 (0.065)	-0.014 (0.061)	-0.090 (0.100)	-0.073 (0.095)	-0.093 (0.100)	-0.102 (0.094)	0.038 (0.085)	0.022 (0.072)
AI mentor*(\geq median)			0.129 (0.132)	0.121 (0.119)				
Post-treatment obs	2,514	2,514	2,514	2,514	1,251	1,251	1,263	1,263
C) Completed panel, winsorized at the 99% level								
AI mentor	-0.020 (0.065)	-0.007 (0.061)	-0.096 (0.097)	-0.075 (0.092)	-0.098 (0.097)	-0.103 (0.091)	0.055 (0.088)	0.048 (0.076)
AI mentor*(\geq median)			0.151 (0.132)	0.139 (0.119)				
Post-treatment obs	2,488	2,488	2,488	2,488	1,244	1,244	1,244	1,244
D) Completed panel, winsorized at the 95% level								
AI mentor	-0.029 (0.066)	-0.017 (0.061)	-0.099 (0.100)	-0.078 (0.095)	-0.101 (0.100)	-0.106 (0.094)	0.038 (0.086)	0.021 (0.073)
AI mentor*(\geq median)			0.136 (0.132)	0.124 (0.119)				
Post-treatment obs	2,488	2,488	2,488	2,488	1,244	1,244	1,244	1,244
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
Rand. block FE	Y	Y	Y	Y	Y	Y	Y	Y
Pre-treat. covariates	N	Y	N	Y	N	Y	N	Y

This table reports the performance effects of the AI assistant. Outcomes are reported for an inverse-hyperbolic sine transformation. $\dagger p < 0.10$, $* p < 0.05$, $** p < 0.01$, $*** p < 0.001$. All tests are two-tailed. Cols. 1-2 report average treatment effects. Cols. 3-4 report heterogeneous treatment effects based on whether the entrepreneur has pre-treatment performance that was above median. Cols 5-6 report average effects on below-median entrepreneurs, and 7-8 report effects on above-median entrepreneurs. Panels A and B report effects using all post-treatment observations, and panels C and D report effects for participants who complete the entire panel. Panels A and C report effects winsorized at the 99% level, and panels B and D report effects winsorized at the 95% level. Covariates (in even-numbered columns) were selected using a double-LASSO procedure. Standard errors are clustered at the individual level.

Table A7: Effects on performance index component: Weekly profits

	Full sample				Below median		Above median	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A) All observations, winsorized at the 99% level								
AI mentor	0.016 (0.077)	0.023 (0.058)	-0.123** (0.042)	-0.122** (0.041)	-0.124** (0.042)	-0.123** (0.040)	0.149 (0.147)	0.157 (0.110)
AI mentor*(\geq median)			0.274† (0.153)	0.288* (0.117)				
Post-treatment obs	2,514	2,514	2,514	2,514	1,251	1,251	1,263	1,263
B) All observations, winsorized at the 95% level								
AI mentor	-0.058 (0.065)	-0.031 (0.059)	-0.180** (0.057)	-0.165** (0.055)	-0.181** (0.057)	-0.173** (0.053)	0.048 (0.116)	0.081 (0.104)
AI mentor*(\geq median)			0.228† (0.129)	0.256* (0.117)				
Post-treatment obs	2,514	2,514	2,514	2,514	1,251	1,251	1,263	1,263
C) Completed panel, winsorized at the 99% level								
AI mentor	0.014 (0.077)	0.021 (0.058)	-0.117** (0.043)	-0.118** (0.042)	-0.117** (0.043)	-0.115** (0.041)	0.145 (0.149)	0.158 (0.111)
AI mentor*(\geq median)			0.262† (0.155)	0.280* (0.119)				
Post-treatment obs	2,488	2,488	2,488	2,488	1,244	1,244	1,244	1,244
D) Completed panel, winsorized at the 95% level								
AI mentor	-0.063 (0.065)	-0.035 (0.059)	-0.172** (0.059)	-0.159** (0.057)	-0.171** (0.059)	-0.164** (0.055)	0.047 (0.117)	0.082 (0.105)
AI mentor*(\geq median)			0.218† (0.131)	0.249* (0.118)				
Post-treatment obs	2,488	2,488	2,488	2,488	1,244	1,244	1,244	1,244
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
Rand. block FE	Y	Y	Y	Y	Y	Y	Y	Y
Pre-treat. covariates	N	Y	N	Y	N	Y	N	Y

This table reports the performance effects of the AI assistant on weekly profits. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All tests are two-tailed. Cols. 1-2 report average treatment effects. Cols. 3-4 report heterogeneous treatment effects based on whether the entrepreneur has pre-treatment performance that was above median. Cols 5-6 report average effects on below-median entrepreneurs, and 7-8 report effects on above-median entrepreneurs. Panels A and B report effects using all post-treatment observations, and panels C and D report effects for participants who complete the entire panel. Panels A and C report effects winsorized at the 99% level, and panels B and D report effects winsorized at the 95% level. Covariates (in even-numbered columns) were selected using a double-LASSO procedure. Standard errors are clustered at the individual level.

Table A8: Effects on performance index component: Monthly profits

	Full sample				Below median		Above median	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A) All observations, winsorized at the 99% level								
AI mentor	0.032 (0.077)	0.047 (0.060)	-0.122** (0.044)	-0.109* (0.043)	-0.123** (0.044)	-0.112** (0.040)	0.178 (0.147)	0.201† (0.111)
AI mentor*(\geq median)			0.304* (0.153)	0.309** (0.118)				
Post-treatment obs	2,514	2,514	2,514	2,514	1,251	1,251	1,263	1,263
B) All observations, winsorized at the 95% level								
AI mentor	-0.048 (0.064)	-0.021 (0.059)	-0.166** (0.057)	-0.145** (0.055)	-0.166** (0.057)	-0.154** (0.053)	0.052 (0.115)	0.075 (0.104)
AI mentor*(\geq median)			0.217† (0.128)	0.235* (0.117)				
Post-treatment obs	2,514	2,514	2,514	2,514	1,251	1,251	1,263	1,263
C) Completed panel, winsorized at the 99% level								
AI mentor	0.029 (0.078)	0.046 (0.061)	-0.117* (0.046)	-0.106* (0.044)	-0.116* (0.046)	-0.106* (0.041)	0.177 (0.149)	0.205† (0.112)
AI mentor*(\geq median)			0.293† (0.155)	0.305* (0.120)				
Post-treatment obs	2,488	2,488	2,488	2,488	1,244	1,244	1,244	1,244
D) Completed panel, winsorized at the 95% level								
AI mentor	-0.053 (0.065)	-0.025 (0.059)	-0.155** (0.059)	-0.137* (0.057)	-0.153** (0.059)	-0.143* (0.055)	0.050 (0.116)	0.076 (0.104)
AI mentor*(\geq median)			0.204 (0.130)	0.224† (0.118)				
Post-treatment obs	2,488	2,488	2,488	2,488	1,244	1,244	1,244	1,244
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
Rand. block FE	Y	Y	Y	Y	Y	Y	Y	Y
Pre-treat. covariates	N	Y	N	Y	N	Y	N	Y

This table reports the performance effects of the AI assistant on monthly profits. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All tests are two-tailed. Cols. 1-2 report average treatment effects. Cols. 3-4 report heterogeneous treatment effects based on whether the entrepreneur has pre-treatment performance that was above median. Cols 5-6 report average effects on below-median entrepreneurs, and 7-8 report effects on above-median entrepreneurs. Panels A and B report effects using all post-treatment observations, and panels C and D report effects for participants who complete the entire panel. Panels A and C report effects winsorized at the 99% level, and panels B and D report effects winsorized at the 95% level. Covariates (in even-numbered columns) were selected using a double-LASSO procedure. Standard errors are clustered at the individual level.

Table A9: Effects on performance index component: Weekly revenue

	Full sample				Below median		Above median	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A) All observations, winsorized at the 99% level								
AI mentor	0.004 (0.070)	0.022 (0.057)	-0.073** (0.028)	-0.050† (0.029)	-0.073* (0.028)	-0.068* (0.027)	0.078 (0.136)	0.089 (0.108)
AI mentor*(\geq median)			0.152 (0.139)	0.144 (0.110)				
Post-treatment obs	2,514	2,514	2,514	2,514	1,251	1,251	1,263	1,263
B) All observations, winsorized at the 95% level								
AI mentor	-0.051 (0.060)	-0.011 (0.056)	-0.132* (0.052)	-0.099* (0.050)	-0.132* (0.052)	-0.113* (0.049)	0.022 (0.110)	0.071 (0.100)
AI mentor*(\geq median)			0.154 (0.121)	0.172 (0.111)				
Post-treatment obs	2,514	2,514	2,514	2,514	1,251	1,251	1,263	1,263
C) Completed panel, winsorized at the 99% level								
AI mentor	0.001 (0.070)	0.024 (0.057)	-0.075** (0.028)	-0.053† (0.030)	-0.074** (0.028)	-0.069* (0.027)	0.077 (0.138)	0.106 (0.108)
AI mentor*(\geq median)			0.151 (0.141)	0.153 (0.111)				
Post-treatment obs	2,488	2,488	2,488	2,488	1,244	1,244	1,244	1,244
D) Completed panel, winsorized at the 95% level								
AI mentor	-0.058 (0.061)	-0.016 (0.056)	-0.136** (0.052)	-0.105* (0.050)	-0.135** (0.052)	-0.118* (0.049)	0.023 (0.111)	0.072 (0.100)
AI mentor*(\geq median)			0.157 (0.122)	0.177 (0.112)				
Post-treatment obs	2,488	2,488	2,488	2,488	1,244	1,244	1,244	1,244
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
Rand. block FE	Y	Y	Y	Y	Y	Y	Y	Y
Pre-treat. covariates	N	Y	N	Y	N	Y	N	Y

This table reports the performance effects of the AI assistant on weekly revenue. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All tests are two-tailed. Cols. 1-2 report average treatment effects. Cols. 3-4 report heterogeneous treatment effects based on whether the entrepreneur has pre-treatment performance that was above median. Cols 5-6 report average effects on below-median entrepreneurs, and 7-8 report effects on above-median entrepreneurs. Panels A and B report effects using all post-treatment observations, and panels C and D report effects for participants who complete the entire panel. Panels A and C report effects winsorized at the 99% level, and panels B and D report effects winsorized at the 95% level. Covariates (in even-numbered columns) were selected using a double-LASSO procedure. Standard errors are clustered at the individual level.

Table A10: Effects on performance index component: Monthly revenue

	Full sample				Below median		Above median	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A) All observations, winsorized at the 99% level								
AI mentor	0.089 (0.089)	0.114† (0.067)	-0.073† (0.039)	-0.051 (0.039)	-0.074† (0.039)	-0.062† (0.037)	0.246 (0.173)	0.276* (0.127)
AI mentor*(\geq median)			0.322† (0.177)	0.328* (0.133)				
Post-treatment obs	2,514	2,514	2,514	2,514	1,251	1,251	1,263	1,263
B) All observations, winsorized at the 95% level								
AI mentor	-0.007 (0.066)	0.034 (0.061)	-0.109† (0.057)	-0.084 (0.055)	-0.110† (0.057)	-0.089† (0.053)	0.085 (0.119)	0.143 (0.107)
AI mentor*(\geq median)			0.194 (0.132)	0.228† (0.120)				
Post-treatment obs	2,514	2,514	2,514	2,514	1,251	1,251	1,263	1,263
C) Completed panel, winsorized at the 99% level								
AI mentor	0.087 (0.090)	0.117† (0.069)	-0.071† (0.040)	-0.048 (0.039)	-0.070† (0.040)	-0.059 (0.037)	0.247 (0.175)	0.288* (0.131)
AI mentor*(\geq median)			0.317† (0.179)	0.329* (0.136)				
Post-treatment obs	2,488	2,488	2,488	2,488	1,244	1,244	1,244	1,244
D) Completed panel, winsorized at the 95% level								
AI mentor	-0.012 (0.067)	0.032 (0.061)	-0.103† (0.059)	-0.079 (0.055)	-0.101† (0.059)	-0.084 (0.054)	0.081 (0.120)	0.141 (0.109)
AI mentor*(\geq median)			0.183 (0.134)	0.220† (0.122)				
Post-treatment obs	2,488	2,488	2,488	2,488	1,244	1,244	1,244	1,244
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
Rand. block FE	Y	Y	Y	Y	Y	Y	Y	Y
Pre-treat. covariates	N	Y	N	Y	N	Y	N	Y

This table reports the performance effects of the AI assistant on monthly revenue. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All tests are two-tailed. Cols. 1-2 report average treatment effects. Cols. 3-4 report heterogeneous treatment effects based on whether the entrepreneur has pre-treatment performance that was above median. Cols 5-6 report average effects on below-median entrepreneurs, and 7-8 report effects on above-median entrepreneurs. Panels A and B report effects using all post-treatment observations, and panels C and D report effects for participants who complete the entire panel. Panels A and C report effects winsorized at the 99% level, and panels B and D report effects winsorized at the 95% level. Covariates (in even-numbered columns) were selected using a double-LASSO procedure. Standard errors are clustered at the individual level.

Table A11: Probability of 0 profits or revenue by experimental condition

	p(outcome=0)		
	Full sample	Control	AI
Monthly profits			
Survey 4	0.016	0.019	0.013
Survey 5	0.008	0.006	0.009
Survey 6	0.008	0.006	0.010
Survey 7	0.008	0.006	0.010
Weekly profits			
Survey 4	0.016	0.019	0.013
Survey 5	0.008	0.006	0.009
Survey 6	0.010	0.010	0.010
Survey 7	0.011	0.010	0.013
Monthly revenue			
Survey 4	0.014	0.019	0.009
Survey 5	0.009	0.006	0.013
Survey 6	0.008	0.006	0.010
Survey 7	0.008	0.006	0.010
Weekly revenue			
Survey 4	0.014	0.019	0.009
Survey 5	0.008	0.006	0.009
Survey 6	0.008	0.006	0.010
Survey 7	0.008	0.006	0.010

This table reports the probability of reporting 0 weekly/monthly profits/revenue in each of the four post-treatment periods.

Table A12: Calculating effect sizes in percentages

	Below median	Above median
A) All observations, winsorized at the 99% level		
Profit - Month	-8.09	15.43
Profit - Week	-7.25	13.84
Revenue - Month	-8.85	16.89
Revenue - Week	-7.36	14.04
Average profit	-7.67	14.64
Average revenue	-8.10	15.46
Average (profit and revenue)	-7.89	15.05
B) Completed panel, winsorized at the 99% level		
Profit - Month	-7.77	16.71
Profit - Week	-6.97	14.98
Revenue - Month	-8.50	18.27
Revenue - Week	-7.07	15.19
Average profit	-7.37	15.84
Average revenue	-7.78	16.73
Average (profit and revenue)	-7.58	16.29

This tables presents percent estimates of our experimental treatment effects for above and below median initial performance entrepreneurs. All outcomes were winsorized at the 99% level. For rows 1-4 and 8-11 we calculate percent estimates (1) taking the treatment effect estimates on our performance index, (2) multiplying this treatment effect estimate by the standard deviation of each outcome, and (3) dividing by the mean pre-treatment value of the outcome. Rows 5-7 and 12-14 capture the average of these effect estimates across outcomes within each panel.

Table A13: Heterogeneous Treatment Effects by Gender

	(1)	(2)
A) All observations, winsorized at the 99% level		
AI mentor	-0.004 (0.085)	0.089 (0.065)
AI mentor*(Female)	0.125 (0.162)	-0.119 (0.112)
Post-treatment obs	2,514	2,514
B) All observations, winsorized at the 95% level		
AI mentor	-0.060 (0.070)	-0.009 (0.064)
AI mentor*(Female)	0.059 (0.123)	-0.070 (0.106)
Post-treatment obs	2,514	2,514
C) Completed panel, winsorized at the 99% level		
AI mentor	-0.005 (0.085)	0.094 (0.067)
AI mentor*(Female)	0.121 (0.165)	-0.131 (0.114)
Post-treatment obs	2,488	2,488
D) Completed panel, winsorized at the 95% level		
AI mentor	-0.061 (0.070)	-0.010 (0.064)
AI mentor*(Female)	0.049 (0.126)	-0.081 (0.108)
Post-treatment obs	2,488	2,488
Time FE	Y	Y
Rand. block FE	Y	Y
Pre-treat. covariates	N	Y

This table reports heterogeneous treatment effects of the AI assistant by gender. $\dagger p < 0.10$, $* p < 0.05$, $** p < 0.01$, $*** p < 0.001$. All tests are two-tailed. Cols. 1 reports a base model of heterogeneous effects without covariate adjustment. Col. 2 reports covariates selected using a double-LASSO procedure. Panels A and B report effects using all post-treatment observations, and panels C and D report effects for participants who complete the entire panel. Panels A and C report effects winsorized at the 99% level, and panels B and D report effects winsorized at the 95% level. Standard errors are clustered at the individual level.

Table A14: Heterogeneous Treatment Effects by Initial ChatGPT Use

	(1)	(2)
A) All observations, winsorized at the 99% level		
AI mentor	0.031 (0.075)	0.057 (0.063)
AI mentor*(Freq. ChatGPT use)	0.029 (0.199)	-0.022 (0.121)
Post-treatment obs	2,514	2,514
B) All observations, winsorized at the 95% level		
AI mentor	-0.036 (0.063)	-0.014 (0.058)
AI mentor*(Freq. ChatGPT use)	-0.010 (0.145)	-0.070 (0.126)
Post-treatment obs	2,514	2,514
C) Completed panel, winsorized at the 99% level		
AI mentor	0.028 (0.076)	0.058 (0.064)
AI mentor*(Freq. ChatGPT use)	0.029 (0.200)	-0.017 (0.123)
Post-treatment obs	2,488	2,488
D) Completed panel, winsorized at the 95% level		
AI mentor	-0.042 (0.064)	-0.018 (0.059)
AI mentor*(Freq. ChatGPT use)	-0.009 (0.146)	-0.071 (0.127)
Post-treatment obs	2,488	2,488
Time FE	Y	Y
Rand. block FE	Y	Y
Pre-treat. covariates	N	Y

This table reports heterogeneous treatment effects of the AI assistant by GPT use, reported as whether a participant reported using ChatGPT at least once a week prior to treatment. $\dagger p < 0.10$, $* p < 0.05$, $** p < 0.01$, $*** p < 0.001$. All tests are two-tailed. Cols. 1 reports a base model of heterogeneous effects without covariate adjustment. Col. 2 reports covariates selected using a double-LASSO procedure. Panels A and B report effects using all post-treatment observations, and panels C and D report effects for participants who complete the entire panel. Panels A and C report effects winsorized at the 99% level, and panels B and D report effects winsorized at the 95% level. Standard errors are clustered at the individual level.

Table A15: Effects of AI assistant (DID specification)

	Full sample				Below median		Above median	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A) All observations, winsorized at the 99% level								
AI mentor	0.050 (0.055)	0.050 (0.056)	-0.092** (0.033)	-0.068* (0.034)	-0.092** (0.033)	-0.092** (0.033)	0.191† (0.105)	0.188† (0.106)
AI mentor*(\geq median)			0.283* (0.110)	0.236* (0.111)				
N Entrepreneurs	637	637	637	637	318	318	319	319
B) All observations, winsorized at the 95% level								
AI mentor	0.005 (0.053)	0.005 (0.054)	-0.132** (0.048)	-0.100* (0.050)	-0.132** (0.048)	-0.132** (0.048)	0.143 (0.094)	0.141 (0.095)
AI mentor*(\geq median)			0.275** (0.106)	0.212* (0.106)				
N Entrepreneurs	637	637	637	637	318	318	319	319
C) Completed panel, winsorized at the 99% level								
AI mentor	0.049 (0.057)	0.049 (0.057)	-0.089** (0.033)	-0.074* (0.034)	-0.089** (0.033)	-0.089** (0.034)	0.186† (0.108)	0.186† (0.108)
AI mentor*(\geq median)			0.274* (0.113)	0.246* (0.113)				
N Entrepreneurs	622	622	622	622	311	311	311	311
D) Completed panel, winsorized at the 95% level								
AI mentor	-0.001 (0.054)	-0.001 (0.054)	-0.130** (0.049)	-0.108* (0.050)	-0.130** (0.049)	-0.130** (0.049)	0.126 (0.096)	0.126 (0.096)
AI mentor*(\geq median)			0.256* (0.108)	0.216* (0.108)				
N Entrepreneurs	622	622	622	622	311	311	311	311
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
Rand. block FE	Y	Y	Y	Y	Y	Y	Y	Y
Pre-treat. covariates	N	Y	N	Y	N	Y	N	Y

This table reports the performance effects of the AI assistant from a difference-in-difference specification. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All tests are two-tailed. Cols. 1-2 report average treatment effects. Cols. 3-4 report heterogeneous treatment effects based on whether the entrepreneur has pre-treatment performance that was above median. Cols 5-6 report average effects on below-median entrepreneurs, and 7-8 report effects on above-median entrepreneurs. Panels A and B report effects using all post-treatment observations, and panels C and D report effects for participants who complete the entire panel. Panels A and C report effects winsorized at the 99% level, and panels B and D report effects winsorized at the 95% level. Covariates (in even-numbered columns) were selected using a double-LASSO procedure. Standard errors are clustered at the individual level.

Table A16: Estimation of HTEs using Generic Machine Learning methods. The coefficient β_2 captures the overall heterogeneity of treatment effects. $GATE_1$ is the group average treatment effect for entrepreneurs predicted to have treatment effects below the median of predicted treatment effects. $GATE_2$ is for entrepreneurs with above median predicted treatment effects. If there is no predictable treatment effect heterogeneity, then $GATE_1$ and $GATE_2$ will be near zero and statistically insignificant. The final three rows in each panel show differences in the fraction of entrepreneurs in each group that are women, had used ChatGPT before our study, and are baseline high performers. Panel A shows the results when we only use our three pre-registered heterogeneity variables to predict HTEs. Panel B includes a broader set of baseline variables including location, sector, age, and more. See [Appendix J](#) for further details.

Panel A: Pre-registered heterogeneity variables				
	Estimate	CI lower	CI upper	p-value
ATE	0.038	-0.082	0.158	0.465
β_2	0.972	0.475	1.493	< 0.0001
$GATE_1$ (Below Median)	-0.269	-0.427	-0.116	< 0.0001
$GATE_2$ (Above Median)	0.347	0.144	0.555	< 0.0001
$GATE_2 - GATE_1$	0.609	0.351	0.879	< 0.0001
Female ₂ - Female ₁	0.0007	-0.058	0.059	0.919
Used GPT ₂ - Used GPT ₁	-0.041	-0.101	0.019	0.127
High Performer ₂ - High Performer ₁	0.338	0.279	0.398	< 0.0001
Panel B: Full set of variables				
	Estimate	CI lower	CI upper	p-value
ATE	0.039	-0.056	0.136	0.345
β_2	1.487	1.029	1.943	< 0.0001
$GATE_1$ (Below Median)	-0.282	-0.423	-0.142	< 0.0001
$GATE_2$ (Above Median)	0.368	0.171	0.578	< 0.0001
$GATE_2 - GATE_1$	0.641	0.383	0.911	< 0.0001
Female ₂ - Female ₁	-0.056	-0.115	0.003	0.034
Used GPT ₂ - Used GPT ₁	0.009	-0.052	0.069	0.739
High Performer ₂ - High Performer ₁	0.232	0.170	0.293	< 0.0001

Errors clustered at the entrepreneur level. 95% Confidence Intervals.

Table A17: Summary statistics for the AI question and answer data. Panel A reports the average number of messages sent by low and high-performing entrepreneurs to the AI assistant along with the average number of numeric followups and the average number of words in each question. Panel B reports the number of pieces of AI generated advice sent to low and high-performing entrepreneurs during our study along with the average number of words generated per message by the AI. Pieces of advice correspond to the 3-5 actionable strategies described in our system prompt. Across our entire sample the average entrepreneur sent 17.1 messages, 4.96 of these messages were numeric followups, and they on average received 23.5 pieces of AI generated advice.

Panel A: Questions Sent							
(≥Median)	Entrepreneurs w/ Questions	Number of Messages Sent to the AI				Avg. Num. Followups	Num. Words Per Question
		Average	S.D.	Median	Max		
0	145	18.6	40.8	9	402	5.34	13.5
1	136	15.5	19.1	8.5	105	4.57	14.4

Panel B: Answers Generated							
(≥Median)	Entrepreneurs w/ Answers	Pieces of Advice Generated				Num. Words	
		Average	Median	S.D.	Max	Per Answer	
0	145	21.5	13	29.8	247	168	
1	136	25.6	11	39.2	315	166	

Table A18: Summary statistics for the questions, answers, and business changes text data. Panel A shows that low and high performers were equally likely to ask at least one question and that they sent in questions totalling roughly 110 words. Panel B shows that both groups received over 2,000 words worth of AI advice. Finally, Panel C shows that treated/control and low/high performers are equally likely to describe making a business change and that on average they use about 50 words to describe these changes.

Panel A: All Questions							
(≥Median)	Entrepreneurs	Has Question(s)?	Num. Words Sent to the AI				Max
			Average	Std. Dev.	Median	Max	
0	161	0.882	106	180	55	1,506	
1	158	0.842	115	186	64	1,524	

Panel B: All Answers							
(≥Median)	Entrepreneurs	Has Answer(s)?	Num. AI Generated Words				Max
			Average	Std. Dev.	Median	Max	
0	161	0.888	2,740	6,352	1,125	64,554	
1	158	0.842	2,077	3,026	972	19,207	

Panel C: All Business Changes							
AI	(≥Median)	Entrepreneurs	Has Text?	Num. Words to Describe Changes			
				Average	Std. Dev.	Median	Max
0	0	156	0.821	46.1	44.8	35.5	219
0	1	159	0.862	53.6	54.0	42	358
1	0	161	0.851	44.9	42.9	34	182
1	1	158	0.854	57.1	53.2	40	224

Table A19: While word embeddings created from the questions asked and AI advice cannot predict whether an entrepreneur is a low/high performer word embeddings created from the self-reported business changes can predict whether an entrepreneur is a low/high performer and whether they receive the AI treatment or not.

Model	(1)	(2)	(3)	(4)	(5)
Data	<i>Questions</i>	<i>AI Advice</i>	<i>Changes</i>	<i>Changes</i>	<i>Changes</i>
Prediction	High Performer	High Performer	High Performer	AI Treated	AI Treated \times High Performer
Classes	2	2	2	2	4
Entrepreneurs	280	276	537	537	537
No Information Rate	0.518	0.518	0.505	0.505	0.255
Accuracy	0.492	0.482	0.617	0.617	0.453
95% CI	[0.42, 0.57]	[0.41, 0.56]	[0.52, 0.71]	[0.52, 0.71]	[0.36, 0.55]
P-Value [Acc > NIR]	0.741	0.860	0.013	0.013	0.000008
Mcnemar's P-Value	0.547	0.134	0.755	0.532	0.494

Table A20: An entrepreneur's treatment status and whether they are a low or high performer at baseline does not predict whether they ask a question, receive answer text, or describe making a business change. They also do not predict the number of words in question, answer, or business changes text.

	Changes Text		Questions Text		Answers Text	
	Has Text? (1) OLS	Num. Word (2) Poisson	Has Text? (3) OLS	Num. Words (4) Poisson	Has Text? (5) OLS	Num. Words (6) Poisson
Constant	0.8205*** (0.0308)	3.831*** (0.0778)	0.8820*** (0.0255)	4.662*** (0.1336)	0.8882*** (0.0249)	7.916*** (0.1827)
AI	0.0304 (0.0418)	-0.0274 (0.1083)				
(\geq Median)	0.0411 (0.0413)	0.1510 (0.1115)	-0.0402 (0.0387)	0.0813 (0.1857)	-0.0464 (0.0383)	-0.2770 (0.2163)
AI \times (\geq Median)	-0.0376 (0.0574)	0.0896 (0.1537)				
Subsample	All	All	AI	AI	AI	AI
Observations	634	634	319	319	319	319

Heteroskedasticity-robust standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A21: Regressions showing that the content of the AI advice impacted the types of business changes made by an entrepreneur. Panel A shows that words that appear in the AI’s advice text increase the chance the word appears in the entrepreneur’s description of their business changes. Panel B shows that when the value of a word embedding element increases in the AI’s advice text the value of the same element increases for the entrepreneur’s business changes text. [Figure A11](#) shows how controlling for the content of the question ask can allow us to recover the causal impact of the advice text on business changes.

Panel A: Word-entrepreneur level

Dependent Variable: Model:	(1)	(2)	In Changes Text			
			(3)	(4)	(5)	(6)
Constant	0.1844*** (0.0126)	0.1740*** (0.0123)				
In AI Answer Text	4.175*** (0.3971)	2.836*** (0.2497)	3.017*** (0.2522)	2.649*** (0.2635)	2.649*** (0.2639)	3.417*** (0.3394)
In Question Text		15.48*** (1.717)	15.55*** (1.700)	15.40*** (1.962)	15.40*** (1.965)	15.62*** (2.025)
Answer \times Above Median				0.7678** (0.3365)		
Question \times Above Median				0.2110 (2.088)		
Performers Subsample	All	All	All	All	Low	High
<i>Fixed-effects</i>						
Entrepreneur			Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations	4,071,078	4,071,078	4,071,078	4,071,078	2,054,682	2,016,396

Panel B: Embedding-entrepreneur level

Dependent Variable: Model:	(1)	(2)	Changes Embedding Value			
			(3)	(4)	(5)	(6)
Constant	0.0005 (0.0107)	0.0005 (0.0086)				
Answer Embedding Value	0.2631*** (0.0156)	0.1274*** (0.0102)	0.1274*** (0.0102)	0.1266*** (0.0113)	0.1266*** (0.0113)	0.1280*** (0.0115)
Question Embedding Value		0.3450*** (0.0127)	0.3451*** (0.0127)	0.3320*** (0.0156)	0.3320*** (0.0156)	0.3588*** (0.0152)
Answer \times Above Median				0.0014 (0.0101)		
Question \times Above Median				0.0267 (0.0175)		
Performers Subsample	All	All	All	All	Low	High
<i>Fixed-effects</i>						
Entrepreneur			Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations	731,136	728,064	728,064	728,064	371,712	356,352

Clustered (Entrepreneur & Embedding Dimension) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A22: Panel A and B regress whether an entrepreneur's business change text matches our regular expression patterns capturing whether using the AI assistant or business training guide is mentioned. Panel C is similar but for whether the text matches a regular expression capturing whether the entrepreneur engaged in discounting or advertising.

Panel A: Mentions the AI?

Dependent Variable:	Thanks AI Mentor?			
Model:	(1)	(2)	(3)	(4)
Constant	0.6349 (0.4482)	0.6410 (0.6410)	0.6410 (0.6410)	0.6289 (0.6289)
AI	6.262*** (1.490)	1.843 (1.388)	1.843 (1.388)	10.76*** (2.612)
(\geq Median)		-0.0121 (0.8980)		
(\geq Median) \times AI		8.920*** (2.958)		
Subsample	All	All	(<Median)	(\geq Median)
Observations	634	634	317	317

Panel B: Mentions the Business Guide?

Dependent Variable:	Thanks ILO Guide?			
Model:	(1)	(2)	(3)	(4)
Constant	0.6349 (0.4482)	0.6410 (0.6410)	0.6410 (0.6410)	0.6289 (0.6289)
AI	-0.0080 (0.6300)	-0.6410 (0.6410)	-0.6410 (0.6410)	0.6369 (1.092)
(\geq Median)		-0.0121 (0.8980)		
(\geq Median) \times AI		1.278 (1.266)		
Subsample	All	All	(\geq Median)	(\geq Median)
Observations	634	634	317	317

Panel C: Discounting or Advertising?

Dependent Variable:	Discounts or Ads?			
Model:	(1)	(2)	(3)	(4)
Constant	45.71*** (2.811)	39.10*** (3.919)	39.10*** (3.919)	52.20*** (3.974)
AI	5.069 (3.970)	11.83** (5.566)	11.83** (5.566)	-1.568 (5.631)
(\geq Median)		13.10** (5.582)		
(\geq Median) \times AI		-13.40* (7.918)		
Subsample	All	All	(<Median)	(\geq Median)
Observations	634	634	317	317

Heteroskedasticity-robust standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A23: Poisson regressions showing that high performers using AI use more uncommon tailored words when describing their business changes. “detergent,” “headphones,” “peanut,” “rabbit” and “uganda.” Common words include “advertising,” “customers,” “market,” “price,” and “vegetables.” Given that entrepreneurs who write more will mechanically have more uncommon words, we control for total number of words used. We also focus on the intensive margin, restricting our sample to entrepreneurs who use at least one word in their business changes text for which we can calculate whether it is uncommon or not.

Dependent Variable:		Num. Uncommon Words		
Model:	(1)	(2)	(3)	(4)
Constant	1.055*** (0.0593)	1.097*** (0.0766)	0.8569*** (0.0934)	1.136*** (0.0668)
AI	0.0953* (0.0574)	-0.0360 (0.0815)	-0.0148 (0.0829)	0.1805*** (0.0686)
(≥Median)		-0.0560 (0.0867)		
(≥Median) × AI		0.2234* (0.1160)		
Subsample	All	All	(≥Median)	(≥Median)
Controls	Num. Words	Num. Words	Num. Words	Num. Words
Observations	537	537	265	272

Poisson regressions.

Heteroskedasticity-robust standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A24: Example pairs of business changes whose differences illustrate the semantic differences captured by our treatment effect embedding for high performers

Control Example	AI Treated Example
Chicken Example Similarity to high and low treatment effect embeddings: 3.06, -0.07	
<p>The biggest change I made was advertising through the internet and also doing deliveries. Talking to customers who want one month chicks for them to make their orders before I go for the one day chicks. I have decided to rare one day old chicks and sell them at one month making a good profit from them. Making delivery of the one month chicks to the customer preference and also if the customer wants more older chicks at a fee. Advertising my business on social customer including whatsapp Facebook because they have connected me to many new customers. Instead of buying one day chicks I have use an incubator to cut the cost of buying the one day old chicks. I have started keeping improved chicks for sale after one month because I found its at that age that most customers want them. Keeping and rearing chicks for one month and then selling them out on profit. The social media advertising and also pay on delivery mode. I have started using mostly social media to advertise my product hence reaching many customers. I have introduced a new half kg package to my honey business because customers have been asking if there is half kg packaging. I have decided that I will be delivered order to the customer at a small fee.</p>	<p>Improving the quality of chicks in my chicken roster has brought high quality chicken to my folder. I can now sell a chick that used to go for kshs 150 at 250 to 300 and this is a very great opportunity for me. I ma grateful for the Chat AI. I have used WhatsApp regularly to get information about market and business ideas. I have bought new chicks that were recommended by whe WhatsApp AI. I have regularly used bottles to drip feed my chicks. Selling of broilers and chick mash has been of great impact to my business. I don't rely on the chick and egg sales anymore, but get income from selling broilers and chick mash. Reducing slightly the prices of my commodities to help tilt the competition in my favour. I have started selling broilers and chick mash. Providing chick vaccination to farmers.</p>
Food Example Similarity to high and low treatment effect embeddings: 3.67, 0.7	
<p>Baking business. I've started baking honey filled muffins. Honey baked muffins. Home delivery of my products. Baking and supply of honey baked muffins.</p>	<p>Use of new technology of AI mentor. By doing more benchmarking on the competitors. Cooking oil, cereals. Use of AI mentor.</p>
Bulk Example Similarity to high and low treatment effect embeddings: 3.5, 1.39	
<p>Adding more products into the business. Delivering gas stove door to door. Oreo biscuits. Selling cakes in wholesale. Delivering. Refilling empty gas container. Delivering door to door. Advertising what am selling. The product is called bourbon. Delivering door to door services.</p>	<p>It has boost on sales and customer satisfaction. Improved on customer care skills after getting some tips and advice from AImentor and introduction of discounts to customers who purchased the products in bulk. These strategies have helped in customer loyalty. Introduction of groceries products into the business to reduce inconveniencies to my clients. Yes, it has attracted new customers hence increase in sales and profit in the business. Introduced a new product line i.e selling groceries to reduce inconvenience to my clients.</p>
Power Example Similarity to high and low treatment effect embeddings: 3.32, -0.4	
<p>I have started going to the customer instead of waiting for them to come to me. I have started selling hair bonnets. I have started a delivery service. Delivery of products to customers has really made a difference in my business. Making myself available at any time for customers. I've started selling hair bonnets. Delivery.</p>	<p>AI teachings. Not Applicable. Saloon. Barbering. Use of AI to give me advice when needed. I have inserted a small salon beside my cyber to get more customers and profit. Use of solar when there's no electricity. Salon. A solon for women hair preparations. I got advice from AI mentor on other ways of getting power when there is electricity power black out. Having a salon in my barber shop and people being massaged after services. Selling magazines. Barbecue. Barbecue</p>

Control Example	AI Treated Example
Carwash Example Similarity to high and low treatment effect embeddings: 2.75, 1.84	
The use of new computers which has made my work really easy. I've been making my own products like soaps and detergents which is making profits for my business. I tried some new body oils. I've tried door to door delivery. Selling of soaps and detergents. I have been getting huge profits from my new business since I'm doing everything by myself. Selling of the soaps and detergents. Making of soaps and detergents. Door to door delivery. Making of soaps and detergents	I employed someone, for the smooth runnings. I gave my customers special offers, when they wash their vehicles three times in a week, I wash for the free during the weekend. I bought a coloured printer and replaced the black and white one. I expanded my car wash, added a cyber a place for soft drinks thanks to the AI tool you gave me. Trying a new washing detergent. I opened my own till numbers so that customers can pay using mobile money. There is this new car washing detergent that I tried and the customers love it. We have been doing deliveries of the washed cars to their owners.
AI Market Research Example Similarity to high and low treatment effect embeddings: 2.64, -0.08	
Introduction of goods delivery. Initially, clients would pick their goods at the shop but now, they are delivered at their destinations t. Using Facebook ads to advertise the business. Initially i was selling cooking pots but i introduced other kitchen ware sets like cups and plates. Delivery of goods to customers. Engaging clients. Use of media in advertising. Male clothes. Delivery of products.	The incorporation of repairs in my shop has attracted more customers increasing both my income and sales. With the help of AI, I have been able to do more market research and discovered new ways of advertising my products. I have added cellphones, speakers and electric wires to my business. I have started electronics repairs e.g. Phones and TVs alongside my business. The sale and refill of gas cylinders has really added a fortune to my operation. I have added the sale of Gas cylinders in my shop. I have started free delivery for Items worth 5000 and above.
Cyber shop discount buying Example Similarity to high and low treatment effect embeddings: 2.44, -0.04	
Having own website which people can order and have goods delivered in their door steps. Created a website where people can order from. Selling of electronics. Home delivery. Doing home delivery has attracted customers and having website where customers can order from the comfort of their houses. Having a website to display products online and from where customers can place their order and have the goods delivered at their doorsteps. Selling electronics. Having home delivery using a motorbike. Home delivery , the number of customers ordering products to be delivered at their doorsteps have increased. Having business website where customers order for products pay through mobile money and have delivered at doorsteps. Business website where business products displayed online. Home delivery , motorbike that delivers at customers doorstep.	Buying in bulk saves input in resources. I now buy stock in bulk at the least price possible. Entertainment gadgets. Provide entertainment services for adults and children. Negotiating with suppliers which has led to lowering of prices increasing number of customers. I sell goods bought in bulk at a lower price. I have begun renting out gaming accessories in my cyber cafe' following advice from your AI. I have started offering cashbacks to loyal customers. I have set up a gaming arcade in town. I charge clients to play in gaming simulators. With proper advice, I begun with purchasing only pairs of accessories. Now my clients have enough gaming equipment to engage in. Entertainment arcade and gaming.

Control Example	AI Treated Example
Location Change Example Similarity to high and low treatment effect embeddings: 2.81, 0.24	
<p>I have changed location, i was not making maximum profits because the place i was in was a street with few people. Advertising on social media. I started the perfume refill business. Providing education consultancy. Introducing a new product. I have decided to advertise and sell my products on tiktok and facebook. I was selling perfumes , refillables only..i decided to include designer perfumes.</p>	<p>Business strategic.I change the location. I change the location to more suitable place. Technology. Marketing policy. Value addition in help of chat bot advice. Value addition of my products to meet customer needs. Chat bot table very resourceful. Familiarizing with customers and looking for demand curves.</p>
Gas Cylinders Example Similarity to high and low treatment effect embeddings: 2.34, 0.16	
<p>Home delivery has made it easier for customers with a busy schedule. Interact with clients to know the trend that is in the market that they like. Started selling clothing. Started doing home delivery to customers that are far. Online business has really impacted my business positively, got a number of potential clients. Installed CCTV in my business premise to help with security. I have become a mobile money agent too. I offer cash/cashless money transfer. I'm a mobile money agent now too.</p>	<p>Bringing government digital services(E-CITIZEN) has greatly increased customer traffic in my business premises which has had a direct positive impact in sales volume. I've brought ChatGPT on board cause of it's immense importance especially with regards to a wide array of information it offers. I have started retailing LPG gas cylinders. I've incorporated government digital services(E-CITIZEN) like filing tax returns, huduma services etc. The introduction of government digital services (E-CITIZEN) has greatly increased customer traffic in my business premises which has translated to more sales across my business chain. * Courting new customers with discounted price for every purchase they make on the first day. I commenced selling of LPG gas cylinders alongside my M-PESA business. I've brought on board Government digital services (E-CITIZEN) like renewal of certificate of good conduct, driving licenses etc. Retailing of LPG gas cylinders. Onboarding government digital services (E-CITIZEN) has really increased customer traffic in my business premises translating into more sales. Offering first time customers products on discounted prices has really proved worthwhile. *I've partnered with 'SUNKING' a solar products dealer to sell their merchandise. *I've started (on a pilot basis) providing government digital services (E-CITIZEN) to the public.</p>

Table A25: Example pairs of business changes whose differences illustrate the semantic differences captured by our treatment effect embedding for low performers

Control Example	AI Treated Example
Lower Price Example Similarity to high and low treatment effect embeddings: 0.84, 3.75	
Increase in profit in the business. Sales of eggs. Marketing my goods online. Selling of clothes. Transport service. Selling of clothes.	Giving so offers by reducing the selling price. To advertise in social media. Increase of sales. Advertising on social media platforms. Oraimo brand charger. Giving offers to some products.
Marketing on WhatsApp Example Similarity to high and low treatment effect embeddings: -0.11, 2.94	
The new ways of watering has enabled most of my crops not to get dried up and hence selling more seedlings thus increasing the profits. My business strategies are like engaging my consumers on the best foods I sell to them. I have implemented new cans of watering my crops. I have implemented new services like mulching to prevent crops from drying up. The biggest change is applying shade to my newly seedlings in the seedbed. This have helped them from drying up by preventing direct sunlight. My new business strategies are like carrying out crop rotation in my farm. This will help in increasing crop production. No have implemented new products like planting seeds. I have implemented new services like methods of preserving harvested products. My new activities are by cooking and selling some of the products of my businesses. Opening early and closing late. Opening businesses as early as possible so as to attend to all my customers. The new products include fork jembes and panga. Selling by means of digital or online. Tailoring	Advertising my products through WhatsApp has increased number of customers buying my products. I have ensure that my customers never leave I usually do follow up in order to make sure my customers are satisfied. I have advertising my products through WhatsApp and Facebook. Giving some extra discount on my customers and also giving some products offer to my customers increase the number of my customers. I have used WhatsApp and Facebook to market my products and adding some small offers to my customers. I have realized that most of my customers required something to eat while in my market so I have brought some soda to bread for my customers to eat. I have introduce some extra appreciation offer for a customer buying my products in bulky. Door delivery products for my customers who gives orders through phone. Selling my products on my special offers. Giving special offers To my Royal customers. I have introduce black polythene paper into my market which many customers required. Providing customers offer when buy products to certain amount.
Marketing vs. ads Example Similarity to high and low treatment effect embeddings: 1.12, 3.33	
Online marketing. Getting to know what customers wants, precise record keeping and online marketing. Groceries. Cyber services. Reduced owing to reduce debts. Precise record keeping, balance sheet of the business. Cosmetics. Printing services. Marketing. Good marketing. Groceries. Providing loan at an interest.	Advertisement and discounts. Offering after sales service. Training of staff. Training of staff. Discounts. Offering discounts.
More price discounts Example Similarity to high and low treatment effect embeddings: -0.49, 2.65	
Budgeting and data keeping. Budgeting and data keeping. Fruits. Mpesa services. Increased profits because I can now manage cashflow appropriately without possible losses. Planning and budgeting. Maize. Mpesa services. Fruit selling. I introduced selling of gapes ,oranges at my shop and it's adding up. Budgeting. Budgeting , planning , Recording. Mpesa services.	Posting online has brought more customers. A discount for items above ksh 1000. I started selling shoes in my clothes shop. Posting online my products. Afflicted customers. Afflicted customers. Mpesa alongside my shop. Discount for items above 1000. Offers for goods above 1500. Offers for goods above 1500. Selling shoes in my mtumba warehouse. Posting it on social media. Cereal outside my shop
Yet more ads and discounts 1 Example Similarity to high and low treatment effect embeddings: 0.45, 2.44	
Consulting my customers. grocery. selling groceries. using of technology. selling of cosmetics products. Cosmetics. perfumes,. Grocery. Cosmetic shop. dealings with cosmetics products.	Isuing discounts on some products. Painting my shop as advertising strategy. Discount. Packaging.

Control Example	AI Treated Example
Yet more ads and discounts 2 Example Similarity to high and low treatment effect embeddings: 0.48, 2.61	
Collection not used oil and reselling it. Selling used oil. Advertising the product. That business has a lot of profit and not so many people have ventured into it. Selling of used oil. This oil is been used by. People to treat timber and use it in powersaw. Middleman in buying and selling of secondhand. Farm implements and tractors. Stocking the products. Selling of waste oil.	Introducing discounts. Selling them in affordable prices. Seats. Comfortable cotton seats. Branding my products. Customer satisfaction through discounts. Electronics. Aftersales services. To satisfy customers and offer best services to them. Offering best discounts. Opening of more branches. Wireless CCTV cameras. Free transportation. To empower customers to buy your products
More free stuff Example Similarity to high and low treatment effect embeddings: -0.81, 2.37	
The mitumbas business. I have started selling materials for making vitenge. And including mitumbas. Inclusion of mitumba. To include selling shoes. Including mitumbas in my boutique. Bodaboda service. Including mitumbas in my business. Including mitumbas in shoes selling. Selling credit. Started layering poultry	The free delivery for my products is really working out for me. To work within the budget. I have tried the new enriched Amara oil. I have offered free delivery for my customers. Free delivery services to customers within range. I have offered free delivery for my clients. New Imperial leather for men. Hair dressing. It is a gaming house containing a play station 5 and Xbox. Free delivery services to my customers. Free delivery services. New saving gel. Free delivery services. Rearing and selling chicken
Milk Example Similarity to high and low treatment effect embeddings: 0.97, 2.58	
Selling of milk is highly profitable and increasing of products in general has made my business grow. Selling milk. Selling of rice. Selling firewood. Selling of milk.	Advertising the new products in my shop. Selling my products at a relatively lower prices and offering discounts to my loyal customers. I also made efforts to advertise my products. I started selling eggs in my shop and found that the demand was appealing, I'm yet to add more stock to gain more. Customer relations and offering discounts to the goods I sell. General shop. Discounts usually encourage my customers to be back and back to my business. Giving discounts to all my sales. I introduced milk sale in my business. Considered customer feedback, I engaged some of my loyal customers to know more about services I usually offer. I started to sell milk alongside my business I'm now running both
Billboard Example Similarity to high and low treatment effect embeddings: 0.04, 2.35	
Including some of the potential customers and suppliers in decision making, also acquiring a loan to expand the business. Selling of eggs both in wholesale sale and retail. Mpesa shop. Involving workers in decision making and also being consultant. Hardware stores. Selling cereals. Mpesa. It's making profits and Also establishing new customers. Making New customers and increasing sales Also making new profits	The use of billboard for advertising my business i.e products and services. I used a billboard to advertise my business which was attracted by many. The use of a different styling gel brand have been liked by most of my clients. I learned about a new hairstyle and discovered that it is loved by many customers. Offering discounts to my customers. I have been offering discounts to my customers. I have tried using different hair foods and am good with my customers. I have used different hair food for blowdrying and customers have liked it.

Table A26: Topic Counts

Phrase	Count	Phrase	Count
Community Engagement	55	Executive Summary	9
Market Research	49	Feedback Mechanism	9
Poultry Farming	42	Fish Farming	9
Loyalty Program	39	Food Business	9
Inventory Management	31	Friends and Family	9
Customer Feedback	29	Market Demand	9
Location	28	Marketing & Sales Strategy	9
Value Addition	27	Online Presence	9
Diversify Your Product Range	25	Quality Assurance	9
Leverage Social Media	25	Risk Management	9
Partnerships	25	Second-hand Clothes Business (Mitumba)	9
Pricing Strategy	25	Soap Making Business	9
Loyalty Programs	24	Customer Loyalty Programs	8
Crowdfunding	23	Dropshipping	8
Diversification	21	Education	8
Customer Service	19	Freelance Services	8
Marketing Strategy	19	Laundry Services	8
Budgeting	18	Licenses and Permits	8
Improve Customer Service	18	Location Selection	8
Local Partnerships	18	Networking	8
Offer Value-Added Services	18	Personal Savings	8
Partnerships with Local Businesses	18	Promotions & Discounts	8
Social Media Marketing	17	Understand Your Customers	8
Social Media Presence	17	Vaccination	8
Customer Loyalty Program	16	Word-of-Mouth Marketing	8
Referral Program	16	Breed Selection	7
Agriculture	15	Cleaning Services	7
Business Plan	15	Educate Yourself	7
Improve Your Marketing	15	Feeding	7
Microfinance Institutions	14	Freelancing	7
Optimize Operations	14	Grants and Competitions	7
Savings	14	Healthcare	7
Create Customer Loyalty Programs	13	Improve Product Display	7
Online Tutoring	13	Laundry Service	7
Competition	12	Marketing	7
Crop Rotation	12	Offer Excellent Customer Service	7
Promote Your Business Locally	12	Offer Promotions	7
Community Involvement	11	Promotion	7
Competitor Analysis	11	Provide Excellent Customer Service	7
Legal Requirements	11	Regular Communication	7
Negotiate with Suppliers	11	Rent	7
Bulk Purchasing	10	Seek Professional Help	7
Market Analysis	10	Social Media Management	7
Microfinance Institutions (MFIs)	10	Start Small	7
Nutrition	10	Supplier Relationships	7
Promotions	10	Training	7
Record Keeping	10	Tutoring Services	7
Affiliate Marketing	9	Affordable Pricing	6
Email Marketing	9	Cash Flow Management	6
Emergency Fund	9	Choose the Right Breed	6

Phrase	Count	Phrase	Count
Cleanliness	6	Regular Updates	5
Company Description	6	Regulatory Compliance	5
Consistency	6	Retail Shop	5
Content Creation	6	Seeds	5
Cost Control	6	Set Clear Goals	5
Cost Reduction	6	Skills	5
Customer Referral Program	6	Soap Making	5
Customer Relationship Management (CRM)	6	Training & Education	5
Customer Retention	6	Transparency	5
Customer Service Excellence	6	Vaccination Program	5
Develop a Business Plan	6	Ventilation	5
Diversify Your Products	6	[Your Name]	5
Employee Training	6	Additional Services	4
Equipment	6	Adjust Your Plan as Needed	4
Excellent Customer Service	6	Ask for Feedback	4
Financial Projections	6	Biosecurity Measures	4
Health Management	6	Bundling	4
Market Trends	6	Business Registration	4
Regular Audits	6	Buy in Bulk	4
Separate Business and Personal Finances	6	Chamas (Investment Groups)	4
Staff Training	6	Clean Environment	4
Unique Selling Proposition (USP)	6	Comfortable Environment	4
Agribusiness	5	Competitive Pricing Strategy	4
Bootstrapping	5	Conduct a SWOT Analysis	4
Bundling Products	5	Consult a Vet	4
Clear Communication	5	Customer Education	4
Communication	5	Customer Engagement	4
Competitive Pricing	5	Demand Forecasting	4
Contact Information	5	Digital Marketing	4
Dairy Farming	5	Direct Selling	4
Direct Marketing	5	Disease Management	4
Enhance Customer Service	5	Diversify Income Streams	4
Freelance Writing	5	Diversify Services	4
Government Programs	5	Educate Customers	4
Identify Your Target Market	5	Employee Engagement	4
Improve Visibility	5	Exceptional Customer Service	4
Irrigation	5	Fruit & Vegetable Stall	4
Labour	5	Fruit & Vegetable Stand	4
Land Preparation	5	Government Funds	4
Loans	5	Grower Feed	4
Mobile Food Cart	5	Improve Marketing Efforts	4
Mobile Money Agent	5	Improve Product Presentation	4
Observation	5	Leverage Technology	4
Offer Competitive Pricing	5	Licensing	4
Offer Delivery Services	5	Local Marketing	4
Operating Expenses	5	Marketing and Promotion	4
Organization & Management Structure	5	Mentorship Programs	4
Pest and Disease Management	5	Monitor Progress Regularly	4
Product Bundling	5	Objective	4
Quality Control	5	Offer Additional Services	4

C Experimental Design and Timeline

Our field experiment lasted approximately five months, and involved 640 Kenyan entrepreneurs and seven separate performance surveys. It was structured in three stages:

1. **Recruitment and pre-treatment surveys:** Initially, entrepreneurs were recruited online and screened using three pre-treatment surveys.
2. **Assignment to interventions:** Participants were then assigned either to an AI assistantship program or to receive a standard business training guide.
3. **Performance tracking:** Finally, we assessed each participant’s subsequent business performance through four post-treatment surveys.

We provide details on these stages below.

Stage 1: Recruitment and pre-treatment surveys

Participant recruitment took place over the Meta ad platform (Facebook penetration in Kenya is over 80% ([Wamuyu, 2020](#))). A partner survey organization created advertisements offering a paid survey targeting entrepreneurs and businesses. The advertisements were shown throughout Kenya and the only restriction on the advertisement was it was only to be shown to individuals over the age of 18 (see [Figure A1](#)).

Individuals who clicked on our ad were directed to a Qualtrics survey, which was the first of three pre-treatment surveys we conducted with participants. These pre-treatment surveys aimed to screen out individuals: (i) who did not own and operate enterprises; (ii) who would be likely to attrit over the long time horizon of our study; (iii) who were inattentive respondents.

Our analytic sample predominantly consists of relatively young businesses, with 61% having been in operation for one year or less, and only 16% having been operating for over three years. The businesses were primarily concentrated in three sectors: Food and Beverage (27%), Clothing (20.8%), and Agriculture (18.9%). A majority of these businesses (64%) operated without any employees. Pre-treatment the median firm in our sample earned a monthly profit of 14,000 Kenyan Shillings (about \$90 USD) and had a monthly revenue of 30,000 Kenyan Shillings (about \$190 USD).

Participants came from 44 of Kenya’s 47 counties, with the largest number of participants coming from Nairobi (133 participants), Kiambu (53 participants), and Nakuru (47 participants). Nairobi’s 4-million-person capitol has the highest concentration of study participants (20%). Participants in our sample generally exhibited a high level of education, with only 10% not having any college education. The sample demonstrated a low level of gender diversity, with women accounting for only 32% of the participants, although this may in part reflect a lower level of business ownership among women in Kenya ([The World Bank Group, 2018](#)). Additionally, the average age of participants in our study was 26 years old. See [Table A1](#) for additional details on our sample.

Participants who completed the first survey were contacted via SMS messages through a “Shortcode” (a shortened phone number used to bulk send text messages) and also over email inviting them to participate in a second survey. These individuals were sent invitations to a second survey via a personalized Qualtrics link that could only be used one time.⁸ Those completing the second survey were sent a third pre-treatment survey over WhatsApp. Following the completion of the pre-treatment surveys but prior to random assignment, we pre-registered the following exclusion criteria:

1. Participants who do not run a business or do not state that they are willing to take subsequent surveys
2. Participants who tried to take any Qualtrics survey more than once, as identified by Qualtrics duplicate score or by duplicated phone numbers or emails
3. Participants who fail the following attention check:

The next question is about the following problem. In questionnaires like ours, sometimes there are participants who do not carefully read the questions and just quickly click through the survey. To show that you read our questions carefully, please enter orange as your answer to the next question. What is your favorite color? {red; yellow; green; blue; black; orange; white; purple; pink}

4. Participants who stated that their weekly profits (revenue) were greater than their monthly profits (revenue), or who stated that their profits were greater than revenue (separately for weekly and monthly measures).

⁸All subsequent survey waves also used similar one-time-use personalized Qualtrics links.

Stage 2: Treatment

After the third pre-treatment survey wave, we randomly assigned participants to one of the following experimental conditions.

1. **AI-mentor.** Participants assigned to the AI-mentorship condition received access to our WhatsApp-based AI-mentor. The AI assistant was powered by GPT-4 and designed to give business mentorship and advice to Kenyan small and medium businesses.
2. **Control.** Participants assigned to the control condition received access to widely used business training materials developed for small and medium businesses by the International Labour Organization.

Randomization was stratified on Gender \times (Quartile of pre-treatment business performance), for a total of $2 \times 4 = 8$ strata.

Onboarding survey. Before participants gained access to the business guide or the AI assistant, they were required to complete a brief *Onboarding Survey*. The first page of the onboarding survey provided a concise summary of the expected time to complete the survey and the mobile airtime compensation they would receive upon completion, which offset the data costs of taking the online survey. Notably, the first page did not disclose the participant’s treatment group. Participants who did not proceed beyond the first page, where treatment status was undisclosed, are considered pre-treatment dropouts. Our analytic sample is comprised of 640 participants who advanced to the second page of the Training Survey, where they were informed of their treatment status. These participants had already successfully completed three rounds of pre-treatment surveys and passed pre-registered attention checks.

Stage 3: Post-treatment surveys

Following treatment, we conducted four additional survey waves in which participants were contacted via SMS messages. Each of these waves contained a battery of questions exploring mechanisms in addition to our set of performance measures. See [Appendix E](#) for an overview of the contents of each survey.

Survey Incentives

Participants completing the first survey were sent a small airtime transfer of KSH 30 as a participation incentive. These credits could be used to purchase phone credits and cellular data to access the internet. To reduce attrition we used an increasing survey incentive scheme: The second pre-treatment survey paid KSH 50, and the third paid KSH 100. All post-treatment surveys paid KSH 200.

Attrition

Our analytic sample, defined by pre-registered exclusion criteria (see [Appendix C](#)) is comprised of 640 participants. A total of 18 (3%) of participants attrit from the survey post-treatment (i.e., did not complete our final survey, but may have completed at least one pre-treatment survey).⁹ The final row of [Table A1](#) reports the final sample size by experimental condition. In total 11 participants attrited from the “AI” condition and 7 attrited from the control condition. The difference in attrition across conditions is 1.2 pp ($p = 0.35$).

⁹Studies such as [Banerjee et al. \(2015\)](#) report attrition rates of approximately 9%.

D Detailed Intervention Overview

AI assistant

Our AI assistant was powered by the GPT-4 and Whatsapp APIs. WhatsApp has very high adoption in Kenya (e.g., [Wamuyu \(2020\)](#) report that 89% of Kenyans use WhatsApp and 81.7% use Facebook). Participants interacted with the AI assistant over WhatsApp by texting our WhatsApp number. While the number was public (anyone can text the number), we programmed the AI assistant only to respond to participants on a designated access list. Thus, there was no risk that control participants could interact with the AI assistant. The AI assistant used the standard GPT-4 corpus and was contextually tailored using prompt engineering alone. All parameter values (e.g., “temperature”) were set to GPT-4’s default values.

Development of the AI assistant

We used an iterative process to develop our AI assistant over the course of several months. This process was structured into three, interconnected phases.

1. In the first phase, we worked with a team of research assistants to craft prompts aimed at providing strategic advice to Kenyan entrepreneurs. This involved a two-step process: first, a team member posed basic business questions to GPT-4. Then, another evaluated the responses for quality and relevance. We cycled through this process repeatedly, refining the prompts in each cycle based on the feedback received in the cycle prior.
2. The second stage involved a team member manually forwarding questions asked by a pilot sample of Kenyan entrepreneurs to GPT-4. As part of this process, our team and the sample of Kenyan entrepreneurs evaluated the quality and form of the GPT-produced responses. As above, we progressively revised our system prompt based on feedback.
3. The third and final cycle took place after we had built an early version of the AI assistant that was able to communicate with participants over WhatsApp. Here, we recruited a new sample of Kenyan pilot participants who were given early access to

the AI assistant. In addition to asking questions directly to AI assistant, these pilot participants completed surveys that captured more detailed feedback on the tool, which we again used to iteratively revise our system prompt.

Based on this process, our final product was an AI assistant instructed to adhere to the following guidelines:

- For each query, the AI assistant provided 3 to 5 tailored, actionable, enumerated strategies.
- After each response, the AI assistant prompted further engagement with: “Enter a number to explore more about one of these points on [topic name], or ask a new question,” ensuring conversational clarity and flow.
- When a user replied with a number, indicating interest in a specific strategy, the AI assistant offered 3 to 5 additional actionable tips related to that strategy.

The AI assistant was endowed with the personality of “an expert business mentor specialized in guiding Kenyan entrepreneurs with limited resources” whose goal was to “improve business performance and profitability by providing actionable, contextually relevant, specific, detailed advice that produces sustained improvements in competitive advantage.”

AI assistant System Prompt

Below is the final version of the system prompt that we provide to GPT-4 in order to create the AI business mentor:

As an expert business mentor, your specialization is guiding Kenyan entrepreneurs with limited resources. Your mission is to improve business performance and profitability by providing actionable, contextually relevant, specific, detailed advice that produces sustained improvements in competitive advantage. Interactions follow these steps:

- [1] When a question is asked, respond by learning about specific business needs relating to the business location, type, etc. Then, provide 3 to 5 actionable, easy-to-implement strategies tailored to the business in question. Offer a clear explanation for each strategy, including benefits & step-by-step guidance on implementation. Enumerate these options with [#].
- [2] After each message, prompt further discussion with the statement: "Enter a number to learn more about one of these points on **topic name** or ask a new question." Offer a clear explanation for each tip, including the benefits & how to implement them.
- [3] If a number is given in the response, it indicates the desire to explore the corresponding strategy further. Offer 3 to 5 additional actionable tips related to the chosen strategy. Enumerate these options with [#].
- [4] Again, after each message, prompt further discussion with the statement: "Enter a number to learn more about one of these points on **(bold topic name)** or ask a new question."

Communicate in the language of inquiry. Use simple non-technical English. **Bold** key advice and message topics. Avoid high-tech or costly solutions as your interlocutor is a Kenyan entrepreneur with limited resources. Avoid jargon. As a business mentor, do not discuss health, religion, politics, or current events. Only discuss business topics.

The most critical rule to remember: Whenever a number is entered, ALWAYS interpret it in reference to your last message. E.g., if the user responds "3", it refers to point "3" in your most recent (last) message. If the user responds "2", it refers to point "2" in your most recent message.

ILO guide

In the control group participants were provided with a guide developed by the International Labor Organization. Specifically, we texted (over WhatsApp) the first three PDFs (out of a series of 9 PDFs in the guide) to participants. We also provided a link where participants could access the other PDFs. The full series of ILO guide PDFs can be accessed at https://www.ilo.org/empent/areas/start-and-improve-your-business/WCMS_192062/lang--en/index.htm.

Training Materials

AI assistant Training

Here at O-Survey, we have built a special robot that can give you advice to help you with your business. This AI-powered robot, or “AI assistant chatbot” is designed to act like a business mentor that you can chat with any time you want. It provides custom advice to improve the performance of your business. We are going to give you access to the AI assistant chatbot for free over WhatsApp. You can interact with the AI assistant chatbot on the same WhatsApp number that sent you this survey. You may find it helpful to chat with this AI assistant at least a few times a week (you can talk to it more if you want). We will send you a message over WhatsApp in 1-2 days when you can begin talking with the AI assistant. Now you will spend 5-10 minutes learning about how to make the best use of the AI assistant chatbot.

Here are a few things you should know about the AI assistant chatbot:

- The advice is provided by an AI system (not by a human mentor). Sometimes, the advice may be incorrect or bad. Make sure to think about the advice before you use it.
- The AI assistant chatbot doesn’t know what’s happening in the world right now. So, you can’t ask it things like “How hot was it yesterday?”
- Sometimes, it might take up to 2 minutes for the AI assistant chatbot to reply. Just wait a bit before asking another question.

Here are some tips to get the best advice from our AI assistant chatbot:

1) Be detailed: The more you tell the AI assistant chatbot about your situation, the better it can help. For example, if you want to make more money as a farmer, don’t just ask “How can I make more money?” Instead, tell it about your farm, what crops you’re growing, and any problems you’re facing, like this: “I’m a sugarcane farmer in Nyando county with a 2-acre plot. We’ve been having a drought that’s hurting my crops and profits. How can I make more money without spending more?”

Here is another example of a good question: “I own a small hair salon in Nairobi with about 2-4 customers a day. I am trying to find ways to promote my business. I already put up posters near my shop. What are some other ways that I could gain new customers and get

the word out about my business?”

This question is good because it provides details on who is asking the question (“I own a small hair salon in Nairobi with about 2-4 clients a day”), and it provides details on the problem and what you have already tried (“I already put up posters near my shop”). Then it ends with the question (“What are some other ways that I could gain new customers and get the word out about my business?”).

2) Ask follow-up questions: When the AI assistant chatbot gives you advice, you can ask follow-up questions. Here is an example response from the AI assistant chatbot to the previous question about promoting a business:

Question: I own a small hair salon in Nairobi with about 2-4 clients a day. I am trying to find ways to promote my business. I already put up posters near my shop. What are some other ways that I could gain new customers and get the word out about my business?

Answer: “Here are three practical strategies:

- Leverage social media: Create accounts on popular platforms like Facebook or Instagram. Regularly post attractive pictures of hairstyles, promotions, or events at your salon to engage with potential clients.
- Offer referral incentives: Encourage existing customers to refer friends by giving them discounts or other rewards for each successful referral.
- Host community events: Organize small gatherings like beauty workshops or charity events at your salon so people can experience the quality of service firsthand while contributing positively to the local community.

Please let me know if you’d like more information about any of these marketing techniques (1-3) or if you have another question related to growing your business.”

You will get the most benefit if you ask follow-up questions to the AI assistant chatbot. For example, you could say “2. What are some ways that I can use referral incentives to get new customers to my barbershop?” to learn more about the second option.

To summarize, to get the best help from our AI assistant chatbot:

- Remember to give lots of details about your situation
- Ask follow-up questions about the advice it gives.

In addition to asking followup questions, you can also ask for a new question at any time.

Remember, you will be given access to the AI assistant over WhatsApp in 1-2 days.

Control (ILO Guide) Training

We would like to help you access resources that could help your business. Specifically, we would like to provide you with business training manuals developed by the International Labor Organization (ILO). Note that we are not affiliated with the ILO.

We request that you please do not share the materials. To get the most out of these materials, you may find it helpful to review them at least a few times a week.

Now let's learn about the materials. Here is an overview of some of the materials we are providing:

1. A guide to coming up with new business ideas
2. A guide to starting a new business
3. A guide for developing a new business plan

While some of these materials are focused on new businesses, they can also be used to evaluate and develop current businesses as well. These materials will be sent to you over WhatsApp 1-2 days after completing the survey.

Now you will spend 5-10 minutes learning about how to make the best use of the materials.

1) Scan the materials: Each of the documents we will share has a “Contents” page that provides an overview of the contents offered in the specific guide. For example, the guide “Generate Your Business Idea” has the following sections:

- PART I – YOU AS AN ENTREPRENEUR
- PART II – A MARKET IS WAITING FOR YOU
- PART III – YOUR OWN BUSINESS IDEA LIST
- PART IV – THE BEST BUSINESS IDEAS FOR YOU
- PART V – YOUR OWN BUSINESS IDEA

Other guides allow you to evaluate your business ideas and more!

2) Business planning: Creating a business plan can help you manage your business, understand who your customers are, what your products or services are, and much more! Careful planning will help you get the most out of your business. The business plan includes guides on:

- Market Research
- Sales Estimation
- Product Costing
- Sales Planning
- Start-up capital

The more details of a plan you develop, the better you will be able to understand the challenges your business will face.

After reading the “Start Your Business” guide you will know how to:

- Describe the content of a Business Plan
- Consolidate your business idea
- Translate your business idea into a completed Business Plan
- Assess your readiness for starting a business

And much more! Remember, you will be sent the materials over WhatsApp in 1-2 days.

Reminders

In both experimental conditions, we provided participants with intermittent reminders to engage with the business manual or AI assistant for each condition.

E Pre-analysis Plan

E.1 Overview of Pre-registration

We pre-registered our study design and analysis on osf.io after the third pre-treatment survey. The following dimensions were pre-registered:

Outcomes: [Subsection E.3](#) outlines all of our pre-registered survey outcomes:

- (i) Performance (revenue and profits)
- (ii) Management Practices
- (iii) Time Management
- (iv) Technology use
- (v) Innovation activities
- (vi) Information: Access Stigma
- (vii) Information: Relevance
- (viii) Information: Accessibility and Quality

We also pre-registered the construction of our performance index and the level of winsorization we would apply our outcome variables. A more detailed overview of our survey measures is included below.

Statistical analysis: We pre-registered our primary ANCOVA regression specification (see below).

Heterogeneous effects: We pre-registered three dimensions with respect to which we would measure treatment effect heterogeneity:

- (i) Initial performance
- (ii) Gender
- (iii) Pre-treatment ChatGPT use

We elaborate on each component of the pre-registration below.

E.2 Econometric Strategy

Our main pre-registered estimation strategy is an ANCOVA variant of a simple OLS regression (McKenzie, 2012):

$$y_{i,t} = \alpha + \beta \text{AI}_i + \gamma y_{i,-t} + \zeta X_{i,-t} + \sum_S \theta_s \mathbb{1}(i \in s) + \sum_T \tau_j \mathbb{1}(j = t) + \varepsilon_{i,t} \quad (1)$$

where $y_{i,t}$ is an outcome variable measured for participant i at post-treatment period t . α is the constant, and $\text{AI}_i \in \{0, 1\}$ is a dichotomous dummy taking a value of 1 if i is assigned to receive access to the AI assistant. β captures the average causal effect of access to the AI assistant in post-treatment periods. $y_{i,-t}$ is the pre-treatment values of outcome y in pre-treatment periods ($-t$). $X_{i,-t}$ is a vector of baseline control variables. θ_s captures stratification dummy variables and τ_j captures post-treatment time dummy variables. $\varepsilon_{i,t}$ is the error term. Standard errors are clustered at the individual level.

Heterogeneous treatment effects: We estimate heterogeneous treatment effects for three pre-registered dimensions: gender, baseline business performance, and whether participants report frequent use of ChatGPT at baseline. In practice, we operationalized baseline business performance as a median split of a standardized performance index over the three pre-treatment periods, and we operationalized frequent ChatGPT use as whether the person uses ChatGPT at least once a week. We did not provide the heterogeneous treatment effects estimating equation in our pre-analysis plan but for clarity we report the parallel heterogeneous treatment effects estimation equation to [Equation 1](#) below:

$$y_{i,t} = \alpha + \beta_1 \text{AI}_i + \beta_2 \text{Het}_i + \beta_3 (\text{AI}_i \times \text{Het}_i) + \gamma y_{i,-t} + \zeta X_{i,-t} + \sum_S \theta_s \mathbb{1}(i \in s) + \sum_T \tau_j \mathbb{1}(j = t) + \varepsilon_{i,t} \quad (2)$$

where Het_i is the relevant dimension of heterogeneity, and β_3 captures the heterogeneous treatment effect.

Covariate adjustment: We follow a principled approach to covariate selection using the

double post-LASSO method (Belloni, Chernozhukov, and Hansen, 2014). This approach involves first running a LASSO to identify outcome-relevant covariates, and second, to identify variables predictive of treatment assignment. The union of this set of covariates is then included as controls in our regression analysis. The same double-LASSO-selected covariates are used in the average treatment effect and the heterogeneous treatment effect variations of each model. Additionally, we always control for randomization strata blocks and time period.

The following pre-treatment variables are included in our covariate selection: performance, management practices (sum), oldest business age, sector, children (count), age, education, GPT use, “Big Five” personality measures (agreeableness, conscientiousness, extroversion, openness, neuroticism), and county. See Table A5 for details covariate selection for our main regression results.

E.3 Overview of Measures

Business performance. We measure business performance using a combination of measures of revenue and profits. Following Anderson and McKenzie (2022), we apply the following transformations:

1. Winsorize the outcome at the 95th percentile (and 99th percentile in a separate specification).
2. After applying the step above, generate an average z -score following Kling, Liebman, and Katz (2007) by subtracting the mean and then dividing by the standard deviation of the winsorized vector of pre-treatment outcomes. Finally we average across the three performance measures: (a) Weekly sales; (b) Weekly profits; (c) Monthly sales; (d) Monthly profits.

Pre-treatment performance. We calculate pre-treatment performance by calculating the average value of the four performance outcomes (weekly and monthly profits and revenue) across the three pre-treatment periods. We then calculate a z -score of winsorized average performance for each of the four measures. Finally, we calculate the average of these four standardized pre-treatment performance measures. For log-like pre-treatment performance measures the first transformation applied is the inverse hyperbolic sine transformation.

Business practices. We create an average z score out of a set of sixteen management practices based on [McKenzie and Woodruff \(2017\)](#): In the following set of questions, we will ask you about certain business practices that you may or may not be implementing in your business. Please only respond ‘Yes’ if you have implemented this business practice at least one time in the last 3 months. It is okay if you are not doing every activity we describe and not all the practices we ask about may make sense for all businesses to do. We are just trying to assess where your business stands now. Over the last three months...

- [M1] ...have you visited at least one of your competitor’s businesses to see what prices they are charging? {No,Yes}
- [M2] ...have you visited at least one of your competitor’s businesses to see what products your competitors have available for sale? {No,Yes}
- [M3] ...you asked existing customers whether there are any other products they would like your business to sell or produce? {No,Yes}
- [M4] ...have you talked with at least one former customer to find out why they have stopped buying from your business? {No,Yes}
- [M5] ...have you asked a supplier about which products are selling well in your business’s industry? {No,Yes}
- [M6] ...have you ever attracted new customers with a special offer? {No,Yes}
- [M7] ...have you marketed or advertised in any form? {No,Yes}
- [M8] ...have you made use of digital social media for marketing purposes? (e.g., Facebook, WhatsApp, Twitter, LinkedIn) {No,Yes}
- [M9] ...have you attempted to negotiate with a supplier/vendor for a lower price on raw material/stock/inputs to your business? {No,Yes}
- [M10] ...did you keep track of every sale made by the business (all money into the business)? {No,Yes}
- [M11] ...did you keep track of every purchase/expenditure (all money out) made by the business? {No,Yes}
- [M12] ...did you use records regularly to know whether sales of a particular product are increasing or decreasing from one month to another? {No,Yes}
- [M13] ...did you know which goods, products, or services make the business the most profit per item selling? {No,Yes}

[M14] ...have you worked out the cost to your business for each of your main products or services you sell? {No,Yes}

[M15] ...did you have a written budget, which states how much is owed or should be spent each month on rent, electricity, equipment maintenance, transport, advertising, and other indirect costs to business? {No,Yes}

[M16] ...have you set targets for how much you will sell? {No,Yes}

Time management

[TM1] In the last seven days, how effectively did you feel you managed your time in running your business? {I managed my time running my business very poorly, Poorly, Neither poorly or well, Well, I managed my time running my business very well}

[TM2] In the last seven days, how many of your planned tasks were you able to accomplish? {I accomplished none of my planned tasks, A few of them, About half of them, Most of them, I accomplished all of my planned tasks}

[TM3] In the last seven days, how often did you find yourself working outside of your planned hours or feeling rushed? {Never worked outside planned hours or felt rushed, Rarely, Sometimes, Often, Always worked outside planned hours or felt rushed}

Innovation

[In1] In the past 30 days, have you considered any new products, services, or business process ideas for your business? No (I have NOT considered any new products, services, or business process ideas), Yes (I HAVE considered new products, services, or business process ideas)

[In2] In the past 30 days, have you tested or experimented with any new products, services, or business process ideas for your business? No (I have NOT tested or experimented with new products, services, or business process ideas), Yes (I HAVE tested or experimented with new products, services, or business process ideas)

[In3] In the past 30 days, have you implemented any new products, services, or business process ideas for your business? No (I have NOT implemented with new products, services, or business process ideas), Yes (I HAVE implemented with new products, services, or business process ideas)

Technology use

- [TE1] How frequently do you use technology (like computers, applications, software) in running your business? I never use technology (like computers, applications, software) in running my business {Rarely, Sometimes, Often, I very often use technology (like computers, applications, software) in running my business}
- [TE2] To what extent do you think technology has impacted your business performance like your growth in profits or customer base? {Technology has had no effect, Technology has had a small effect, Technology has had a moderate effect, Technology has had a large effect, Technology has had a very large effect}
- [TE3] In the last 30 days, have you started using any new technologies in your business? {No (no new technologies), Yes, one new technology, Yes, two new technologies, Yes, three or more new technologies}

Information: Quantity and quality

- [IQ1] In the last 30 days, how easy has it been for you to get advice, mentorship, or information for your business when you need it? {Very difficult to get advice, mentorship, or information, Difficult, Neutral, Easy, Very easy to get advice, mentorship, or information}
- [IQ2] In the last 30 days, how would you rate the quality of advice, mentorship or information you received for your business (across all sources you receive advice, mentorship, and information from)? {The quality of advice, mentorship, and information I receive is Very poor, Poor, Neutral, Good, The quality of advice, mentorship, and information I receive is very good}

Information: Stigma and negative judgement

- [IS1] In the past 30 days, how often have you felt judged or looked down upon when asking for advice, mentorship, or information for your business? {I have always felt judged or looked down upon when asking for advice, mentorship, or information for my business, Almost always, Sometimes, Almost never, I have never felt judged or looked down upon when asking for advice, mentorship, or information for my business}
- [IS2] In the past 30 days, has fear of negative judgement ever prevented you from seeking advice or mentorship or information for your business? {Fear of negative judge-

ment has always prevented me from seeking advice or mentorship or information for my business, Almost always, Sometimes, Almost never, Fear of negative judgment has never prevented me from seeking advice or mentorship or information for my business}

Information: Relevance of mentorship/advice

[IR1] In the past 30 days, is the advice, mentorship, or information you have received directly relevant to your current business needs? {The advice I received is never relevant to my current business needs, Rarely, Sometimes, Often, The advice I received is always relevant to my current business needs}

[IR2] In the past 30 days, how frequently have you found the advice, mentorship or information you received to be modern and up-to-date with current business trends and practices? {The advice I received is never modern and up to date, Rarely, Sometimes, Often, The advice I received is always modern and up to date}

E.4 Explanation of Preferred Functional Forms

In our pre-registration, we pre-specified four different variations on a performance index: 95 and 99% winsorization and a log-like inverse hyperbolic sine transformation $y' = \log(y + (y^2 + 1)^{1/2})$ transformation of these two winsorizations. In the paper body, we focus on the raw outcome variable and report the raw unit results as our preferred specification for two reasons.

First, recent work has shown that the inverse hyperbolic sine transformation (and other similar log-like transformations) is highly sensitive to artefactual statistical features like scaling of the outcome (e.g., measuring performance in units of thousands of KSH), and also the presence of zeros (Chen and Roth, 2023; McKenzie, 2023). We observe exactly this sensitivity to zeros in our data. Figure A9 shows the distribution of our performance indices, winsorized at the 95% level on the left and the 99% on the right. The top row depicts raw units and the bottom depicts log-like units. We observe that for the log-like index, there is an extreme left tail that arises from zeros in the performance distribution, with outcome values close to or above -10 standard deviations.

Second, for the purpose of studying inequality, it makes more sense to focus on raw

changes as proportional changes can reflect radically different performance impacts depending on whether the entrepreneur comes from the top or bottom of the distribution.

F Comparison to existing literature

In this section, we provide additional information on how to interpret our results in relation to the emerging literature on the impact of generative AI on productivity and performance. While several other papers have now found positive effects of generative AI tools on a variety of tasks, we are unable to reject the null hypothesis that the generative AI assistant would have no effect on business performance.

To interpret this null result, we conducted exploratory post-hoc equivalence tests with the aim of rejecting the presence of treatment effects greater than or equal to a certain minimum effect size threshold (Simonsohn, 2015; Lakens, 2017). We focus on a comparison to Noy and Zhang (2023) because, to our knowledge, it is the only peer-reviewed and published experiment testing the impact of generative AI on business performance. We set this threshold to 0.26 s.d., which is the minimum detectable effect size that Noy and Zhang (2023) can reliably detect, and strongly reject the presence of a treatment effect greater than or equal to this threshold (threshold = 0.26, $Z = -3.89$, $p = 0.00005$). In other words, we can clearly establish that the effect of AI on our measure of business performance is less than the productivity improvements reported in Noy and Zhang (2023).

Table A2 positions our work more broadly relative to an emerging set of papers on the impact of AI on productivity. We note that, aside from Noy and Zhang (2023), these papers have not been published in a peer-reviewed journal. In contrast to our results, each of these experiments finds significant impacts on productivity as a result of access to generative AI. Additionally, unlike our paper, each of these studies finds that performance improvements are driven by gains among the lowest-performing businesses, which suggests, in contrast to our work, that generative AI could lead to reductions in inequality.

Importantly, we note that in each of these experiments, participants do not have control over which tasks they are assigned. In our experiment, participants have discretion over which questions they ask and which advice they choose to implement. We argue that these

features could lead to an increase in inequality as better-performing businesses are better able to ascertain which advice is likely to be worth following.

G Ethical Considerations

In our field experiment, we found no overall impact on business performance from the treatment intervention. However, we did detect heterogeneous treatment effects with respect to initial performance levels; high performers experienced a performance increase due to AI advice, while low performers experienced a performance decrease. These unexpected negative results gave us pause. Given that our treatment appears to have led to worse performance for some entrepreneurs in our sample, in this section we outline why our experiment is ethical, above and beyond having received approval from and followed the policies of the UC Berkeley’s Institutional Review Board.

First, according to [Asiedu et al. \(2021\)](#), a key ethical consideration is policy equipoise, which involves meaningful uncertainty about the efficacy of each treatment arm. Indeed, if there was consensus that one treatment arm would be superior to the other, then it would *not* be ethical to randomize it unless the treatment was costly and therefore scarce. In the context of our experiment, there was uncertainty regarding participants’ benefits from each arm of the study. To the best of our knowledge, we are the first to test whether AI advice can increase entrepreneurial performance, especially in a context like Kenya. That said, recent research on generative AI and research on advising and mentoring between entrepreneurs suggested that an AI assistant had the potential to increase business performance. While we believed there were benefits, there was no consensus at the time we deployed the experiment that one treatment would be superior to the other. Given the presence of equipoise, we argue it was more ethical to test the impact of AI with a few hundred entrepreneurs in a well-designed study, rather than not testing and making policy decisions that could a) deny millions of entrepreneurs access to a useful tool if the treatment effect was positive or b) encourage adoption of a technology that hurts productivity by millions of entrepreneurs if the treatment effect was negative ([Glennerster and Takavarasha, 2014](#)).

Beyond the presence of equipoise, we also ensured all participants in our study were well-informed and free to decide whether to participate or not. Our participants were also free to decide whether to continue using our AI assistant, and it was easy to stop using it, consistent with the ethical standards around participant choice.

While our experiment revealed that one particular subgroup did not benefit from access to the AI assistant, negative treatment effects for subpopulations commonly occur in experiments conducted in the social and medical sciences ([Karlan, Knight, and Udry, 2015](#)). Given the growing use of generative AI, understanding the impact of such tools on society is especially valuable.

H Accounting for Spillovers

In this appendix, we examine the plausibility of two types of spillovers that could impact our results: treatment access spillovers, where non-treated individuals might gain access to the AI assistant, and performance spillovers (Holtz et al., 2023), in which performance improvements (or declines) for one individual come at the direct cost (or benefit) of another.

Treatment Access Spillovers

The AI assistant’s WhatsApp contact number is publicly available, allowing anyone to text in. Without strict controls, this raises the possibility that untreated participants could interact with the AI assistant. To address this, we developed our AI assistant in such a way that it only responds to messages from individuals on a pre-specified access list. Only members of the research team were able to modify this access list.

Performance Spillovers

Another concern relates to performance spillovers within our experimental sample. For instance, the heterogeneous treatment effects we observe could potentially be explained by the high-performing entrepreneurs in our sample capturing profits/revenue from the low-performing entrepreneurs in our sample. However, we consider this unlikely for two reasons. First, our advertisement on Meta targeted a broad audience, drawing participants from across Kenya, a country of over 50 million people where Facebook has very high market penetration (Wamuyu, 2020). Consequently, our recruitment strategy yielded a geographically diverse sample. Table A1 illustrates the distribution of respondents over Kenyan counties (looking only at those counties with over 50 respondents). Second, negative spillovers are overwhelmingly more likely to occur between individuals who are located in the same geography *and who work in the same industry*. In Figure A3, we further stratify our sample by county and sector. Because our sample is drawn from all over Kenya, every sector-location combination has fewer than 40 participants (note that some participants have businesses in multiple sectors). Only two sector-location combinations have over 20 individuals: Nairobi × Food and Nairobi × Clothing. As a point of reference, Nairobi has a population of over

4,000,000 and the Nairobi metropolitan area spans over 250 square miles.

It *is* possible that performance spillovers occur between entrepreneurs in our sample and their competitors that are *not* in our sample. Keeping this in mind, the low level of technology saturation and infeasibility of the diffusion of our AI assistant are key to interpreting our results. If there were a greater level of overall generative AI saturation in Kenya, our results may not hold. For instance, it is conceivable that with high levels of generative AI adoption, firm practices could uniformly improve and/or converge; this could lead to greater price competition and lower firm profits. Conversely, widespread adoption of generative AI could drive positive performance spillovers stemming from the diffusion of best-use practices among businesses. While our experiment is unable to shed light on these dynamics, we consider this a fruitful area for future research.

I Detailed Case Examples

Example 1: Dairy Products Business Owner (Below Median)

In this conversation, a milk seller sought advice on various aspects of their business. Initially, they faced a dilemma when a price hike from their supplier forced them to increase their retail prices, leading to the loss of some customers. They receive advice from the AI assistant on methods to maintain customer loyalty during this period, including transparent communication about the price increases, adding value to their products, ensuring quality, engaging with customers, and introducing flexible pricing options like smaller quantities and bulk discounts.

The entrepreneur then expressed an interest in diversifying their product range to maximize profits. They were advised to consider selling items that complement milk, such as tea, coffee, cereals, bread, pastries, sweeteners, and other dairy products like cheese and yogurt. However, they faced challenges in finding reliable suppliers for these new products and received poor feedback from customers. In response, the AI assistant offered further suggestions, including online directories to find suppliers, attending trade shows, incentivizing customer feedback, and maintaining an open dialogue with customers to understand their preferences and improve product offerings.

Finally, the entrepreneur sought ideas for further expanding their business after their attempt to sell yogurt was not successful. They were advised to explore options like cheese making, starting a milk delivery service, selling milk-based snacks or desserts, expanding their wholesale business, and conducting training workshops on dairy products. Specific strategies for expanding into milk-based snacks and desserts were also discussed.

Q1. *Hello, I am [REDACTED]. I sell milk both wholesale and retail, in the past 1 week the price hiked from the supplier, so..I also decided to hike my price too in order to get some reasonable profit but now some customers have already left. What will I do to make them understand the prices are high from the production level?*

A1. Hello [REDACTED], it's a common challenge for businesses when costs increase. Here are some strategies to help your customers understand the situation:

1. ****Transparent Communication****: Inform your customers about the increased prices from suppliers and explain that you had to adjust yours as well. This can be done through signage at your shop or via social media if you have an online presence. The key is to be open and honest about why prices have increased.
2. ****Value Addition****: Enhance the perceived value of your milk products by offering something extra with each purchase such as loyalty points, discounts on bulk purchases, or small freebies like a packet of tea leaves or sugar.
3. ****Quality Assurance****: Assure them that despite the price hike, the quality of your milk remains high. If possible, share information about how you ensure quality control in sourcing and storing your milk.
4. ****Customer Engagement****: Engage more with your customers by asking for their understanding and support during these tough times while promising to review prices once things normalize at production level.
5. ****Flexible Pricing Options****: Offer different pricing options where possible, such as smaller quantities at lower prices which may still be affordable for most of your customers.

Enter a number to learn more about one of these points on **customer retention strategies** or ask a new question.

Q2. 5

A2. Great choice, [REDACTED]. Implementing flexible pricing options can be a practical approach to retain customers during price hikes. Here are some additional tips on how you can do this:

1. ****Small Quantity Pricing****: Offer smaller quantities of milk at lower prices. This could mean selling in half-liter or quarter-liter measurements for those who cannot afford a full liter due to the increased price.
2. ****Bulk Discounts****: Provide discounts for customers buying in bulk. For example, if a customer buys 10 liters, they could get an extra liter free or at a reduced rate.
3. ****Loyalty Program****: Establish a loyalty program where regular customers earn points with each purchase that can be redeemed for discounts or free products in the future.
4. ****Payment Plans****: If feasible and applicable to your business model, consider offering payment plans for trusted regular customers who buy large quantities regularly but are struggling with the sudden increase in price.
5. ****Promotional Offers****: Run occasional promotional offers such as 'buy one get one free' or 'price reductions on certain days' to keep customers interested and engaged despite the general price hike.

Enter a number to learn more about one of these points on **flexible pricing strategies** or ask a new question.

Q3. *I'm [REDACTED], I sell milk both wholesale and retail but now, I want to add something else to the business in order to maximise my profit range. I don't know what will blend well with milk.*

A3. Hello [REDACTED], it's great that you're thinking about diversifying your product range. Here are some suggestions of products that typically complement milk sales:

1. ****Tea and Coffee****: These are commonly used with milk. You could consider selling various types of tea leaves or coffee beans.
2. ****Cereal and Porridge****: Many people enjoy these with milk for breakfast. Offering a variety of cereals or porridge mixes could attract more customers.
3. ****Bread and Pastries****: These items often go well with milk-based beverages for breakfast or snacks.
4. ****Sweeteners****: Consider selling honey, sugar, or other sweeteners which are frequently added to milk.
5. ****Dairy Products****: You can expand into other dairy products like cheese, yogurt, or butter which have a similar supply chain as milk.

Remember to research each potential addition thoroughly to understand the demand in your area, the cost implications, and how it might affect your operations.

Enter a number to learn more about one of these points on **product diversification** or ask a new question.

Q4. 5

A4. Excellent choice, [REDACTED]. Expanding into other dairy products can be a natural extension of your current business. Here are some actionable steps:

1. ****Market Research****: Conduct surveys or informal chats with your customers to understand which dairy products they'd like to see in your store.
2. ****Supplier Search****: Look for reliable suppliers who provide high-quality cheese, yogurt, and butter at competitive prices.
3. ****Product Testing****: Initially introduce a small quantity of the new product to gauge customer response before fully stocking it.
4. ****Promotion****: Inform your customers about the new products through signage, social media, or direct communication when they visit your shop.
5. ****Feedback Loop****: Regularly ask for feedback from customers on these new products and adjust based on their responses - whether it's changing the brand, adjusting pricing or discontinuing if it doesn't meet expectations.

Remember that diversifying into dairy requires proper storage facilities due to perishability and quality control needs.

Enter a number to learn more about one of these points on **diversifying into other dairy products** or ask a new question.

Q5. 2

A5. Absolutely, finding the right suppliers is crucial when diversifying into other dairy products. Here are some steps to help you find reliable suppliers:

1. ****Identify Potential Suppliers****: Start by making a list of potential suppliers. You can do this by searching online, asking for recommendations from other business owners, or visiting local markets.
2. ****Evaluate Product Quality****: Request samples from these suppliers and assess their product quality. Remember, the quality of your products will reflect on your business reputation.
3. ****Compare Prices****: Ask for price lists from each supplier and compare them considering both the cost and quality of the products.
4. ****Check Reliability****: Find out about their delivery schedule, order minimums, return policy etc., to ensure they meet your specific needs.
5. ****Build Relationships****: Once you've chosen a supplier or two, build strong relationships with them as it can lead to better prices, more favorable terms and improved service over time.

Remember that a good supplier does more than just deliver products - they should be partners in helping your business succeed.

Enter a number to learn more about one of these points on **finding reliable dairy product suppliers** or ask a new question.

Q6. *Hello, this is [REDACTED] I sell milk both in wholesale and retail I contacted you asking what I can add in my shop that blends with milk and decided to explore other dairy products like cheese and yogurt now the problems that am facing are: (1). Supplier search i have not yet got a reliable supplier. (2). I have not been getting good feedback on why the customers are not loving the new products*

A6. Hello [REDACTED], I understand your concerns. Let's tackle them one at a time:

1. ****Reliable Supplier Search****: Here are some actionable steps you can take: [1] Use online directories or social media groups for Kenyan businesses to find potential suppliers. [2] Attend local trade shows or farmers markets where you can meet potential suppliers in person. [3] Ask other business owners in your network if they have recommendations for reliable dairy product suppliers.
2. ****Improving Customer Feedback****: [a] Conduct a brief survey with customers who've tried the new products. Ask specific questions about taste, packaging, price, and what they'd like to see improved. [b] Offer samples of the products to customers before they buy - this could help identify any issues early on and also encourage more sales. [c] Consider introducing smaller sizes or trial packs that cost less so that customers can try the products without feeling like it's a big investment.

Remember that changes in business often require patience and iterative improvements based on feedback. Don't be discouraged by initial setbacks.

Enter a number to learn more about one of these points on **supplier search** or **improving customer feedback** or ask a new question.

Q7. 2

A7. You've chosen to explore improving customer feedback further. Here are some additional steps:

1. ****Open Dialogue****: Encourage customers to share their thoughts openly about the new products. Make them feel that their opinion is valued and important for your business improvement.
2. ****Incentivize Feedback****: Offer small incentives like discounts or free samples in exchange for completing a feedback form.
3. ****Direct Observation****: Watch how customers interact with the new products in your store. Do they seem confused? Are they comparing it with other items? This can give you clues about any potential issues.
4. ****Trial Periods****: Run short trial periods for different products, then ask for feedback specifically on those items at the end of each period.
5. ****Transparent Communication****: If you make changes based on customer feedback, let them know! This shows that you value their input and are committed to improving their experience.

Remember, **feedback is an opportunity** to learn more about your customers' needs and preferences, which can guide your business decisions moving forward.

Enter a number to learn more about one of these points on **improving customer feedback** or ask a new question.

Q8. *I sell milk both wholesale and retail in [REDACTED]. .now I wanted to know what else I can add to my business in order to maximize my profits. I tried yoghurts but didn't work.*

A8. Thank you for sharing about your business. It's great to see your interest in diversifying and increasing profits. Here are a few suggestions:

1. ****Cheese Making****: With milk as your primary resource, cheese can be an excellent addition. Cheese has a longer shelf life than milk and can be sold at higher prices. Start with simple types like cottage cheese or mozzarella, which require minimal equipment and skills.
2. ****Milk Delivery Service****: Consider offering home delivery services for regular customers within Nairobi. This service can increase customer loyalty, bring in consistent revenue, and allow you to reach more customers who value convenience.
3. ****Selling Milk-Based Snacks or Desserts****: Depending on your location and customer base, adding snacks or desserts such as ice cream, milkshakes, or custards could attract more retail customers.
4. ****Expand Your Wholesale Business****: Identify other businesses that might need dairy products (like restaurants or bakeries) and offer them competitive prices for bulk purchases of your milk.
5. ****Training Workshops****: Conduct training workshops on the benefits of dairy products or how to make simple dairy-based foods at home (like homemade yogurt). This not only promotes your products but also creates an additional income stream from workshop fees.

Enter a number to learn more about one of these points on **business expansion** or ask a new question.

Q9. 3

A9. Great choice! Selling milk-based snacks or desserts can attract more customers and increase your profit margin. Here are some steps to implement this strategy:

1. **Market Research**: Start by understanding what kind of milk-based snacks or desserts are popular in your area. You could do a quick survey with your current customers or observe what's selling well in nearby food businesses.
2. **Simple Recipes**: Choose simple recipes that require minimal additional ingredients and equipment. Puddings, custards, or simple ice creams could be good starting points.
3. **Pricing**: Price these items competitively while ensuring you still make a profit after considering the cost of ingredients, preparation time, and utilities.
4. **Promotion**: Use signage at your shop to promote these new products and consider offering samples to regular customers to get their feedback and generate interest.
5. **Quality Control**: Ensure the quality of your desserts is consistently high so that customers keep coming back for more.

Enter a number to learn more about one of these points on *milk-based snacks/desserts* or ask a new question.

Q10. 1

A10. Absolutely, conducting market research is a crucial first step. Here are some actionable steps:

1. **Surveys**: Create simple surveys asking your customers about their favorite milk-based snacks or desserts and if they'd be interested in buying them from your shop. This could be done informally when they visit your shop or through a more structured approach like questionnaires.
2. **Observation**: Visit other food businesses in your area that sell milk-based snacks or desserts. Note which items seem popular, how they're priced, and any unique selling points.
3. **Community Engagement**: Engage with local community groups or online forums to understand what types of desserts people in Nairobi enjoy and are currently trending.
4. **Trial Products**: Once you've gathered this information, select one or two products that seem promising and start offering them on a trial basis to gauge customer response.
5. **Feedback Collection**: Collect feedback from the customers who try these new products - did they enjoy it? Would they buy it again? How much would they be willing to pay for it?

Remember: The goal of market research is not just to identify what products you should offer but also to understand how you can differentiate yourself from competitors.

Enter a number to learn more about one of these points on *market research* or ask a new question.

Example 2: Fast Food Joint Owner (Above Median)

In this conversation, the owner of a fast food joint located near a matatu terminus sought advice on differentiating their business in a competitive environment. The key strategies discussed included differentiating the menu with unique or special dishes, improving customer service, implementing a loyalty program, engaging in community activities, and maintaining high standards of cleanliness and hygiene.

The discussion delved into the specifics of establishing a loyalty program, with the AI assistant offering suggestions like a point-based system, special offers for regular customers, and birthday or anniversary rewards. The entrepreneur raised concerns about the cost implications of differentiating the menu and the complexities of community involvement. The AI assistant then suggested using affordable local ingredients, offering smaller portions, running specials during off-peak hours, using simple loyalty systems like punch cards, and engaging in low-cost community activities.

The entrepreneur was also concerned about the risks and customer reception related to changing the menu and implementing new strategies. The AI assistant recommended testing new dishes gradually and seeking customer feedback, especially from regulars. The importance of using both in-person and digital channels to collect reliable feedback was emphasized, considering that some customers are transient. The entrepreneur and the AI assistant also discussed the risks associated with implementing new business strategies, with the AI assistant suggesting thorough market research, pilot testing, financial analysis, risk assessment, and regular feedback review.

Q1. *I have fast food joint that I run just adjacent to a matatu terminus. There are several such food joints at that particular place and it's fair to say competition is stiff. How can I overcome the competitive business environment and be ahead of the rest of the pack?*

A1. Thank you for sharing about your business. Operating a fast food joint in a highly competitive environment like a matatu terminus can indeed be challenging, but there are several strategies you could consider to stand out:

1. ****Differentiate Your Menu****: Offer unique or special dishes that aren't available at the other joints. This could be traditional Kenyan foods with a twist, or it could involve using locally sourced ingredients that others don't use. The benefit is that customers will associate your place with these special offerings and choose you over competitors.
2. ****Improve Customer Service****: Prioritize excellent customer service to make people feel welcome and appreciated when they visit your joint. This might involve training staff on how to interact with customers positively, quickly serving orders, and handling any complaints promptly and professionally.
3. ****Loyalty Program****: Implement a loyalty program where customers earn points for every purchase which they can later redeem for free meals or discounts. It encourages repeat business as customers will want to accumulate more points.
4. ****Community Involvement****: Engage in local community events or sponsor local initiatives such as sports teams or school activities. This raises your profile within the community and makes people more likely to support your business.
5. ****Cleanliness & Hygiene****: Ensure high standards of cleanliness both inside the kitchen and in the dining area because hygiene matters greatly in food businesses.

Enter a number to learn more about one of these points on **improving competitiveness** or ask a new question.

Q2. 3

A2. Great, let's delve into how you can implement a ****Loyalty Program**** in your fast food joint:

1. ****Point-Based System****: Implement a simple point-based system where customers earn points for every shilling spent. Once they accumulate a certain number of points, they can redeem them for free meals or discounts. You could use simple tools like punch cards or digital apps that are inexpensive to manage.
2. ****Special Offers for Regular Customers****: Identify your regular customers and offer them exclusive deals or discounts. This not only rewards their loyalty but also encourages them to keep coming back.
3. ****Birthday/Anniversary Rewards****: Offer special treats or discounts to customers on their birthdays or anniversaries if they share this information with you. This personalized touch can make customers feel valued and appreciated.
4. ****Referral Program****: Encourage your loyal customers to bring in new clients by giving them incentives such as a discount or free meal when someone they refer makes their first purchase.
5. ****Feedback Reward****: Encourage customer feedback by offering small rewards for completing surveys about the quality of food and service at your joint.

Remember, the goal of any loyalty program is to make your customers feel valued so that they continue patronizing your business over others.

Enter a number to learn more about one of these points on **loyalty programs** or ask a new question.

- Q3. 1. **Differentiating my menu** means that I'm incurring a higher cost of preparing meals. This in turn pushes the menu prices upwards which may scare away the potential customers considering the fact that the potential customers are low or middle income earners. 2. About **loyalty programs**, I'm not well conversant with that field. 3. **Community involvement** may require adequate resources and may be capital intensive.

A3. Thank you for your detailed response. Let's delve into these topics a bit more:

1. ****Differentiating Your Menu****: You're correct that introducing unique dishes may increase costs. However, this doesn't always have to result in higher menu prices. Here are some tips: [1] Focus on dishes that use affordable local ingredients but are prepared in a distinctive way. [2] Consider offering smaller portions at lower prices for those who want to try something new without paying a lot. [3] Run specials during off-peak hours to attract customers and offset the cost of unique dishes.

2. ****Loyalty Program****: A loyalty program is essentially a reward system for frequent customers. Here's how you could implement one with minimal costs: [1] Use simple paper punch cards where each purchase gets a punch, and after a certain number of punches, the customer gets a free meal or discount. [2] Keep track of repeat customers manually and offer them discounts or free items after they've made several purchases.

3. ****Community Involvement****: While it can be resource-intensive, there are ways to engage with the community without spending much money: [1] Offer your venue as a meeting place for local groups during slow business hours. [2] Sponsor small-scale local events like neighborhood cleanups by providing refreshments - this helps build goodwill within the community.

Enter a number to learn more about one of these points on **cost-effective strategies** or ask a new question.

- Q4. *If we focus on having a menu that stands out from what others are offering, by what margin may I lose, retain or increase my customer base. And what is the likelihood of having rewards for customer's referrals?*

A4. The impact of differentiating your menu on your customer base can be significant, but it's hard to quantify without specific data. However, here are some general insights:

1. ****Potential Increase in Customers****: Offering unique dishes can attract food lovers who are always looking for something new and exciting to try. If your new offerings resonate with the taste preferences of the local population, you could see a substantial increase in customers.

2. ****Retention of Existing Customers****: If you continue offering popular items while introducing new ones, you're likely to retain most of your existing customers. They might even appreciate the added variety.

3. ****Possible Losses****: There is always a risk that some customers may not like changes to the menu and choose other joints over yours. It's essential to introduce changes gradually and get feedback from customers before making any drastic alterations.

As for referral rewards:

4. Referral programs can be highly effective because people trust recommendations from friends or family more than advertising messages from businesses. If executed well, this strategy could lead to a significant increase in new customers without requiring large investments in marketing.

To implement these strategies:

A) For **Menu Differentiation**, start by researching what foods or flavors are trending locally or globally that align with your capabilities and resources. Test these out as specials before adding them permanently to your menu based on customer response.

B) For **Referral Rewards**, consider giving both the referring customer and the referred friend a discount or free item on their next purchase. Make sure it's easy for customers to share referrals by providing them with simple instructions when they make purchases.

Enter a number to learn more about one of these points on **menu differentiation** or **referral rewards** or ask a new question.

Q5. *I'll need to contemplate on this two strategies and evaluate the outcomes both on a short and long term basis.*

A5. Absolutely, taking the time to contemplate and evaluate potential outcomes is a crucial step in making strategic decisions for your business. Remember that both **menu differentiation** and **referral rewards** are strategies aimed at improving customer retention and attracting new customers, which are key aspects of business growth.

When evaluating these strategies:

1. Consider your resources: Do you have the capacity to introduce new menu items or manage a referral program effectively?

2. Listen to your customers: What do they want or need? Would they appreciate a more diverse menu or would incentives for referrals motivate them?

3. Monitor results: Once implemented, track the success of these strategies through increased sales, more repeat customers, or an increase in new customers.

4. Be ready to adjust: If something isn't working as well as expected, don't be afraid to make changes.

Remember that successful businesses are those that adapt and evolve based on their environment and customer needs.

Feel free to return if you have any other questions or need further clarification on any point related to improving your fast food joint's competitive advantage

Q6. *Customer feedback when venturing into new ideas is the key to monitor the reception of a new menu and also the trust built through referral programs may work well with the business and the projected profit margins.*

A6. Absolutely, you're on the right track. Customer feedback is invaluable when introducing new ideas like a different menu or a referral program. Here's how you can effectively use customer feedback and trust building:

1. ****Feedback Channels****: Establish easy ways for customers to give feedback such as suggestion boxes at your joint, conducting informal surveys, or simply talking to them about their experience.

2. ****Respond to Feedback****: Show customers that you value their opinions by acting on their suggestions where feasible and communicating changes back to them.
3. ****Trust Building in Referral Programs****: Make sure the process is transparent and rewards are delivered as promised. This will build trust not only with the referring customer but also with the new ones coming in through referrals.
4. ****Evaluate Profit Margins****: Regularly review your costs versus revenues after implementing these strategies to ensure they're contributing positively towards your profit margins.
5. ****Pilot Test****: Before rolling out any major changes, consider doing a small-scale pilot test to gauge reception and adjust based on feedback received.

Remember, open communication with your customers helps foster loyalty which can lead to increased profitability in the long run.

Enter a number to learn more about one of these points on **customer feedback** and **trust building** or ask a new question.

Q7. *Together with my team, we shall deliberate on few important tips I've come across here and decide the way forward in best way to implement them.*

A7. That sounds like a great plan. Collaborative decision-making often leads to more effective strategies and smoother implementation. Here are some tips for your deliberation process:

1. ****Set Clear Goals****: Clearly define what you want to achieve with each strategy, such as increasing customer retention, attracting new customers, or boosting profit margins.
2. ****Assign Roles****: Assign team members specific roles in the implementation of each strategy based on their skills and responsibilities.
3. ****Plan for Feedback****: Plan how you will collect and analyze customer feedback once the changes are implemented.
4. ****Prepare for Adjustments****: Be prepared to make adjustments based on the outcomes of your initial efforts.
5. ****Schedule Follow-ups****: Schedule regular follow-up meetings to assess progress and make necessary adjustments.

Remember that it's important to have everyone on board with these changes so they can be effectively implemented and managed over time.

Enter a number to learn more about one of these points on **team deliberation** or ask a new question.

Q8. *Good insight. I'll come back here sooner with some more insights on the progress made if any.*

A8. That's great to hear! Remember, the process of implementing new strategies often involves learning and adjusting along the way. Don't hesitate to return if you have any questions or need further advice as you progress. Wishing you and your team all the best in enhancing your business performance and profitability.

Q9. *Hi... If I'm running a food joint near a matatu terminus, selling fast food, snacks and beverages, how can I diversify the services I offer*

A9. Hello! Diversifying your services can be a great way to attract more customers and increase revenue. Here are some ideas:

1. ****Delivery Service****: You could start offering delivery service for customers who may not have time to visit your food joint in person but want a quick meal.
2. ****Catering Services****: Consider providing catering services for local events or businesses, which can significantly boost sales.
3. ****Meal Deals****: Introduce special meal deals during specific hours of the day, such as breakfast combos or late-night snacks.
4. ****Seasonal Menu Items****: Offering seasonal menu items that align with holidays or local events can attract new customers and keep regulars interested.
5. ****Loyalty Program****: Implement a loyalty program where regular customers earn points they can redeem for meals or discounts.

Each strategy has its own benefits and methods of implementation, so consider what would work best for your specific business situation and resources.

Enter a number to learn more about one of these points on **service diversification** or ask a new question.

Q10. *How do I increase the returns on my cafeteria?*

A10. Understanding your business needs is crucial. I'm assuming your cafeteria is in Kenya, but could you provide more details? What's the size of your cafeteria, and where is it located? What's the competition like? Any specific challenges you are facing?

Based on general information, here are some strategies to increase returns:

1. ****Improve Menu****: Review your menu to ensure it appeals to local tastes and preferences. Make sure you're offering dishes that not only taste good but also have high-profit margins. Experiment with different dishes and keep track of which ones sell best.
2. ****Cost Control****: Manage food waste by implementing inventory management systems and portion control measures. This can help reduce costs and increase profitability.
3. ****Customer Experience****: Enhance the customer experience by training staff in customer service skills, improving café ambiance or cleanliness, or even offering free Wi-Fi if feasible.
4. ****Promotions & Loyalty Programs****: Implement a loyalty program or offer special promotions during slow periods to attract more customers.
5. ****Community Engagement****: Get involved with the local community through sponsorships or events which can enhance visibility and build a loyal customer base.

Enter a number to learn more about one of these points on **increasing cafeteria returns** or ask a new question.

Q11. *I have tried to offer a new special meal menu in my restaurant and I must admit that the returns have been quite fair. Is it okay if I do a whole menu overhaul and start offering some different special meals altogether, even though it would make the price of each plate to be higher than before. I'm stuck between giving new menu or staying with current ones for the fear of losing the current clientele.*

A11. That's a great question and it's normal to feel this way when considering major changes in your business. Here are some strategies you can consider:

1. ****Test the Waters****: Before overhauling the entire menu, consider introducing new special meals one at a time. This will allow you to gauge customer response without risking too much.
2. ****Customer Feedback****: Ask your customers directly for their opinion on potential new dishes or even higher prices. You could use surveys or informal conversations.
3. ****Market Research****: Look at what similar restaurants in your area are doing successfully, especially those with a similar clientele.
4. ****Pricing Strategy****: Consider if there's room to slightly increase prices of current dishes instead of introducing more expensive ones that might deter customers.
5. ****Special Events****: Host special events where only the new menu is available, like themed nights or tasting events, to introduce customers to these dishes and gather feedback.

Enter a number to learn more about one of these points on **menu overhaul** or ask a new question.

Q12. 1

A12. Great, you're interested in the strategy of testing the waters with new special meals. Here are some additional tips on how to do that:

1. ****Pilot Dish****: Start by introducing one or two new dishes as a special offer and observe customer reactions.
2. ****Promote****: Make sure your customers know about the new dish through table tents, menu boards, or even verbally when they're ordering.
3. ****Feedback****: Ask for feedback directly from those who try the new dishes. This could be done informally by waitstaff or through a small comment card.
4. ****Evaluate****: Track sales of the new dishes and compare them to your regular items. If they perform well, consider adding them permanently.
5. ****Adjust****: Based on feedback and sales data, adjust the dish if needed before fully integrating it into your menu.

Enter a number to learn more about one of these points on **testing new special meals** or ask a new question.

Q13. *As for the customer feedback, sometimes it may not be reliable since some of the clients we host are passengers on transit so some opinion they may give may not apply to everyday's consideration.*

A13. You're correct, transient customers may provide feedback that doesn't necessarily reflect the preferences of your regular clientele. Here are some strategies to ensure you're getting reliable feedback:

1. ****Target Regular Customers****: Make a point to solicit feedback from your regulars who understand your current menu and can provide insights into potential changes.
2. ****Digital Feedback****: If you have an online presence (e.g., social media), use these platforms to gather feedback as people might be more open and detailed in their responses.
3. ****Feedback Cards****: Leave these at tables or the checkout counter for customers to fill out before they leave. This will give you a wider range of opinions.

4. ****Observation****: Watch what dishes get ordered most often and which ones come back with leftovers – this can be very telling about what’s working and what isn’t.
5. ****Trial Periods****: As mentioned earlier, introduce new items slowly and see how they perform over a set period before making permanent decisions.

Enter a number to learn more about one of these points on **collecting reliable customer feedback** or ask a new question.

Q14. *A shift in business strategies may sometimes backfire. Implementing new ideas may fail the business often. How can I avoid the setbacks off trying new ideas and taking risks?*

A14. Trying new ideas and taking risks is a part of business growth, but it’s important to manage these risks wisely. Here are some strategies:

1. ****Market Research****: Before implementing a new idea, conduct thorough market research to understand its potential impact on your customers and competitors.
2. ****Pilot Testing****: Instead of implementing changes across the entire business, try them in a small controlled environment first to gauge their effectiveness.
3. ****Financial Analysis****: Consider the financial implications of the new strategy. Calculate potential returns against costs and assess if it’s financially viable.
4. ****Risk Assessment****: Identify potential pitfalls or challenges with your new strategy and develop contingency plans to address them.
5. ****Feedback & Review****: Regularly review the performance of the new strategy, seek feedback from customers and staff, and be ready to make adjustments as needed.

Enter a number to learn more about one of these points on **risk management when trying new business strategies** or ask a new question.

J Generic Machine Learning for Heterogeneous Treatment Effects

We use “Generic ML” as further assess the robustness of the heterogeneous treatment effects we estimate in our experiment (Chernozhukov et al., 2018). These tools are now being used by applied researchers to test for heterogeneous treatment effects in their experiments and our write up here builds on the excellent overview and application of this new method in Davies et al. (2024).

The Generic ML approach beings by splitting the data into a training and test set. Using the training set, it uses “generic” machine learning methods (e.g. Random Forests, Lasso, ...) to build models that predict treatment effects for each observation in the data. These models are then used to predict treatment effects \hat{S}_i for each observation in the test set. These estimated treatment effects allow for three types of analysis.

First, if there is treatment effect heterogeneity *and* the Generic ML model can predict these differences decently, then in the following “best linear predictor” (BLP) regression $Beta_2$ should be positive and statistically significant:

$$Y_i = \beta_1 \times T_i + \beta_2 \times T_i \times \hat{S}_i + X_i$$

Here T_i is an dummy for whether the individual is treated. If there is minimal to no heterogeneity, then the impact of T_i and is simply the value of β_1 for all observations and so the estimate of b_2 should be near 0. If there is heterogeneity, and the model has learned to predict it from the variables included in the training set, then β_2 should be positive as the estimates \hat{S}_i will help T_i better explain who benefits, who doesn’t, and who is harmed by the treatment. In this regard, β_2 serves as a test for any the presence of any type of heterogeneity as long as the mode

Second, conditional on there being evidence for treatment effect heterogeneity, we can use the estimates of S to group observations into those who are estimated to have more positive treatment effects and those with more negative values. For simplicity, we focus on a median split, but results are consistent when we use quartiles or quintiles instead. For

the below- and above-median predicted treatment effect groups, we then calculate the group average treatment effect (GATE). These groups allow us to map where positive, negative, and null treatment effects are concentrated within the distribution of estimated treatment effects.

Third and finally, for each group, we can calculate the average value of the covariates we used to predict treatment effect heterogeneity to understand how the treatment effect heterogeneity identified by the BLP and GATE estimates varies with observables.

Turning to our specific application, we estimate generic ML models using two different sets of data. First, we use the three measures we pre-specified to test for treatment effect heterogeneity: our baseline index of performance, the gender of the entrepreneur. and whether the entrepreneur had used ChatGPT before the experiment. Second, we expand the set of data to test for any potential sort of heterogeneity. In particular, we train a model on the following variables:

- **Index and above median indicator of pre-treatment performance**
- **Index of pre-treatment management practices**
- **Business age:** {Less than 1 year old, Between 1-2 years old, Between 2-3 years old, 3 years or older}
- **Business sector:** {Agriculture, Cosmetics, Carpentry, Jewelry, Multi-sector, Clothing, Food, Retail, Transportation, Information Technology, Other}
- **Number of children**
- **Age of the entrepreneur**
- **Education level of the entrepreneur:** {Completed college, Started college, No college}
- **Entrepreneur has used ChatGPT**
- **Gender**
- **Big Five personality scores:** {Agreeableness, Conscientiousness, Extraversion, Openness, Emotional Stability}
- **Location:** {Machakos, Bomet, Nyeri, Kisii, Mombasa, Kisumu, Kericho, Bungoma, Kakamega, UasinGishu, Nakuru, Kiambu, Nairobi, Other}

For both data sources, we use the R package “GenericML” to train models using Random Forest and Lasso algorithms with 200 splits of the data, though using more algorithms or additional splits has virtually no impact on our estimates. Following [Chernozhukov et al. \(2018\)](#), who also analyze an experiment with blocking and multiple post-treatment periods, and include fixed effects at the block-time level and cluster our standard errors at the entrepreneur level.

Table [Appendix J](#) presents our Generic ML estimates. Panel A presents estimates for models using our pre-specified set of variables. Panel B includes a broader set of baseline variables. In both panels the estimate ATE is just under 0.04 S.D. and far from statistical significance. In contrast, the estimate on β_2 is far from zero and is highly statistical significant, telling us that there is meaningful treatment effect heterogeneity in the data.

The next three rows present GATE estimates. The estimate for GATE₁ in both panels suggest that the average treatment effect for entrepreneurs with the lowest treatment effects is negative; for GATE₂ that the average effect is quite positive. The difference between these estimates is also statistically significant.

The final three rows in each panel show that across the GATE₁ and GATE₂ groups there are no statistically differences in the fraction of female entrepreneurs or entrepreneurs who have used ChatGPT before our study. In contrast, entrepreneurs the Generic ML algorithm estimates to have positive returns from our treatment are much more likely to be high rather than low performers. This result is consistent with our pre-registered heterogeneity analysis.

K Text Analysis

This section describes our text data and analysis in detail. [Subsection K.1](#) provides an overview of our three sources of text data along with summary statistics on these datasets. [Subsection K.2](#) walks through the construction of [Table 1](#) which shows the topic distribution of questions asked by the entrepreneurs in our study; this sub-section also describes the topic distribution for the AI generated responses. [Subsection K.3](#) walks through how we test for differences in the text for low and high performers, and for treated versus control entrepreneurs, by using word embedding vectors to predict performance and treatment

status. [Subsection K.4](#) describes our word- and embedding element-level analyses testing if the content of the AI’s advice impacts the business changes made by an entrepreneur. [Subsection K.5](#) details how we constructed our treatment effect embedding vectors to generate qualitative insights about how treated versus control entrepreneurs make different business changes. Finally, [Subsection K.6](#) describes our quantitative tests of the qualitative insights discussed in [Subsection K.5](#).

K.1 Overview and Summary Statistics

Our text analyses leverages three sets of data. First, we use data on the quantity and content of messages each entrepreneur sent to the AI assistant. [Table A17](#) Panel A presents summary statistics at the message level. These messages include questions asked of the AI, numeric followups (e.g. the number “3” in [Figure 1](#)), and non-business and other messages sent to the AI. That said, for most of our analysis we combine the set of messages sent to the AI into a single text “blob” per entrepreneur. Summary statistics for the overall set of text sent to the AI assistant are shown in Panel B of [Table A18](#). When we estimate the word embedding vector for the question asked of the AI assistant, we use this full set of messages sent to the AI by each entrepreneur.

Second, we use data on the responses sent by the AI to the entrepreneurs. Our AI assistant was completely responsive, so entrepreneurs only received messages from the AI when they sent in a message. These responses were often more structured, as [Figure 1](#) shows, and if possible they returned 3-5 pieces of advice. [Table A17](#) Panel B provides summary statistics on the number of pieces of advice generated and the length of the messages entrepreneurs receive. Panel B of [Table A18](#) shows summary statistics for the full set of text sent to each entrepreneur by the AI. As with the messages sent, we analyze the AI’s responses and estimate word embeddings by using the concatenating the full set of advice sent by the AI to the entrepreneur.

Third, we use data on the business changes each entrepreneur describes making. As part of our post-treatment surveys, we asked each entrepreneur to describe the changes they made to their business. If an entrepreneur did not make any changes, they did not have to describe any changes. [Table A20](#) shows that treatment and performance status does not

predict whether an entrepreneur makes any changes. In particular, entrepreneurs were asked the following survey questions during the final three survey waves:

- Please describe in detail any new products that you have implemented in the last 30 days. If you have not implemented any new products, you can leave this question blank.
- Please describe in detail any new services that you have implemented in the last 30 days. If you have not implemented any new services, you can leave this question blank.
- Please describe in detail any new business processes or strategies that you have implemented in the last 30 days. If you have not implemented any new business processes or strategies, you can leave this question blank.
- In the last 30 days, which change that you made to your business has had the largest impact on performance (for example, sales and profits)? Please describe it in as much detail as possible.

As with the messages sent and advice received, we analyze the business changes made—including our estimates of word embedding vectors—by first concatenating the full set of changes into a single text field.

Crucially, [Table A20](#) shows that treatment and performance (i.e., high vs. low) status does not predict whether an entrepreneur has question, answer, or business change text. As a consequence, we can analyze differences in text usage without worrying that our estimates are the by product of differences in who ask questions, who receives answers, and who describes making business changes.

K.2 Question and Answer Topic Analysis

K.2.1 Taxonomy Generation for Business Messages

To understand the topics the messages sent to and answers generated by the AI, we follow an approach outlined in [Choi et al. \(2023\)](#) and use GPT-4 to identify topics within the business related questions posed by entrepreneurs to the AI assistant. This process has four stages:

Stage 1: We had two human coders reader through all for the entire set of 4,810 identifying messages that were not numeric (e.g. following up with 1), not fragments and follow ups (e.g. “tell me more” “Can you explain how I might do 3?”), and that were not focused on business topics (e.g., a business owner who asked about how to quit smoking). Given that entrepreneurs ask a large number of follow ups we end up left with 1,392 messages that are business focused and have enough context to be classified into topics.

Stage 2: Using this sample of “business messages” we instructed GPT-4 using the prompt below to generate 3 to 20 topics to classify the messages in the corpus. We begin with 3 topics and incrementally increase to 20 topics. The end result was a set of 207 topics across all messages. For robustness, we randomly selected about 10% of the assigned topics and manually reviewed them to check the topics were sensible.

Listed below are queries submitted by SMEs entrepreneurs in Kenya to an AI Business Mentor designed to help them in their daily business operations.

The AI Business Mentor is designed to provide guidance and support to SMEs entrepreneurs in Kenya by answering their queries and providing them with valuable insights and advice. For the queries presented, provide **[num.topics]** high-level business topic in less than 4 words.

For example, if the query is “How can I improve my sales?”, the corresponding 4 topics could be [“Sales Strategy”, “Customer Acquisition”, “Marketing Campaigns”]

Return the topics for each query in a JSON format with the queryID, query and the corresponding topics.

The query list can be found below:

query_list

Stage 3: We then took the est of 207 latent topics and instructed GPT-4 to merge or eliminate redundant topics, narrowing them down to a more concise list. We then manually refined this list to remove overly specific topics, settling on a set of 10 topics. The topics are:

1. Financial Management
2. Marketing Strategies
3. Risk Management
4. Customer Acquisition & Retention
5. Business Planning & Growth
6. Technology Integration & Digital Marketing
7. Inventory & Operations Management
8. Market Research & Analysis
9. Livestock & Farm Management
10. Entrepreneurial Opportunities & Business Ideas

Stage 4: Using these ten high-level topics, we instructed GPT-4 to classify each business message into one of these categories. Messages could not be classified into multiple topics. The prompt we used for this classification is included below:

Listed below are queries submitted by SMEs entrepreneurs in Kenya to an AI Business Mentor designed to help them in their daily business operations. The AI Business Mentor is designed to provide guidance and support to SMEs entrepreneurs in Kenya by answering their queries and providing them with valuable insights and advice.

Provide a classification of the query using ONE high-level business topic from the list of topics below and categorize the queries accordingly. For example, if the query is "How can I improve my sales?", the corresponding topics could be ["Sales Strategy"] or ["Customer Acquisition"] or ["Marketing Campaigns"]

Here is the list of topics to choose from:

1. Financial Management
2. Marketing Strategies
3. Risk Management
4. Customer Acquisition & Retention
5. Business Planning & Growth
6. Technology Integration & Digital Marketing
7. Inventory & Operations Management
8. Market Research & Analysis
9. Livestock & Farm Management
10. Entrepreneurial Opportunities & Business Ideas

Here is the query list that you need to categorize: [query_list]

Return the topics for each query in a JSON format with the queryID, query and the topic.

We then used these topics to generate [Table 1](#) which shows the distribution of topics by treatment and performance status. Below we include the AI generated responses that correspond to the questions asked in [Table 1](#) to illustrate AI answers for each of the 10 topics.

K.2.2 AI generated responses for the example questions shown in [Table 1](#).

Here we include the AI generated responses that correspond to the questions asked in [Table 1](#) to illustrate AI answers for each topic of the 10 topics.

Entrepreneurial Opportunities & Business Ideas:

AI Response: “That’s a great idea! Setting up a cyber, printing and photocopy shop can be quite profitable in a town like Kericho. Here are some strategies you might find useful:

1. ****Research the Market****: Understand your potential customers’ needs and preferences. Find out where other similar businesses are located and what services they offer. This will help you identify gaps that you can fill to differentiate your business.

- Step 1: Identify potential customers (students, local businesses, government offices etc.)
- Step 2: Conduct informal interviews or surveys to understand their needs
- Step 3: Visit competitor shops to see their service offerings and pricing.

2. ****Location is Key****: Choose a location with high foot traffic such as near schools, colleges or office buildings.

- Step 1: Look for available spaces in these areas
- Step 2: Consider factors like rent costs, security, accessibility when choosing a location.

3. ****Diversify Your Services****: Apart from cyber services, printing and photocopying, consider offering additional services such as scanning, binding or laminating documents.

- Step 1: Assess the demand for these additional services through market research
- Step 2: Calculate the cost of adding these services against potential profit.

4. ****Quality Customer Service****: Ensure that your shop offers excellent customer service to retain existing customers and attract new ones.

- Step 1: Train yourself and any staff on customer handling skills
- Step 2: Implement measures like quick service delivery times or loyalty programs.

5. ****Maintenance of Equipment****: Regularly maintain your equipment to ensure smooth operations and reduce downtime due to breakdowns.

- Step 1: Have regular check-ups for your desktop & printer
- Step 2: Keep contacts of reliable technicians who can quickly fix any issues

Enter a number to learn more about one of these points on **setting up a cyber shop** or ask a new question.”

Livestock & Farm Management:

AI Response: “Understanding your situation as a maize farmer in Kiambu county, it’s clear that pests and diseases are a significant challenge. Here are some strategies to increase your profitability without spending more:

1. ****Crop Rotation****: This is a method of controlling pests and diseases by changing the type of crop grown in each field from season to season.

Benefits: It can break the life cycle of pests and diseases specific to maize, reducing their impact on your crops.

Implementation: In one planting season, grow maize. In the next, switch to another crop like beans or potatoes.

2. ****Intercropping****: This involves growing two or more different types of crops together in the same field.

Benefits: Some combinations of crops can deter pests or improve soil health, increasing yield without additional inputs.

Implementation: Consider intercropping maize with legumes such as beans or peas which also enriches the soil with nitrogen.

3. ****Natural Pest Control****: Use natural predators or homemade pesticides to control pest populations.

***Benefits*:** Reduces costs associated with buying commercial pesticides and is safer for the environment.

***Implementation*:** Introduce natural predators like ladybugs into your farm or use neem-based solutions as they are effective against many pests.

4. ****Proper Sanitation Practices**:** Keep your farm clean and free from plant debris where pests can breed.

***Benefits*:** Helps prevent infestations before they become problematic.

***Implementation*:** Regularly remove plant waste from your fields, especially after harvests.

5. ****Use Disease-Resistant Varieties**:** Plant varieties that have been bred to resist common diseases affecting maize in Kenya.

***Benefits*:** These varieties reduce losses due to disease while requiring fewer chemical treatments.

***Implementation*:** Consult agricultural extension services for advice on suitable disease-resistant varieties available locally.

Enter a number to learn more about one of these points on **improving profitability in maize farming** or ask a new question.”

Financial Management:

AI Response: “Managing debtors can be a challenging part of running a business. Here are some strategies you might consider:

1. ****Communicate Clearly**:** Politely but firmly explain to your customers that due to financial constraints, you’re unable to extend credit at this time.

2. ****Prepayment Policy**:** Implement a policy where customers must pay before receiving their coffee. This is standard in many businesses and most customers will understand.

3. ****Limited Credit**:** If you want to maintain some flexibility, consider setting a limit on how much credit one customer can have at any given time.

4. ****Loyalty Program**:** Instead of offering credit, consider creating a loyalty program where after purchasing a certain number of coffees, the customer gets one free.

5. ****Payment Reminders**:** For existing debtors, gently remind them of their outstanding balance and request payment.

Enter a number to learn more about one of these points on **managing debtors** or ask a new question.”

Business Planning & Growth:

AI Response: “That’s a great question! It sounds like you’re entering a competitive market, but there are always opportunities to differentiate yourself. Let’s explore some strategies that could help your t-shirt business stand out:

1. ****Unique Designs**:** Collaborate with local artists or students who have an artistic flair to create unique designs for your t-shirts. This will make your product distinctive and appealing to those who want something different from the mainstream options.

2. ****Quality Materials**:** Invest in good quality materials for your t-shirts. While this might slightly increase the cost of production, it can justify a higher selling price as customers often associate better quality with higher prices.

3. ****Targeted Marketing**:** Understand your target audience within the school and tailor your marketing messages towards them. For example, if you’re targeting sports teams, emphasize how comfortable and durable your t-shirts are during physical activities.

4. ****Customer Service Excellence****: Provide excellent customer service by being responsive to inquiries, handling returns/exchanges professionally, and delivering orders on time.

5. ****Competitive Pricing Strategy****: Price your products competitively by understanding what other suppliers are charging and offer occasional discounts or promotions to attract customers.

Enter a number to learn more about one of these points on **business differentiation** or ask a new question.”

Customer Acquisition & Retention:

AI Response: “Thank you for sharing your situation. It sounds like the competition is getting tough and we need to find ways to stand out and attract customers back to your motorcycle spares shop. Here are a few strategies:

1. ****Customer Loyalty Program****: Introduce a customer loyalty program where repeat customers get points for every purchase, which can be redeemed for discounts or free products in the future. This will encourage them to keep coming back.

2. ****Promote Your Business Locally****: Collaborate with local businesses and community events to promote your shop. You could sponsor a local event or offer special discounts to members of certain organizations.

3. ****Improve Customer Service****: Ensure that every customer who walks into your shop has an excellent experience. Train your staff on effective communication, product knowledge, and problem-solving skills.

4. ****Offer Value-Added Services****: Consider offering services such as free delivery within the city or basic maintenance checks at a minimal fee (or even free with larger purchases). This could make shopping at your store more convenient and attractive.

5. ****Social Media Presence****: Create accounts on popular social media platforms (like Facebook) specifically tailored for Kenyan audiences where you can share updates about new stock arrivals, promotions, or useful tips related to motorcycles and their maintenance.

Enter a number to learn more about one of these points on **improving business performance** or ask a new question.”

Marketing Strategies:

AI Response: “Thank you for sharing about your business. It’s great that you’re already using posters to promote it. Here are a few more strategies that could help attract new customers:

1. ****Leverage Social Media****: Create accounts on platforms like Facebook, Instagram, and Twitter where you can share pictures of your food, tea, and hotel environment. Engage with followers by responding to their comments and messages promptly. Regularly post updates about any new items or special offers.

Benefits: Social media is a powerful tool for reaching out to potential customers far beyond your immediate location.

Implementation: Start with one platform where most of your target audience likely spends their time (for example, Facebook). Set up a business page and start posting regular updates – photos of the day’s specials, behind-the-scenes glimpses into the kitchen or brewing process etc.

2. ****Customer Loyalty Program****: Implement a loyalty program where customers earn points each time they visit or make purchases which they can redeem later for free items.

Benefits: This encourages repeat visits from existing customers while also attracting new ones who are interested in earning rewards.

Implementation: Keep it simple - perhaps a stamp card system where after 10 purchases the customer gets a free meal or drink.

3. **Community Involvement**: Participate in local events or sponsor small community projects. This not only increases visibility but also builds goodwill within the community.

Benefits: Being actively involved in community activities helps create positive brand recognition.

Implementation: Look out for local events like festivals, sports events or markets where you could set up a stall offering samples of your tea/food.

4. **Partnerships**: Collaborate with other non-competitive businesses nearby to cross-promote each other's services/products.

Benefits: You get access to an entirely different customer base without spending extra on advertising.

Implementation: Find complementary businesses such as bookshops or craft stores; propose displaying each other's flyers/posters inside respective premises.

5. **Word-of-Mouth Marketing**: Encourage satisfied customers to refer friends and family members by offering them incentives like discounts on future orders when their referrals make purchases at your hotel.

Benefits: Personal recommendations from trusted sources often have higher conversion rates than traditional marketing methods.

Implementation: After providing excellent service, politely ask if they'd be willing to recommend you to others; provide incentive if possible.

Enter a number to learn more about one of these points on *customer acquisition* or ask a new question."

Technology Integration & Digital Marketing:

AI Response: "That's a great question! Before I can provide specific advice, could you please tell me more about your business? For instance, what industry are you in? What type of products or services do you offer? Also, it would be helpful to know if the majority of your customers are local or international."

Inventory & Operations Management:

AI Response: "Thank you for reaching out. Managing stock effectively is crucial to the success of your business. Here are some strategies that can help:

1. **Inventory Audit**: Regularly perform a physical count of your items to ensure what's on paper matches with what's in store. You could do this weekly or monthly depending on the size and turnover of your stock. - **Benefits**: Helps identify any discrepancies early, prevent loss due to theft, damage, or misplacement. - **How-to**: Choose a quiet time when sales activities are minimal. Count each item and record it manually or electronically.

2. **Categorization**: Group similar items together either by type, use, supplier, price range etc. - **Benefits**: Makes it easier to locate items quickly which saves time and improves customer service. - **How-to**: Identify common characteristics among your products and create categories based on them.

3. **Use Stock Cards**: For every item in your shop, maintain a card that records every purchase (inflow) and sale (outflow). - **Benefits**: Gives you real-time information about how much of each product you have at any given time. - **How-to**: Create cards with columns for date, description of transaction (purchase/sale), quantity in/out, balance after transaction.

4. **Implement FIFO (First In First Out)**: Always sell older stock first before newer ones especially for perishable goods. - **Benefits**: Prevents waste due to expiry or spoilage hence saving costs. - **How-to**: Organize your shelves so that older stocks are at the front while newer ones are

at the back.

5. ****Supplier Relations****: Build good relationships with suppliers to get favorable terms such as return policy for unsold goods or timely delivery of new stock. - Benefits: Minimizes risk of holding too much stock and ensures steady supply when needed - How-to: Pay suppliers promptly; communicate openly about needs & challenges; negotiate mutually beneficial agreements

Enter a number to learn more about one of these points on **stock management** or ask a new question.”

Market Research & Analysis:

AI Response: “Understanding your customers’ needs is crucial for expanding your product range and increasing profits. Here are some strategies:

1. ****Conduct Surveys****: Ask your customers directly about the types of products they would like to see in your store. - You can do this informally during their shopping or formally through a written survey.

2. ****Observe Shopping Patterns****: Pay attention to what items people often buy together or ask for frequently. - This can give you insights into other products they might be interested in.

3. ****Research Market Trends****: Look at popular products in similar stores or trends within your specific market. - This can help you identify potential new products that align with customer interests.

4. ****Test New Products****: Consider introducing new items on a trial basis and monitor customer response. - If a product sells well, consider making it a regular item.

5. ****Promote Your New Items****: When you add new products, make sure to promote them to your customers through signage, special deals, or verbal communication. - This helps increase visibility and interest in these items.

Enter a number to learn more about one of these points on **increasing profits** or ask a new question.”

Risk Management:

AI Response: “I understand your concern. Dealing with missing items can be a challenge for many businesses. Here are some strategies you might consider:

1. ****Inventory Management System****: Implement a simple system to track all items in the hotel, from kitchen utensils to furniture. This could be as straightforward as a ledger where entries are made every time an item is used or moved.

2. ****Staff Training****: Train your staff on the importance of managing resources and keeping track of items they use in their daily tasks.

3. ****Regular Audits****: Conduct frequent audits, not just financial but also physical checks of inventory. This helps detect any losses early and identify patterns if certain items frequently go missing.

4. ****Clear Policies on Theft****: Make sure there’s a clear policy that theft by employees will not be tolerated and could lead to termination or legal action.

5. ****Security Measures****: Consider basic security measures like installing lockable storage for valuable items or even surveillance cameras in key areas (if budget allows).

Enter a number to learn more about one of these points on **managing business assets** or ask a new question.”

K.2.3 Taxonomy Generation for AI Mentor Responses

Following the process outlined in [Subsection K.2](#), we generate a new set of topics for the responses generated by the AI business assistant in response to the entrepreneurs' business messages. The result of the first three steps is a list of 12 topics:

1. Financial Management
2. Marketing Strategies
3. Business Growth & Expansion
4. Operational Efficiency
5. Agriculture
6. Customer Engagement
7. Starting a Business
8. Product Differentiation
9. Market Research & Analysis
10. Risk Management
11. E-commerce Platforms
12. Human Resources

[Figure A10](#) shows the distribution of response topics for low and high performers. The distribution largely mirrors what we see in [Table 1](#). We again find no differences in the topics across the two groups. Lastly, for each of the 12 topics below we show an answer generated by the AI that represents the topic.

K.3 Predicting Treatment and Performance Heterogeneity

As discussed in the body of the paper, we test for differences in the questions asked, answers received, and business changes made by using word embedding vectors to predict treatment and performance (i.e., high vs. low) status. To do so, we first take the full text from each of our three data sources and estimate 3,072 dimension embedding vectors for each entrepreneur

using OpenAI’s `text-embedding-3-large` model. Using these vectors we then predict treatment and performance status using simple random forest models.

Specifically, we split our data into a training (80% of the data) and test set (20% of the data). We then use random forest models trained to word embedding vectors representing the question, answer, and changes text to estimate treatment and/or performance status. We report the models’ accuracy on the test sets in [Table A19](#).

K.4 Impact of Advice Text on Business Changes Text

Does the content of the advice an entrepreneur receives impact the types of changes they make to their business? While our experiment does not randomize the content of the AI generated advice, the structure of our intervention allows us to estimate the impact of advice on business changes simply by controlling for the content of the messages sent to the AI by the entrepreneur. Why? Because the AI only learn about the entrepreneur from the messages sent. Since the messages sent to the AI are fully observed, we can fully control for the content of the messages—conditional that our model of the language data generating process holds—to isolate the impact of the advice on the business changes made by the entrepreneur ([Pearl, 2009](#)). [Figure A11](#) illustrates the logic behind the “backdoor criterion.”

Turning to our analysis, in Panel of A of [Table A21](#) we show that when we assume a “bag of words” model that when one of the 12,762 non-stop words in the AI’s responses is present in an entrepreneur’s advice text it strongly influence the words used in the business changes text. Even after controlling for whether a word occurs in the messages sent to the AI by the entrepreneur, we find that when that same word appears in the AI’s responses it is 2.84% points ($p < 0.0001$, 95% CI = [2.345; 3.327]) more likely to be used in the entrepreneur’s business changes text. Given the base rate that a word is used is roughly 0.18 percentage points, the effect corresponds to a roughly 15X increase.

While this approach has the benefit of simplicity, the analysis fundamentally rests on the strong assumptions inherent in the “bag of words” model at the core of this analysis. In short, whether we control for the question words or not, the regression models in Panel A of [Table A21](#) assume that the impact of a word like “price” appearing in the advice text only has a direct impact the word “price” in the business changes text. Put differently, words are

treated as independent of one another (Grimmer and Stewart, 2013). This is clearly a strong simplifying assumption and its easy to think of why it might be violated. For example, the appearance of the word “price” in the advice text is likely to increase the changes of words like “discount” or “lowered” in the business changes text.

To address these concerns, we turn to the word embedding vectors we have for the questions, answers, and business changes text. In short, we swap out the 12,762 words for the 3,072 elements of each embedding vector. Instead of having dummies for whether the word is used, we have continuous values for each embedding element. The advantage of using embedding vectors is that it largely rule out semantic edge cases, for example that an entrepreneur asked a question about “prices” but received advice about and described changes about “pricing.” Panel B of Table A21 shows similar regressions as Panel A, but using the embedding elements. Our pattern of results holds. An increase in an embedding-element in the advice text, even after accounting for the question text, leads to an increase in that value in the embedding element representing the business changes text. These results strongly suggest that the impact of advice text on the business changes text is not merely an artifact of assuming a “bag of words” model.

K.5 Qualitative Analysis using Treatment Effect Embeddings

Subsection K.3 shows that we can use an entrepreneur’s business changes word embedding vector to predict their treatment and performance status. While the results in Table A19 suggests that there are differences in how high and low performers respond to the AI assistant, it does not point to the drivers these differences. To shed light on these differences, we again leverage word embedding methods.

Specifically, we calculate “treatment effect embedding vectors” for high and low performers. For high performers, we calculate the average business change embedding vector for treated and and control entrepreneurs. We then calculate the treatment effect embedding vector by taking the element-wise difference between these vectors. We then do the same for low performers.

Using these estimates, we then find pairs of control and treated business changes text that are similar to these treatment effect embedding vectors. While we can never know an

individual entrepreneur’s counterfactual business changes, by finding pairs of text similar to the treatment effect embedding vector we can shed light on the types of changes access to the AI induced. In [Table A24](#) and [Table A25](#) we qualitatively illustrate how treated versus control entrepreneurs differ in the types of business changes they make. Both tables show example pairs of treated versus control text sampled from the top 1% of most cosine similar text pairs to the high or low treatment effect embedding.

K.6 Quantitative tests of qualitative insights

Building on [Table A24](#) and [Table A25](#), we test if the qualitative insights highlighted in these tables hold when we test them quantitatively. To do so, we rely on simple word matching and counting methods.

For our word matching analysis we construct three measures. First, to test whether the entrepreneur mentions working with and learning from the AI, we take the raw business changes text and match it with the following regular expression pattern:

```
AI|[Cc]hat|[Gg][Pp][Tt]|[Aa][Ii][Mm]entor|
[Aa]rtificial[Ii]ntelligence
```

We classify an entrepreneur who matches this pattern as having described working and learning from the AI. To test if the entrepreneur thanks the PDF guide we distributed to the control, we test if the raw business changes text matches the following regular expression pattern:

```
[Gg]uide|[Pp][Dd][Ff]|[Mm]anual.
```

Third, to test if the entrepreneur offered discounts or invested in advertising we match whether the raw business changes text matches the following regular expression:

```
[Ff]ree|[Dd]iscount|[Rr]educ|[Oo]ffer|[Aa]dver
```

We deliberately do not match against “marketing” as many entrepreneurs how they tried non-paid marketing and in [Table A25](#) we see no differences in the use of this term across the groups, instead the differences appear to be about whether the entrepreneur invests in paying for a billboard or online ads. We include “free” because many entrepreneurs mention offering free delivery and other services that implicitly serve as a discount. Finally, we note that these regular expressions are far from perfect, but are designed to capture the word-level

differences present in [Table A24](#) and [Table A25](#).

We then simply regress using OLS whether an entrepreneur matched the regular expression on the entrepreneur’s treatment and performance status. For these measures, since working with the AI, building on the guide, or lowering prices requires making a business change, We assume that entrepreneurs that do not describe making any business changes do not match any of the regular expressions. [Table A22](#) shows regression estimates using the mentions AI and Guide measures. ?? shows results for whether the entrepreneur mentions discounting or running ad campaigns.

Our final measure is the count of uncommon words used in an entrepreneur’s business changes text. To create this measure, we started by removing capitalization from the business changes text, tokenizing words, and then removing stop words. We then classified words used as uncommon if they occur in 5 or fewer business change descriptions. These words include “150cc,” “arcade,” “biodegradable,” “detergent,” “headphones,” “peanut,” “rabbit” and “uganda.” Common words include “advertising,” “customers,” “market,” “price,” and “vegetables.” Given that entrepreneurs who write more will mechanically have more uncommon words, we calculate both the total number of uncommon words and a control variable that measures the total number of words used in the business changes text. Finally, we focus on the intensive margin, restricting our sample to entrepreneurs who use at least one word in their business changes text.

[Table A23](#) presents Poisson regressions testing if treatment and performance status predict the count of uncommon words used. We use Poisson regressions since the outcome is a highly skewed count. [Figure 5](#) Panel B shows estimate incidence rate ratios.