Regular article

Training microentrepreneurs over Zoom: Experimental evidence from Mexico ☆

Elwyn Davies [a], Peter Deffebach [b], Leonardo Iacovone [a], David McKenzie [c],*

[a] *World Bank, United States of America*
[b] *Boston University, United States of America*
[c] *Development Research Group, World Bank, United States of America*

ARTICLE INFO

ABSTRACT

Standard in-person business training programs are costly and difficult to scale to the millions of microenterprises in the developing world. We conduct an experiment to test the feasibility, cost-savings, and impact of delivering live training sessions over Zoom to microentrepreneurs in Mexico and Guatemala. We demonstrate that it is now feasible to both recruit and train self-employed women online, covering a wide geographic area, with few technology issues. However, the cost-savings over in-person classes are less than expected. Training improved business practices and performance over 2 months, but impacts had dissipated within 6 months.

1. Introduction

Developing countries are the home to millions of microenterprises, which provide an important source of income for the poor. For example, the Mexican Economic Census found 4.1 million firms with zero to ten workers (INEGI, 2016). Many governments and NGOs offer business training programs to try to help owners of these firms improve their business practices and increase their incomes. A recent meta-analysis of business training experiments found statistically significant, but modest, average impacts of a 4.7 percent improvement in sales and 10.1 percent improvement in profits (McKenzie, 2021). The most typical training programs take place in-person in classroom settings, requiring groups of 20–30 business owners to travel to a common location for several days, with an average cost of $177 per participant (Van Lieshout and Mehtha, 2017). This raises questions and skepticism about the cost-effectiveness and scalability of such programs (e.g. Fox and Thomas (2016)), and how they can be expanded to a scale where they can reach thousands or millions of firms.

Digital technology offers the potential to both lower the costs of delivering training, and to enable it to be scaled across a wider geographic area. However, experience in developed countries with asynchronous voluntary learning on massive open online courses (MOOC) platforms such as Coursera shows incredibly large drop-out rates (Rivard, 2013). Moreover, poor entrepreneurs in a developing country setting may face further technological obstacles in accessing online training. Reich and Ruipérez-Valiente (2019) note that almost all of the pre-pandemic growth in MOOC registration and certification came from high-income countries. An alternative to self-paced asynchronous courses without any human interaction is to hold video-based synchronous courses that mimic the format of in-person training, while allowing for remote delivery. The rise of Zoom and other video meeting platforms opens up this possibility, but raises three key questions: how feasible is this with microenterprise owners in a developing country setting? How much of a cost-saving is achieved by offering live classes via Zoom? And is remote training effective in improving business practices and performance?

We implement a field experiment with 2208 female microenterprises recruited via microfinance partner contacts and Facebook advertising throughout Mexico and Guatemala to answer these questions. The recruited businesses are small: 42 percent have no employees, and

only 65 percent kept business records at baseline. The main sectors are food, beauty and clothing, and handicrafts. Firms were randomly allocated into a treatment group of 1513 firms, and control group of 695 firms. The treatment group were offered a business training program taught live in small groups over Zoom in nine 2-hour sessions over 4 weeks, while the control group were provided with four asynchronous online modules. We test both a "top-down" training approach in which the topics are chosen by the implementing NGO, and a "bottom-up" approach in which training topics are requested by the entrepreneurs. In practice these overlap substantially in topics covered and have similar impacts.

We find that it is now technologically feasible to provide business training over Zoom in a developing country setting, with attendance rates not too dissimilar to in-person training. We were able to recruit small businesses from all 32 states in Mexico, and from Guatemala. They were able to connect to classes via their own mobile devices, with few technology issues. Eighty percent of those assigned to training started at least one session, and 61 percent completed all sessions to graduate. This demonstrates the feasibility of scaling to thousands over a wide geographic range. However, we found three factors limit the ability to scale to tens or hundreds of thousands. First, the conversion rate from advertising and approaching entrepreneurs through partners was low, and it required multiple rounds of recruitment to obtain our sample. Second, while the businesses recruited are small and have household incomes much lower than the Mexican average, the women who self-select into training are younger, more educated, and are more likely to have an employee than the average Mexican microenterprise, suggesting limits on reaching the poorest and least educated. Finally, the cost savings from switching from in-person to online training are not that large ($50 versus $62 per participant), reflecting that the main costs of instructor time, recruitment, and material development are similar for on- and offline synchronous training.[1]

Remote training via Zoom does improve business practices and business performance in the period immediately after training, but these impacts do not last. We conducted follow-up surveys approximately 2 and 6 months post-training. After two months, women assigned to our Zoom training treatment have significantly improved their business practices by 5.4 percentage points. Monthly sales are 4100 pesos ($240) higher, which is statistically significant and a 24 percent increase on the control mean, while monthly profits are a statistically insignificant 648 pesos ($38) higher, or 10% of the control mean. The 6-month impacts are all significantly smaller than the 2-month impacts, and are not statistically different from zero. We examine treatment heterogeneity using the traditional interaction approach, by examining quantile treatment effects, and using the generic machine learning approach of Chernozhukov et al. (2020). There is limited predictable heterogeneity in treatment impacts, and we do not find any subgroup has lasting treatment impacts. This lack of persistent impact appears to reflect both the treatment group stopping doing some of the business practices it had adopted, as well as some control group catch-up. In a changing business environment, training appears to have sped up the process of getting firms to examine their records and make budgets, but since it did not significantly improve their marketing or personal initiative, this may explain why it was unable to generate sustained increased in sales.

This paper contributes to literatures on interventions to help the self-employed, and on remote education and training. The main contribution is to the literature on improving business practices and management in firms, reviewed in McKenzie et al. (2021). The majority of this literature has focused on in-person training programs that can be hard to scale. Improvements in digital technology combined with the COVID-19 pandemic have led to different approaches to testing digital delivery.

One approach has been to use asynchronous content. For example, Jin and Sun (2021) offer short training tasks to Chinese online sellers, and Estefan et al. (2023) offer 1–7 min video capsules to Guatemalan chicken franchise owners. Such an approach can scale cheaply, but tends to be short in duration limiting what can be taught, does not allow for interaction with an instructor, and can suffer from low take-up: only 12.6 percent of sellers in Jin and Sun's study finished even one task. Another alternative has been to work with somewhat larger or growth-oriented firms, and offer one-on-one virtual coaching (Anderson et al., 2022), or live online sessions combined with one-on-one coaching (Cusolito et al., 2023). This allows for more interaction and tailored content, but is far more expensive and less likely to scale. Our study compliments these approaches by testing the feasibility, cost, and effectiveness of offering synchronous business training content with a poorer and more typical set of microentrepreneurs.

2. Context, content, sample, and data

We partnered with the Mexican NGO Crea Comunidades de Emprendedores Sociales, which has been providing programs in Mexico since 2008 for women entrepreneurs in economically marginalized areas. They typically offer in-person training courses to women, funded by a range of government and private sector partners. When the COVID-19 pandemic hit, in-person training was no longer feasible, and they were interested in seeing whether they could instead deliver training to women online.

2.1. Recruitment and enrollment in the program

CREA launched their program under the name *Fortalece tu negocio* (Strengthen your business). The program was advertised as a free online course where microenterprise owners could learn resilience, costing, prices, marketing, e-commerce, and making their business plan. Recruitment took place in ten waves spread between November 2020 and November 2021. Recruitment mainly took place through social media channels, the principal one being paid Facebook advertising. In addition the program was advertised on the social media pages of CREA and some of its funders, through SMS messages and emails sent to a sample of firms in a Mexican government database, and through flyers in Mexico City. Guatemala was added as a second country after five rounds of recruitment had taken place, in order to test the feasibility of further geographic expansion. Overall, 65% of the Mexican sample and 86% of the Guatemalan sample were recruited through Facebook.

Facebook usage is high in Mexico, with an estimated 90 million users in 2022, which is 84 percent of the population aged 10 and over.[2] This illustrates the potential for online recruitment to reach large numbers of microenterprises. To participate in the program, individuals had to click on the advertisement and fill out a short pre-registration form indicating interest, and then they were invited to attend an online information session to find out more about the course. They then registered by filling out a form that serves as our baseline data. CREA's paid advertising campaigns were seen by 3.3 million unique viewers. However, as is typical with online advertising, the conversion rate is low: 52,719 (1.6%) of viewers clicked on the link, 10,700 pre-registered, and 2208 registered for the program across all sources (1478 from the Facebook ads). The estimated recruitment cost from advertising was $3.38 per person in our sample.

Table 1 provides summary statistics of the experimental sample, and compares them to a representative sample of Mexican female entrepreneurs taken from the 2023 ENOE, and to a sample of CREA's in-person clients from Iacovone et al. (2018). Online recruitment was

---

Table 1
Balance tests for baseline covariates.

Control variable	Mean (1)	SD (2)	Control N = 695 (3)	Treated N = 1513 (4)	P-value (Joint = .5) (5)	2023 Mexico ENOE Survey Mean (6)	SD (7)	2014 CREA Survey Mean (8)	SD (9)
Years of business operation	4.21	5.32	3.98	4.32	0.099	9.73	10.3	2.14	2.65
Is a family business	0.510	0.500	0.498	0.516	0.592				
Age	40.1	10.1	39.9	40.2	0.418	46.0	14.2	42.3	11.3
Married	0.555	0.497	0.522	0.570	0.055	0.594	0.491		
In State of Mexico or Mexico City	0.276	0.447	0.269	0.280	0.707	0.210	0.407	0.427	0.495
In Guatemala	0.075	0.264	0.079	0.073	0.998				
Attended university	0.485	0.500	0.498	0.479	0.536	0.171	0.376	0.378	0.485
Household earnings > 8000	0.457	0.498	0.450	0.461	0.494	0.448	0.497	0.436	0.496
Sales in past month	7824	15,096	7733	7866	0.796			22,257	38,218
Profits in past month	1966	3617	1983	1958	0.740			8600	14,700
Any employees	0.585	0.493	0.588	0.583	0.663	0.273	0.445	0.400	0.490
Number of employees	1.52	2.40	1.43	1.56	0.224	0.544	2.75	0.968	6.68
Keeps written accounts	0.645	0.479	0.645	0.646	0.858	0.381	0.486	0.733	0.443
Index of marketing practices	0.378	0.211	0.367	0.383	0.036				
Index of accounting practices	0.271	0.235	0.273	0.270	0.975				
Index of planning practices	0.170	0.262	0.164	0.173	0.143				
Food sector	0.320	0.466	0.344	0.309	0.037				
Beauty sector	0.104	0.306	0.099	0.106	0.605				
Handicrafts sector	0.101	0.302	0.095	0.104	0.433				
Service sector	0.302	0.459	0.294	0.306	0.453			0.298	0.458
Essential business	0.189	0.391	0.200	0.184	0.284				

Notes: Baseline characteristics of firms involved in the program shown in first five columns. Characteristics of a representative sample of Mexican female entrepreneurs shown in columns 6 and 7 are from the 2023 ENOE (National Survey of Occupation and Employment). Columns 8 and 9 show characteristics of CREA's in-person training clients taken from a 2014 survey. Not all characteristics are available in these other surveys. *P*-value in Column (5) correspond to the effect of treatment on the baseline covariate, controlling for strata fixed effects.

successful in scaling across a wide geographic range. The businesses are based in all 32 states of Mexico, with only 28% coming from Mexico City and the neighboring State of Mexico. 8% of the women are in Guatemala. The businesses are small, with only 59% having any employees, and an average of 1.5 employees, and average monthly profits of approximately 2000 pesos ($100). Firms were required to be in operation for at least a year to join the program, and the average years of operation is 4 years. Firms cover a heterogeneous mix of industries, but the majority involve women making and selling some sort of product, while 30 percent are in services. The most common sectors are baked and prepared food, beauty, handicrafts, and clothes and accessories. The women running these businesses have an average age of 40, and 48 percent have some university education. At baseline there was plenty of scope to improve their business practices: while 65% said they kept written accounts, they were only doing 38 percent of the marketing practices, 27 percent of the accounting practices, and 17 percent of the planning practices that the training was intended to cover.

We see that the women who are recruited for Zoom training are on average younger and more educated than the average Mexican female entrepreneur, and more likely to have an employee in their firm. They are more similar to the typical in-person clients of CREA in age, but also more educated. In-person clients tend to be concentrated in a few states, whereas online recruitment gives a sample more geographically representative. Household income levels are similar to those of the average microentrepreneur and profits and sales are substantially lower than for in-person clients, although this could reflect the impact of the COVID-19 pandemic on demand.

2.2. Random assignment and training content

Firms were stratified by recruitment wave, country, terciles of baseline sales, and terciles of baseline business practices and then randomized into a control group of 695 businesses, and two training treatment groups of 1513 businesses.

The two treatment groups varied in how the content of their training was determined. The first, which we call 'top-down', is the more standard structure, where the training organization (CREA) determined

which topics should be taught. Trainers covered four modules that covered resilience and self-determination (drawing on aspects of personal initiative training); costs, prices, and finances; marketing and e-commerce; and the business Canvas tool and business model for planning. The second approach, which we call 'bottom-up', had participants collectively meet in their first class and help determine which topics they were most interested in being covered. In practice, there was large overlap between the topics and material in the two treatments, perhaps in part due to the advertising for the program emphasizing certain topics. Appendix A provides more details on the content and overlap of the two types, and shows we cannot reject equality of treatment effects across these two groups. Given the similarity of topics and effect, we therefore pool the two into a single treatment group for our main analysis.

Women selected for treatment were offered a choice among several time slots in order to attend live classes over Zoom. Given that each recruitment round only had between 40 and 120 in each treatment group assignment, and the need to have several class time offerings, this meant that the typical online class only had around 20 participants. Training took place 2 or 3 times a week, typically in the evenings, in nine two-hour sessions conducted over Zoom, for a total of 18 h of training. This was supplemented with several take-home exercises for the participants to do.

The control group was offered an asynchronous online training option, where they could access the slides and webinars of the same content as the 'top-down' treatment through the CREA course platform by setting up an account. At the end of each module there was a small quiz, and entrepreneurs were considered to graduate from the program if they registered and completed all four modules. This enables us to see how much live Zoom classes add value over a zero marginal cost asynchronous option, and also was intended to reduce the risk of attrition by having offered something to all firms. Our prior was that take-up of this offering would be low. This was the case in practice, with only 11 percent of control firms completing at least one module, and fewer than 7 percent completing all four modules.

2.3. COVID-19 context

Our project takes place between November 2020 and July 2022, and so covers a period in which the global COVID-19 pandemic was taking place. While the pandemic limited the willingness of organizations such as CREA to offer large in-person gatherings, Mexico had somewhat limited and loosely enforced shutdowns, which varied by state.[3] By the time our training started and follow-up surveys were taking place, the initial period of most severe shocks and shutdowns had already taken place, and during our follow-up surveys we find 90 percent of firms on average to be open and making sales. In Appendix D we test for heterogeneous impacts by whether firms are classified as essential or non-essential businesses from the point of view of COVID-19 regulations, and find no significant differences. Mexico's economy grew at 4.7% in 2021 and 3.1% in 2022, recovering from the pandemic year of 2020. Therefore, firms were in a situation where they could largely operate, the economy was recovering, and they could use tools taught in the course.

2.4. Data collection and measuring impacts

We worked with Innovations for Poverty Action (IPA) Mexico to conduct two rounds of follow-up surveys. The first took place two months after training started (January 2021–January 2022), and was intended to measure short-term impacts and see whether firm owners had implemented some of the practices they were taught in training. The second took place after 6 to 8 months (August 2021–July 2022), to see if these impacts were sustained. Since the participants were recruited online from across Mexico and Guatemala, follow-up took place through a combination of phone calls and online questionnaires.

After multiple attempts at re-contacting firm owners, we were able to re-interview 1592 of the 2208 entrepreneurs at 2 months (72%, 66% control, 75% treatment) and 1613 at 6 months (73%, 70% control, 74% treatment). Appendix B shows that the sample answering the surveys remain balanced on baseline characteristics. We account for possible bias due to selective attrition in several ways. Our main specification (noted below) uses the post-double-selection lasso of (PDS lasso) of Belloni et al. (2014). This selects covariates that either predict the outcome of interest (which can potentially improve power), or that predict treatment status (which could arise from unbalanced attrition). Appendix B shows our results remain robust to alternative approaches to accounting for attrition, such as probability re-weighting, or using bounding approaches.

We supplement these quantitative surveys with a qualitative survey of 20 treated firms, selected to comprise of a sample of 10 firms that had attended training and had improved their sales and business practices a lot by the 2-month survey, and 10 firms that had attended training but not shown improvement. We also use our own observations from observing training sessions to provide more qualitative information on content and process.

The main primary outcomes of interest are those that are the focus of the majority of the business training literature: whether training gets business owners to adopt new business practices, and whether it improves business performance in terms of profits and sales. Our AEA registry includes a short pre-analysis plan specifying these measures.

Due to the geographically disbursed locations of our firms and the aftermath of the pandemic still limiting fieldwork, our follow-up surveys take place by phone and online, rather than in-person. A potential concern may then be whether answers to questions about business practices will be reliable, and whether there is a tendency for those who have gone through training to overstate the practices

they are using. McKenzie and Woodruff (2017) examine the accuracy of self-reports of business practices in Sri Lanka and Nigeria, by sending experts to do business practice audits and comparing to survey responses. They find a high correlation between audited and self-reported responses, and no systematic differences for those who had been through business training. Bloom et al. (2019) likewise find a high correlation between self-reported mail and online responses to management questions and scores from open-ended double-blind phone discussions. Anderson and McKenzie (2022) find treatment effects on business practices which could be verified through photos and physical proof to be at least as large as the impacts on practices that were completely reliant on self-reports. Such literature provides us with some confidence that responses are reliable. Furthermore, if firm owners who went through training were systematically overreporting their business practices, we would expect this to occur for all practices used in training, whereas we shall see that there is no reported improvement in marketing practices.

We estimate the effect of being assigned to training using the following specification for outcome Y for firm i in randomization stratum s:

$$Y_i = \alpha + \beta \text{Treat}_i + \gamma Y0_i + \sum_{s=1}^{S} \delta_s 1(i \in s) + \phi' X_i + \epsilon_i \qquad (1)$$

This regression includes the lagged outcome variable ($Y0$) where available, dummies for the different randomization strata, and a set of control variables X selected via PDS Lasso. Robust (Eicker–White) standard errors are used. The coefficient of interest β corresponds to the intention-to-treat effect, which is the effect of being offered the live Zoom training, compared to just being offered the asynchronous version in the control group. We also run a stacked version of this equation which pools both survey rounds and allows us to test for equality of treatment effects over time.

3. Feasibility, cost, and effectiveness

We start by examining whether Zoom training is feasible in a developing country microenterprise setting, then discuss the costs of providing this training, before turning to measuring training effectiveness.

3.1. Feasibility, take-up, and attendance

Women who signed up for the course knew that it would be an online program and had managed to watch the short information session telling them some details about the training. Nevertheless, as an emerging technology in a developing country setting, we were still unsure how feasible live Zoom training sessions would be.

IPA Mexico monitored 75 training sessions to observe how frequently technology issues occurred. 3 of the sessions had to be rescheduled due to electricity cuts or a hurricane, but otherwise technical issues involving computers, cameras, and microphones were not a major issue. There were occasional connectivity issues due to slow connection networks or to entrepreneurs' data plans finishing, but these connections were usually reestablished within a few minutes. Participants typically used their mobile phones to join the Zoom sessions. Instructors used the chat and microphone features to get some questions and feedback from participants: in the average monitored session just over 80 percent of participants typed something in the chat, and approximately half turned on their video briefly (they kept it off most of the time to conserve data).

Take-up and attendance were then much higher than has been the case in most voluntary asynchronous courses. 80.7 percent of those assigned to treatment attended at least one session, with a mean of 5.5 sessions attended and 61.4 percent completing the course. This is in line with the average take-up rate for in-person training classes of 65 percent (McKenzie and Woodruff, 2014) and higher than Iacovone et al.

Table 2
ITT effects on primary outcomes at 2 and 6-months.

Dependent variable	2-month endline			6-month endline			
	N (1)	Control mean (2)	ITT (3)	N (4)	Control mean (5)	ITT (6)	Diff. (7)
Index of personal initiative	1592	4.36	0.027 [0.035]				
Score on mock test	1592	0.668	0.024 [0.012]**				
Index of planning practices	1592	0.417	0.117 [0.019]***	1613	0.518	0.009 [0.020]	−0.109 [0.023]***
Index of accounting practices	1592	0.541	0.069 [0.015]***	1613	0.590	0.019 [0.016]	−0.050 [0.017]***
Index of marketing practices	1592	0.487	0.016 [0.014]	1613	0.503	−0.006 [0.014]	−0.022 [0.016]
Index of business practices	1592	0.502	0.054 [0.013]***	1613	0.545	0.007 [0.014]	−0.047 [0.015]***
Sales in past month	1591	17,023	4112 [1461]***	1612	15,339	−981 [1265]	−5,093 [1444]***
Log sales in past month	1372	9.19	0.107 [0.059]*	1384	8.79	−0.057 [0.074]	−0.163 [0.077]**
Profits in past month	1591	6309	648 [506]	1612	4887	−471 [351]	−1119 [504]**

Notes: **Personal initiative** is an index of 7 questions measuring personal initiative, with a higher score denoting more initiative. It was only asked in the 2-month survey; **Score on mock test** is the proportion right on an 11-question knowledge measure, only measured at the 2-month survey; **Index of planning practices** is the proportion of 3 planning practices used; **Index of accounting practices** is the proportion of 10 accounting practices used; **Index of marketing practices** is the proportion of 9 marketing practices used; **Index of business practices** is the proportion of all 22 business practices used; **Sales in past month** is sales measured in Mexican pesos (winsorized at the 99th percentile); **Log sales** is log of sales in the past month for firms with positive sales; **Profits in past month** is profits in Mexican pesos (winsorized at the 1st and 99th percentiles). Regressions control for randomization strata, baseline value of outcome where available, and additional controls selected by pdslasso. Robust standard errors in parentheses.
* Denote significance at the 10 percent level.
** Denote significance at the 5 percent level.
*** Denote significance at the 1 percent level.

(2018) find for CREA's in-person training, where 69% started the course and 45% completed it. Figure A2 shows attendance rates by session. They fall over the course, but not steeply, from 72% for the first session to 60% for the last session. In contrast, take-up from the control group for the asynchronous materials was low: only 11.3% completed the first module, and only 6.6% completed all 4 modules.

3.2. How much cost-saving is there from zoom training?

One motive for considering business training by Zoom is the potential to lower costs. We worked with CREA to collect cost data on provision of training and to compare it to their cost structure when offering in-person training. The estimated cost per participant in Zoom training was $50 in Mexico and $56 in Guatemala. This covers the cost of personnel for recruitment and training, technology costs such as Zoom license fees, data plans, paid Facebook advertising, and other recruitment costs. The costs of personnel, especially the trainer, are the main cost. While in principle these personnel costs could be lower than in person if the online trainers were able to teach more women at the same time, in practice the difficulty of recruiting large numbers of women who all wanted to start and attend sessions at the same time meant that class sizes were similar to an in-person class. As a result, the estimated cost of an in-person class is not that much higher: $62 in Mexico. In-person classes involve some costs for trainer transport and venue rental, and fewer technology costs, but even with in-person training the personnel costs are 79% of total cost. However, this is based on having microentrepreneurs show up in-person at places where CREA already operates. In contrast, if we were to take the geographic spread across all states of Mexico and also in Guatemala, it would be much more expensive for CREA to travel and set up new trainings in all of these locations.

3.3. How effective is zoom training?

Did business owners learn anything from online training, implement what they had learned, and experience changes in business outcomes?

We answer these questions using our 2-month follow-up survey, presenting intention-to-treat estimates in Table 2. There is a statistically significant, but small impact of 2.4 percentage points on business knowledge as measured by an 11-question test. This is consistent with the small effects on test-assessed knowledge seen in in-person business training and financial education programs (Carpena et al., 2019; McKenzie and Puerto, 2021).

To measure the impact on business practices, we implement a slightly modified version of the practices in McKenzie and Woodruff (2017). We measure 22 business practices consisting of 9 marketing practices (e.g. monitors competitor's prices, uses special offers), 10 accounting practices (e.g. keep written records, separates household and business accounts), and 3 planning practices (e.g. has a written budget, sets sales goals). The control group are doing 50.2 percent of these practices, and we find that being assigned to treatment results in a statistically significant improvement of 5.4 percentage points. The magnitude of improvement is similar to the impact of an in-person course like the ILO SIYB course (McKenzie and Woodruff, 2017). The largest improvement comes in planning practices (11.7 percentage points), followed by accounting practices (6.9 percentage points). In contrast, the impact on marketing practices is small (1.6 percentage points) and statistically insignificant. We also find a small and statistically insignificant improvement in an index of personal initiative based on Campos et al. (2017). These results accord with our qualitative interviews, where personal initiative and marketing were the topics least remembered by participants, while finance and planning had the highest recall.

There is some evidence that this improvement in business practices is accompanied by short-term improvements in business sales. When measured in levels, monthly sales increase by 4113 pesos ($240) relative to a control mean of 17,023 pesos, a 24 percent increase. 90 percent of control firms and 91.9 percent of treated firms are open at the time of the 2-month survey (Table C.1), and those that are closed are coded as having zero sales. Taking log sales conditional on being open and making positive sales, the increase is 11.2 percent, which is not statistically significant. Monthly profits increase by 648 pesos
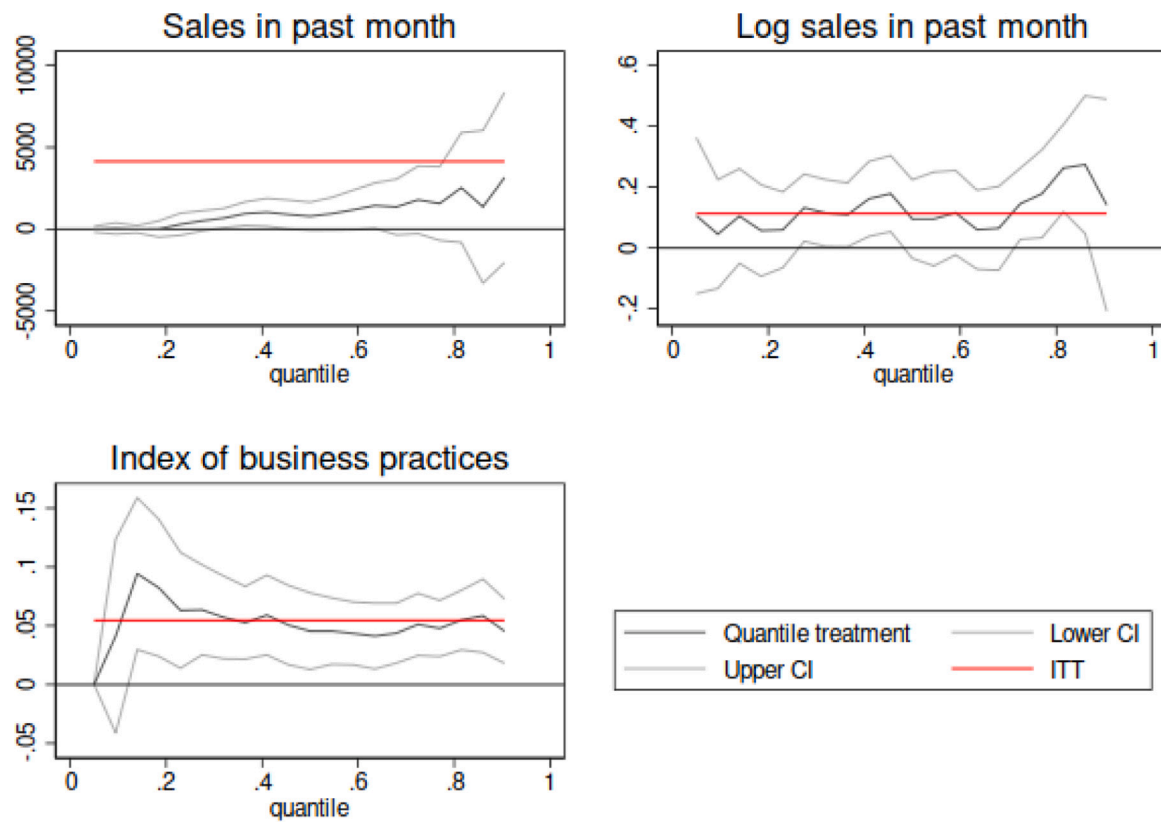
Fig. 1. Quantile treatment effects at 2-months.

($38), which is 10 percent of the control mean and not statistically significant.

In Appendix B we examine robustness of these results to attrition and outliers. We show the impacts are similar if we probability re-weight for attrition, and if we employ the Behaghel et al. (2015) approach of dropping the most difficult to contact treated firms to equalize response rates with control firms. Our business practice results are more robust to the possibility that the additional control group attritors are better than average than is the impact on sales. To examine how much our results are being driven by a few observations, we use the approximate maximum influence perturbation approach of Broderick et al. (2023) to see how sensitive the results are to removing a small fraction of the data. Table B.5 shows our business practice impacts are quite robust to selectively removing data (we would need to selectively drop almost 3.7 percent of the sample to change the sign), whereas the sales impact would change sign by dropping only 1.7 percent of the sample.

As an additional way of seeing whether these impacts are concentrated in a few firms or more widespread, we estimate quantile treatment effects. Fig. 1 plots the impacts and compares them to the ITT impact shown in Table 2. Quantile treatment effects for the level of monthly sales are well below the ITT for all but the top decile, showing that the large magnitude of the average improvement is indeed driven by the top of the distribution. However, we would expect training to be result in a constant percentage increase in sales, than the same level increase in sales regardless of initial firm size. This is the case, with the quantile impacts on log sales relatively constant across all quantiles and similar to the estimated ITT average impact. Likewise, we see the quantile treatment impacts on business practices are positive and significant and of similar magnitude across most of the distribution.

3.4. Do these impacts last?

Columns 4–7 of Table 2 show that none of these impacts persist at 6 months. The estimated impact on business practices has fallen to

0.8 percentage points, which is not statistically different from zero, and is statistically different from the 2-month impact. The estimated impacts on sales and profits are all negative in sign, and not statistically different from zero, and likewise are statistically different from the 2-month impacts. This difference is not a result of changing sample composition: Appendix Table B3 shows the results are similar if we restrict analysis to the balanced panel of firms.

We investigated treatment effect heterogeneity to examine whether training had sustained impacts for some subgroup of the sample, even if the overall impact disappeared. We explored two approaches to examining heterogeneity in Appendix D. The first is to examine treatment interactions with firm and owner characteristics. We find the initial impacts of training appear to be higher for those owners with more personal initiative, but even this subgroup does not have lasting impacts. Second, we use the generic machine learning approach of Chernozhukov et al. (2020) to test whether there is predictable heterogeneity in treatment effect based on a set of baseline covariates, and cannot reject that there is no predictable heterogeneity. The drop-off in treatment effect therefore seems widespread.

Fig. 2 graphically shows this reversal by showing the distribution of changes in business practices, sales, and profits by treatment status between the baseline and 2-month survey, and then between the 2-month survey and 6-month survey. Three results are apparent. First, not only is there a lot of volatility in sales and profits, but we also show that there is considerable churn in business practices. This is not something that has been documented in previous literature. Even in the control group, many firms are starting and stopping practices between survey rounds. Second, between the baseline and two months, we see relatively more treated firms adding business practices and fewer ones dropping them, and relatively more treated firms growing profits and sales than the control group. Third, in contrast, between 2 and 6 months we see relatively more of the control group adopting new practices, whereas more of the treatment group are dropping business practices, and more of the treated experience a drop in sales and profits. This figure also
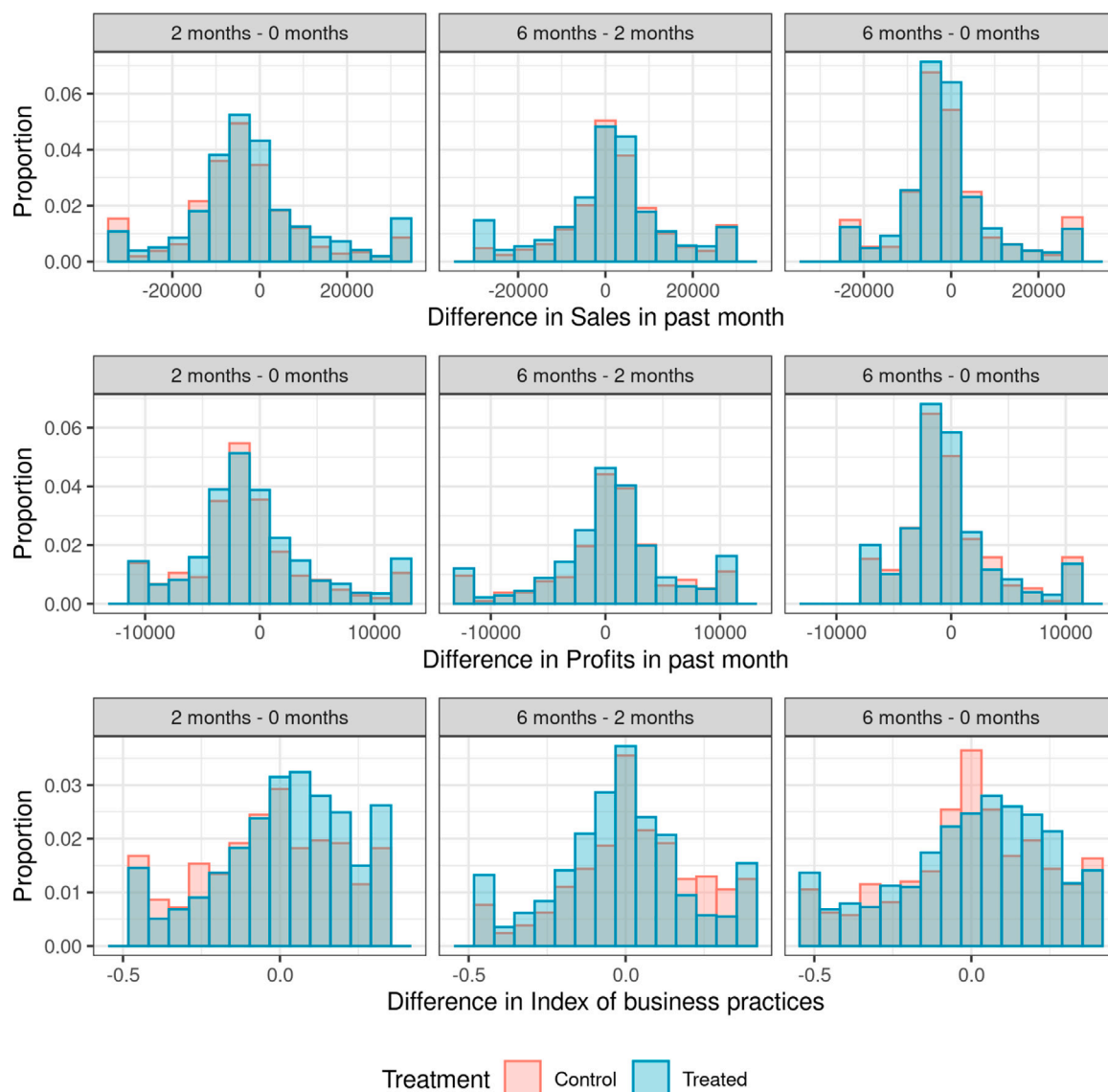
Fig. 2. Distributions of changes between survey waves.

helps show that the difference in 2- and 6-month results is not being driven by a few observations, but is instead visually apparent in the distributions.

We dig deeper into this churn in business practices in Table 3, looking at the individual practices that make up the planning and accounting practice indices (Appendix Table C2 does the same for marketing practices). By looking at specific practices, we can examine whether the lack of sustained impact on business practices is due to the treatment group being more likely to stop doing practices (falling back), or due to the control group being more likely to add new practices (catching up). Column 1 reports the baseline mean doing each practice, and then columns 2 and 3 the 2- and 6-month treatment impacts on that specific practice. Column 4 then documents how much churn there is in the control group's use of each practice between the 2- and 6-month follow-ups. For example, 32 percent of the control group either switched from having a written budget to no longer having a budget, or vice versa.

Columns 5 and 6 of Table 3 then calculate treatment impacts on outcomes of improving that practice between 2 and 6 months, and worsening that practice between 2 and 6 months (the residual category being staying the same). We see significant negative impacts

of treatment on improving business practices, reflecting control group catch-up, and significant positive impacts on worsening practices, reflecting treatment group falling back. For example, the control group is 6.4 percentage points more likely to have started keeping a budget between 2 and 6 months, while the treatment group is 9.0 percentage points more likely to have stopped keeping a budget. Together these two estimates account for the 15.4 percentage point difference between the 2-month and 6-month ITT estimates for having a written budget. The last column then calculates the proportion of this change in treatment effect that comes from control group catch-up as opposed to the treatment group falling back. For example, for having a written budget this is $6.4/15.4 = 0.42$.

This then raises two questions: why did the treatment group stop doing some of these practices, and how was the control group able to adopt some of them without receiving training? We speculate that one reason may be related to the churn in business practices that we observe, which may reflect both the types of individuals who applied to the program and changes in the economy taking place as Mexico came out of the COVID-19 pandemic. Women who responded to advertisements about a business training program may be people who are looking to make changes in their business. In addition, changes in

Table 3
Improvement and worsening of business practices.

Dependent variable	Baseline mean	Outcome ITT		Churn 2-6-month ITT			
	Control	2 months	6 months	Control mean churn	Improve	Worsen	Frac. catch up
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Index of planning practices	0.164	0.122 [0.022]***	−0.002 [0.023]	0.574	−0.094 [0.029]***	0.117 [0.028]***	0.446
Has written budget	0.194	0.157 [0.030]***	0.003 [0.030]	0.322	−0.064 [0.024]***	0.090 [0.022]***	0.418
Has set sales goals for next year	0.213	0.080 [0.029]***	−0.033 [0.030]	0.322	−0.056 [0.023]**	0.057 [0.022]**	0.494
Has budget of approximate costs	0.085	0.127 [0.029]***	0.023 [0.030]	0.358	−0.065 [0.024]***	0.039 [0.022]*	0.622
Index of accounting practices	0.273	0.069 [0.018]***	0.021 [0.019]	0.759	−0.069 [0.031]**	0.093 [0.030]***	0.427
Keeps written records	0.499	0.090 [0.024]***	0.036 [0.024]	0.208	−0.027 [0.019]	0.028 [0.019]	0.491
Records every purchase and sale	0.492	0.095 [0.026]***	0.023 [0.026]	0.231	−0.027 [0.021]	0.046 [0.019]**	0.366
Records how much money business has	0.342	0.046 [0.027]*	0.053 [0.027]**	0.226	0.027 [0.022]	0.019 [0.020]	3.62
Records sales trends	0.224	0.074 [0.029]**	0.020 [0.029]	0.299	−0.011 [0.024]	0.043 [0.021]**	0.204
Calculates sales and expenses	0.345	0.073 [0.028]***	0.026 [0.028]	0.322	−0.025 [0.023]	0.022 [0.022]	0.535
Knows most profitable products	0.380	0.039 [0.026]	0.038 [0.025]	0.246	0.002 [0.021]	0.002 [0.020]	−3.85
Has records showing could pay off loan	0.104	0.053 [0.030]*	−0.004 [0.030]	0.332	−0.023 [0.024]	0.034 [0.023]	0.404
Has documents of annual profits	0.045	0.045 [0.023]*	−0.009 [0.027]	0.261	−0.013 [0.024]	0.040 [0.018]**	0.249
Tracks cash income annually	0.030	0.051 [0.028]*	0.035 [0.030]	0.272	0.026 [0.024]	0.042 [0.020]**	−1.62
Separates household and personal finances	0.268	0.110 [0.028]***	−0.019 [0.031]	0.315	−0.047 [0.020]**	0.082 [0.026]***	0.364

Notes: Column 1 shows baseline means of the indices of planning practices and accounting practices, along with the individual practices that are included in these indices. Columns 2 and 3 show ITT treatment impacts from regressions which include randomization strata fixed effects and control variables selected via pdslasso. The 2 to 6 month churn in column 4 is the proportion of control firms that change the practice between the 2 and 6 month surveys. Column 5 shows the estimated treatment effect on improving (starting) the practice between 2 and 6 months, and Column 6 on worsening (dropping) the practice during this time frame. Column 7 shows the fraction of the change in treatment effect between 2 and 6 months which comes from the control group catching up (being more likely to improve). Robust standard errors in parentheses.

* Denote significance at the 10 percent level.
** Denote significance at the 5 percent level.
*** Denote significance at the 1 percent level.

the economy during the pandemic may have made business owners also seek to do something new. This could explain why the control group adopts new practices over time. However, some of these planning and accounting practices may be ones where firms do not do them all the time, but rather, at some point, decide they need to take a snapshot of what is happening in their business. Training may then have just accelerated this process of trying some of these new practices. In the qualitative surveys, some owners in the treatment group acknowledged that they had had the discipline to implement new practices while training was taking place, but then spoke of losing their 'discipline' and reverting back to their old ways once training was over.

Another reason that firm owners may stop doing practices is that they are unable to detect changes in business performance from using these practices, given the amount of volatility and other factors driving sales and profits. Our analysis finds that even with a sample of over 2000 firms, it is difficult to statistically detect an impact of a 10–13 percent increase on sales and profits. In our qualitative interviews, three out of 10 of the sample we had identified as showing the highest sale growth in the quantitative survey said they believed the program had not increased sales. Moreover, since firms did not increase their use of marketing practices or personal initiative, there could be limited impacts on generating new customers and prolonged sales impacts. In other business training studies in which firms have improved marketing and customer retention, and introduced new products, it has been easier for them to detect new sales coming from these efforts as customers mention a promotion or buy a new product (e.g. McKenzie and Puerto (2021)).

4. How might online training be done better?

Our results demonstrate the feasibility of conducting online training by Zoom, but show the need for improvement in cost-effectiveness and impact. Based on our observations of the training and qualitative interviews with participants, there appear to be several areas where improvement in content and delivery could occur in any future efforts.

In terms of content, we saw that the training did not result in improvements in marketing practices. Nor are firms innovating by digitization or selling new products or services (Table C.1). Hardy and Kagy (2020) have shown how a lack of demand is a key constraint for the growth of many women-owned businesses. Increasing demand through product innovation and better marketing efforts to generate new sales could help the program have greater impact. Revamping the marketing component to provide specific actionable steps suitable for these types of firms could help result in more impact.

In addition to providing less general and strategic content, and more actionable steps, training could also attempt to foster more networking and sharing of knowledge among those participating. There was some interactive participation via audio and chat features in Zoom, and the instructors also set up Whatsapp groups to communicate with the class. But these did not result in much networking or idea sharing between participants (Table C1), instead serving mainly as a means to communicate with the trainer. Incorporating a more structured networking component via the Whatsapp groups, as in Asiedu et al. (2023) could help enhance effectiveness.

Finally, since the main cost is the instructor salary, the main suggestion to improve cost-effectiveness would be to increase the size of online classes. Doubling the typical class from 15–20 students to 30–40 students would likely still allow as much, or even more, interaction, while almost halving the cost. Having the most dynamic and effective trainers train larger groups of women at once offers the potential to greatly improve cost-effectiveness and hence scalability.

5. Conclusions

Women running small businesses throughout Mexico and in Guatemala were able to successfully connect to, and attend, business training sessions by Zoom. Widespread usage of mobile phones and social media has now made using technology to offer programs to thousands of microenterprises possible. The logistics of recruiting microenterprises and scheduling resulted in class sizes that were similar to in-person training, so that personnel costs were not lowered much by holding training online, resulting in relatively limited cost savings, albeit with much greater geographic spread. Future endeavors for implementation at scale needs to include larger class sizes with the best trainers to drive down costs. We found microenterprise owners did implement some of the practices learned, resulting in short-run gains, but they dropped some of these practices and the control group had caught up by 6 months. While there is a tendency to want to spend time on diagnostics and strategic planning, making sure training includes immediately actionable and specific advice that entrepreneurs can use to ensure their business looks different tomorrow than it does today is needed to improve the effectiveness of training content.

CRediT authorship contribution statement

Elwyn Davies: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Software, Writing – original draft. **Peter Deffebach:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Leonardo Iacovone:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft. **David McKenzie:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data and replication code linked in cover page.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jdeveco.2023.103244.

References

Anderson, S., Chintagunta, P., Vilcassim, N., 2022. Virtual collaboration technology and international business coaching: Examining the impact on marketing strategies and sales. Mark. Sci. forthcoming.

Anderson, S., McKenzie, D., 2022. Improving business practices and the boundary of the entrepreneur: A randomized experiment comparing training, consulting, insourcing and outsourcing. J. Polit. Econ. 130, 157–209.

Asiedu, E., Lambon-Quayefio, M., Truffa, F., Wong, A., 2023. Female Entrepreneurship and Professional Networks. Mimeo.

Behaghel, L., Crépon, B., Gurgand, M., Barbanchon, T.L., 2015. Please call again: Correcting nonresponse bias in treatment effect models. Rev. Econ. Stat. 97, 1070–1080.

Belloni, A., Chernozhukov, V., Hansen, C., 2014. Inference on treatment effects after selection among high-dimensional controls. Rev. Econom. Stud. 81, 608–650.

Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Patnaik, M., Saporta-Eksten, I., Van Reenen, J., 2019. What drives differences in management practices? Amer. Econ. Rev. 109, 1648–1683.

Broderick, T., Giordano, R., Meager, R., 2023. An Automatic Finite-Sample Robustness Metric: When Can Dropping a Little Data Make a Big Difference?. Working Paper.

Campos, F., Frese, M., Goldstein, M., Iacovone, L., Johnson, H., McKenzie, D., Mensmann, M., 2017. Teaching personal initiative beats traditional training in boosting small business in West Africa. Science 357, 1287–1290.

Carpena, F., Cole, S., Shapiro, J., Zia, B., 2019. The ABCs of financial education: Experimental evidence on attitudes, behavior, and cognitive biases. Manage. Sci. 65, 346–369.

Chernozhukov, V., Demirer, M., Duflo, E., Fernández-Val, I., 2020. Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments with an Application to Immunization in India. Working Paper.

Cusolito, A., Darova, O., McKenzie, D., 2023. Capacity building as a route to Export market expansion: a six-country experiment in the Western Balkans. J. Int. Econ. forthcoming.

Estefan, A., Improta, M., Ordoñez, R., Winters, P., 2023. Digital Training for Micro-Entrepreneurs: Experimental Evidence from Guatemala. Mimeo.

Fox, L., Thomas, A., 2016. Africa's got work to do: A diagnostic of youth employment challenges in Sub-Saharan Africa. J. Afr. Econ. 25, i16–i36.

Hardy, M., Kagy, G., 2020. It's getting crowded in here: Experimental evidence of demand constraints in the gender profit gap. Econom. J. 130, 2272–2290.

Iacovone, L., Calderón, G., MacGregor, C., 2018. Participating or not? Characteristics of female entrepreneurs participating in and completing an entrepreneurial training program. In: AEA Papers and Proceedings, Vol. 108. pp. 246–251.

INEGI, 2016. Las Empresas En México: Censos Económicos 2014. Instituto Nacional de Estadística y Geografía.

Jin, Y., Sun, Z., 2021. Lifting Growth Barriers for New Firms Evidence from an Entrepreneurship Training Experiment with Two Million Online Businesses. Mimeo.

McKenzie, D., 2021. Small business training to improve management practices in developing countries: Re-assessing the evidence for "training doesn't work". Oxf. Rev. Econ. Policy 37, 276–301.

McKenzie, D., Puerto, S., 2021. Growing markets through business training for female entrepreneurs: A market-level randomized experiment in Kenya. Am. Econom. J.: Appl. Econom. 13, 297–332.

McKenzie, D., Woodruff, C., 2014. What are we learning from business training evaluations around the developing world? In: World Bank Research Observer, Vol. 29. pp. 48–82.

McKenzie, D., Woodruff, C., 2017. Business practices in small firms in developing countries. Manage. Sci. 63, 2967–2981.

McKenzie, D., Woodruff, C., Bjorvatn, K., Bruhn, M., Cai, J., González-Uribe, J., Quinn, S., Sonobe, T., Valdivia, M., 2021. Training Entrepreneurs. In: VoxDevLit, vol. 1, (2).

Reich, J., Ruipérez-Valiente, J.A., 2019. The MOOC pivot. Science 363, 130–131.

Rivard, R., 2013. Measuring the MOOC Dropout Rate. Inside Higher Ed, March 8.

Sheridan, M.B., 2021. Mexico's Pandemic Policy: No Police. No Curfews. No Fines. No Regrets. Washington Post.

Van Lieshout, S., Mehtha, P., 2017. The Next 15 Million: Start and Improve Your Business Global Tracer Study 2011–2015. International Labour Organization.

Further reading

Bryan, G., Karlan, D., Osman, A., 2021. Big Loans to Small Businesses: Predicting Winners and Losers in an Entrepreneurial Lending Experiment. Working Paper.

Ghanem, D., Hirschleifer, S., Ortiz-Becerra, K., 2022. Testing Attrition Bias in Field Experiments. CEGA Working Paper no. 113.