Journal of International Economics 144 (2023) 103794

Contents lists available at ScienceDirect

Journal of International Economics

journal homepage: www.elsevier.com/locate/jie

Full Length Articles

Capacity building as a route to export market expansion: A six-country experiment in the Western Balkans☆

Ana P. Cusolito [a], Ornella Darova [b], David McKenzie [c,*]

[a] Markets and Technology Unit, The World Bank, Washington D.C. 20433, United States
[b] Department of Economics, University of Pennsylvania, 133 S 36th St, Philadelphia, PA 19104, United States
[c] Development Research Group, The World Bank, Washington D.C. 20433, United States

ARTICLE INFO

Article history:
Received 11 December 2022
Received in revised form 29 June 2023
Accepted 30 June 2023
Available online 07 July 2023

Repository link: https://data.mendeley.com/datasets/xkrpv4khnj/1

JEL:
F14
O25
O12
L26
C93

Keywords:
Export market expansion
Training
Consulting
Digital presence
Customer acquisition

ABSTRACT

The limited market size of many small emerging economies is a key constraint to the growth of innovative small and medium enterprises. Exporting offers a potential solution, but firms may struggle to locate and appeal to foreign buyers. We conducted a six-country randomized experiment with 225 firms in the Western Balkans to test the effectiveness of 30 h of live group-based training and 5 h of one-on-one remote consulting in overcoming these constraints. Treated firms used techniques such as search engine optimization and improved Facebook content to increase their digital presence and better reach foreign customers. A year later, we find positive and significant impacts on the number of customers, and a significant intensive margin increase in export sales. Qualitative interviews suggest this improvement came from a combination of sector-specific advice on market expansion, and through an encouragement effect which gave entrepreneurs the confidence to try new sales strategies.

1. Introduction

With a low cost of labor, skilled human capital, and geographical and cultural proximity to Europe, countries in the Western Balkans would seem to have the potential for growth in innovative entrepreneurship. However, a key constraint to the growth of

☆ This study was funded by the European Union under the Western Balkans Enterprise Development and Innovation Facility. We thank Maja Andjelkovic and Alberto Criscuolo from the World Bank for their support and collaboration in implementing this project; Muhamed Semic, Elma Delalic and the Deloitte team for their work implementing the project and responses to our many questions; and Vladimir Panic for excellent research assistance. We also thank comments received from the editor, two anonymous referees, Ana Fernandes, Natasha Kapil, and participants at several seminars. This study was registered in the AEA RCT Registry on May 21, 2021, and the digital object identifier (DOI) is: https://doi.org/10.1257/rct.7711-1.0

* Corresponding author.
E-mail address: dmckenzie@worldbank.org (D. McKenzie).

https://doi.org/10.1016/j.jinteco.2023.103794
0022-1996/

small and medium enterprises (SMEs) with new ideas is limited domestic demand, with populations varying from only 620,000 in Montenegro up to slightly under 7 million in Serbia. Limited demand then reduces the amount of market traction[1] these firms have to show in their early stages, acting as a barrier to their investment readiness when seeking external funding. Expanding market-size by exporting to larger and richer countries offers the potential to vastly expand the customer base for specialized products, and boost individual firm and overall economic growth (Frankel and Romer (1999); Goldberg and Reed (2021)). However, expanding into foreign markets can involve substantial search frictions (Arkolakis, 2008; Anderson and van Wincoop, 2004; Allen, 2014), requiring SMEs to learn the tastes and preferences of foreign customers, and figure out how to market their products to these buyers. Small firms may lack the knowledge, skills, and confidence to try to do this, limiting how much they can export.

We conducted a six-country randomized experiment in Albania, Bosnia and Herzegovina, Kosovo, Montenegro, North Macedonia and Serbia to test the effectiveness of a capacity building program (called *PowerUP*) in overcoming these constraints and helping innovative firms to expand their markets and export more. Our sample consists of 225 SMEs with a mean (median) of 17.6 (8) workers. The majority of the firms are in the service sector and have other businesses as their main customers, such as firms producing customized software solutions and providing information technology solutions for other businesses, and professional services such as legal and energy consulting. One quarter of the sample produces specialized manufacturing products. The majority were doing some exporting already, with their main challenge being how to acquire new customers and expand the amount of exporting they were doing. This was additionally important as firms sought to bounce back from a short-term fall in business due to the onset of the COVID-19 pandemic. These firms were randomized into 113 treated firms and 112 control firms. The treatment group was offered a combination of 30 h of virtual group-based training sessions over Zoom, as well as up to 5 one-on-one meetings with a Deloitte consultant. Training sessions focused heavily on building customers and markets, and using digital marketing tools, with additional sessions on human resources, financial management, and pitching. The control group was offered three webinars, lasting a total of 4 h, which provided a conceptual overview of these topics without providing advice on specific actions to take. Firms in both groups were then invited to submit a pitch deck, with the top eleven pitching in a regional event.

Take-up of the training and one-on-one sessions was initially high for an online program, with 97% of the treatment group attending some sessions. However, attendance fell over time, with the median firm attending 6 out of the 9 days of training. Given the structure of the course, this meant most firms received training on customers and markets, and on using digital tools to reach customers, but fewer were trained in other areas. We find three main results. First, treated firms improved their digital presence in a way that made it easier for potential international customers to find the firm's offerings and be attracted to contact the firm. We independently searched for the firms on Google, and scored their website and Facebook pages on multiple dimensions. Treated firms were 10 percentage points more likely to show up on the first page of search results, and improved their Facebook pages by being 16 percentage points more likely to include posts in English, 10 percentage points more likely to use special offers, and 11 percentage points more likely to include customer testimonials and stories. Second, a year after starting the program, treated firms had significantly increased the number of customers, with a statistically insignificant 12% increase in total sales revenues. Third, the distribution of export sales shifted right, with no significant change at the extensive margin, and a significant increase in export sales at the intensive margin. In contrast, we see no improvements in more general business practices, nor in the quality of their pitches. In-depth qualitative interviews were used to examine how firms had expanded their digital presence and export sales. The improvement in digital presence appears to be a knowledge story, with many firms not being aware of how to use modern digital marketing methods. The improvement in exports is suggested to come through a combination of two channels. The first is sector- and market-specific advice offered by the consultants on how to expand sales in Europe by targeting the right customers and adjusting product offerings to meet customer demand. The second is that the program appeared to encourage firms who were unsure about expanding that they were on the right track and should have the confidence to try new markets. The results point to the potential of capacity-building efforts to help in market expansion.

This paper contributes to three related strands of the literature. The first concerns direct policy efforts to help firms to export. Srhoj et al. (2020) provide a recent survey of the literature on the wide range of programs undertaken by export promotion agencies to attempt to boost exports, which range from grants, credit, guarantee schemes, technical assistance with quality standards, trade fairs, and a range of other activities. Most of this evidence is non-experimental and focused on the manufacturing sector, attempting to match participating firms to similar non-participants (e.g. Görg et al. (2008), Martincus and Carballo (2008), Munch and Schaur (2018), and Cadot et al. (2015)). Overall the evidence has found positive impacts on the extensive margin of trade, with more mixed evidence on the intensive margin. There have been fewer experimental studies of efforts to boost exports. Atkin et al. (2017) work with very small (average one worker) Egyptian rug manufacturers who were not exporting, and directly provide them with initial export orders in order to measure the impact of exporting on productivity. Two experiments that attempted to reduce informational frictions through light-touch interventions (information brochures (Breinlich et al., 2017) or a one-day seminar (Kim et al., 2018)) found no impact on the extensive margin. Iacovone et al. (2023) provided more intensive consulting services to Colombian firms that focused on improving management practices and productivity as an indirect route to increasing exporting, finding no effect. Our paper builds on this literature by directly teaching firms to find and connect with foreign customers, and boosting exports at the intensive margin.

[1] Investors commonly ask for proof that customers are interested in the product or service a firm is selling, and a demonstration of the potential size of this market. In addition to actual sales, this could include other demonstrations of customer interest and customer acquisition such as website traffic, social media likes, customers newsletter or email subscriptions, downloads, and interest from customers in new markets. Greenstein (2011) notes a simple definition of market traction is "quantitative evidence of market demand".

Second, our paper contributes to a literature showing the potential of digital technologies to lower search costs, reduce trade frictions, and expand markets (Bakos, 2001; Goldfarb and Tucker, 2019). Most of the micro-empirical literature has focused on how this facilitates internal market integration, such as the use of mobile telephony to boost internal trade in Uganda (Bergquist et al., 2022), and the use of e-commerce platforms to boost internal trade in China (Jin and Sun, 2021; Couture et al., 2021). A recent exception is work by Carballo et al. (2022) who conduct a non-experimental evaluation of the impact on Peruvian firms of joining an online business platform. They find this platform lowers search costs and boosts export values. Our work compliments these studies and highlights the use of digital marketing tools like search engine optimization, and social media improvements to help connect with foreign customers and boost international trade.

Finally, our work contributes to a broader literature on firm upgrading (Verhoogen, 2022), and specifically the effectiveness of training and consulting capacity-building programs for firms. McKenzie et al. (2021) provide an overview, noting that the majority of these studies have either involved training for micro-firms, or intensive management consulting for larger manufacturing firms, without a focus on exporting. There have been relatively few studies with the types of innovative and service-oriented SMEs that are studied here. In a predecessor to this project (Cusolito et al., 2021), we worked with a different sample of earlier-stage innovative start-ups in five Western Balkans countries on an investment readiness program.[2] That program focused much more on getting entrepreneurs more interested in equity investment, improving their pitching skills, and helping them at an earlier stage of product development. One of the key lessons from that work was that many of these firms lacked sufficient market traction to be attractive to investors, and that domestic markets were typically too small for this sales growth. This project builds on those lessons by focusing on efforts to help firms boost markets. In addition, existing training and consulting programs have typically been conducted in person, raising questions as to whether they can be effectively delivered remotely. This is particularly important when attempting to run programs with specialized firms across multiple cities or countries, where travel to in-person training events can constrain participation. Our work shows the potential, and some of the limits, of providing training and short consulting visits remotely.

2. Program details and experimental design

This program was funded by the European Union through the Western Balkans Enterprise Development and Innovation Facility (EDIF). One of the goals of this program is to help shift the Western Balkans region towards a productivity-based and export-oriented growth model, through supporting the growth of innovative firms. In late 2019, a competitive procurement process was used to select an implementing partner that could build the capacity of innovative firms in the region. The international consulting firm Deloitte was selected as the winner, leading the implementation of the program through its Sarajevo office. The original proposal was to involve in-person workshops and meet-ups, and a focus on building market traction to help firms become more attractive to investors. The onset of the COVID-19 pandemic required a shift in both the modality (from in-person to online) and in the main priority areas for training. The training was re-designed to help firms adapt and bounce back from the COVID-19 pandemic through a focus on helping them to acquire new customers and expand into new markets given limited local demand. The pandemic also led to an increased focus on digital marketing and reaching customers through an enhanced digital presence.

2.1. Program launch, randomization, and sample

The program was launched under the name *PowerUP*, on September 2, 2020 (see appendix Appendix A for a timeline). The call was issued through Deloitte and the World Bank's social media channels, through Facebook and LinkedIn, local media, and through partnering organizations such as entrepreneurship hubs and Chambers of Commerce in each country. To be eligible for participation, firms had to be legally incorporated and registered in one of the six countries (Albania, Bosnia and Herzegovina, Kosovo, Montenegro, North Macedonia, and Serbia); operating for at least one year; have between 2 and 100 employees; be innovative and have a product or service that is already generating revenue in at least one market and be looking to boost sales; and not be conducting activities on a "negative list" such as arms, alcohol or tobacco. Firms had to apply online and fill out an application form. The program was offered free of charge, but companies had to sign a Memorandum of Understanding to commit their time to take part in the project.

A total of 412 firms applied for the program by the deadline of October 4, 2020, of which 250 firms were selected based on the eligibility criteria (the main reasons for rejection were being too small or not innovative enough). These firms were then asked to fill out a diagnostic questionnaire that would provide more details on their current market traction and sales, their business practices, and their investment readiness. The 225 firms that answered this diagnostic then provide the experimental sample used for this study. They were stratified based on country of operation (Bosnia and Herzegovina, Serbia, or other); whether or not they sold a digital product; and into three groups based on a diagnostic score assessed by Deloitte that measured their baseline levels

[2] The firms in that earlier program were on average younger, some were still in the pre-revenue stage, and they were smaller in size (a mean(median) of 6(4) workers compared to 17(8) workers in this study.

Table 1
Summary Statistics on Experimental Sample.

	(1)					(2)	(3)	(4)
	Full sample					Control	Treatment	*P*-value
	Mean	St. Dev.	10th perc.	Median	90th perc.	Mean	Mean	
Firm registered in Serbia	0.23	(0.42)	0.00	0.00	1.00	0.23	0.23	.
Firm registered in Bosnia and Herzegovina	0.31	(0.46)	0.00	0.00	1.00	0.32	0.30	0.13
Sector is ICT	0.36	(0.48)	0.00	0.00	1.00	0.36	0.35	0.97
Diagnostic score for business practices	2.01	(0.82)	1.00	2.00	3.00	2.02	2.01	1.00
Diagnostic score for investment readiness	0.61	(0.15)	0.40	0.62	0.80	0.61	0.61	0.50
English proficiency of 4 or 5 out of 5	0.86	(0.35)	0.00	1.00	1.00	0.86	0.87	0.80
Female (co-)founder(s)	0.28	(0.45)	0.00	0.00	1.00	0.30	0.26	0.70
Average age of (co-)founder(s)	43.13	(11.19)	31.50	41.00	59.50	42.78	43.46	0.81
At least one co-founder with master degree	0.47	(0.50)	0.00	0.00	1.00	0.40	0.55	0.06*
Has a digital product	0.19	(0.39)	0.00	0.00	1.00	0.21	0.18	0.61
Age of firm in 2020	9.54	(9.34)	1.49	6.04	24.80	8.81	10.26	0.32
Sector is Manufacturing	0.24	(0.43)	0.00	0.00	1.00	0.24	0.23	0.81
Sector is Professional Services	0.14	(0.35)	0.00	0.00	1.00	0.13	0.14	0.74
Firm is B2B	0.72	(0.45)	0.00	1.00	1.00	0.70	0.73	0.53
Total number of employees	17.58	(21.39)	3.00	8.00	48.00	17.05	18.10	0.71
2 to 5 Full-time employees	0.40	(0.49)	0.00	0.00	1.00	0.42	0.39	0.73
6 to 10 Full-time employees	0.16	(0.37)	0.00	0.00	1.00	0.13	0.19	0.28
Outside investor (non-family & friends)	0.08	(0.28)	0.00	0.00	0.00	0.10	0.07	0.43
Interested in outside investment	0.84	(0.36)	0.00	1.00	1.00	0.86	0.83	0.52
Have used consulting before	0.47	(0.50)	0.00	0.00	1.00	0.46	0.48	0.85
Company has a Facebook page	0.75	(0.43)	0.00	1.00	1.00	0.78	0.73	0.36
Uses Search Engine Optimization (SEO)	0.25	(0.44)	0.00	0.00	1.00	0.29	0.22	0.27
Has updated comprehensive business plan	0.55	(0.50)	0.00	1.00	1.00	0.51	0.59	0.23
Uses KPIs to track marketing	0.27	(0.44)	0.00	0.00	1.00	0.29	0.25	0.59
Adapts products to export markets	0.62	(0.49)	0.00	1.00	1.00	0.59	0.65	0.39
Firm derives revenue from exports	0.71	(0.45)	0.00	1.00	1.00	0.73	0.69	0.57
Has sales in Europe	0.46	(0.50)	0.00	0.00	1.00	0.46	0.47	0.79
Profit in 2019	83,964	(251321)	−873	15,000	298,261	84,997	82,941	0.95
Profit in first 9 months of 2020	101,334	(370978)	−15,031	14,100	338,827	126,284	76,615	0.32
Customers in Jan 2020	878	(8637)	1.00	12.00	166	1205	553	0.59
Revenue in 2019 below 25,000 Euros	0.16	(0.37)	0.00	0.00	1.00	0.16	0.16	0.98
Revenue in 2019 25,000–100,000 Euros	0.28	(0.45)	0.00	0.00	1.00	0.33	0.24	0.16
Revenue in 2019 100,000–250,000 Euros	0.13	(0.34)	0.00	0.00	1.00	0.12	0.15	0.45
Observations	225					112	113	225

Notes: Baseline characteristics of firms involved in the program. Our baseline data come from the application form and diagnostic sheets submitted by all firms as they entered the program. The omnibus test of joint orthogonality, dummying out missing values, has p-value 0.165. When restricted to set of 141 observations with all non-missing covariates, the p-value is 0.516.

of business practices and investment readiness. Half of the firms were then randomized to treatment and half to control within each stratum,[3] resulting in 113 firms assigned to treatment and 112 firms assigned to control.

We collected baseline data from the application form and a diagnostic questionnaire carried out in October and November 2020. The baseline information offers key insights on the characteristics of the firms included in the sample, with summary statistics provided in Table 1. The selected firms are heterogeneous in sector. Just over one-third (36%) of firms are in the ICT sector, making customized software or providing information technology solutions for other businesses and consumers. Most of these firms are in customized software development. Examples include companies that develop mobile and web applications, enterprise management software, cloud services, and other software consulting for other firms; supply-chain management; and helping firms to use new technologies such as virtual and augmented reality and 3D printing.[4] The second most common industry is manufacturing (24%), which covers a wide range of activities including making specialized metal products, specialized clothing (e.g. safety gear for factories, medical uniforms, fashion items), food products (e.g. traditional food shipped online, organic

[3] The three categories provided (location, digital product, and diagnostic score group) formed 18 strata. There were 214 firms who had answered the diagnostic questionnaire by December 2, and they were randomly assigned within these strata. A further 11 firms that were late in responding were then separately randomized with equal probability to treatment and control within a 19th stratum consisting of 10 of these stragglers, and then the final firm was randomly assigned to treatment as a singleton. In practice this means the final firm will not contribute to identifying the treatment effects when we include controls for randomization strata. This firm only attended one session and then attrited, not providing any follow-up data apart from our independently measured digital presence variables.

[4] These firms were not simply selling mass-market software to consumers through an app store, but rather developing customized tools for that often required considerable interaction with customers. For example, a firm selling digital tools for vineyard management would need to visit customers (wine producers) in person to help determine the system needed for them and to set it up.

processed food products) and wooden products (e.g. handmade window frames and shutters, boats). Professional services accounts for 14% of firms, and includes digital marketing firms, legal consulting, health services, and energy consulting. Four to 5 % of firms are each in retail, tourism, and education. 72% of firms are largely B2B firms, with firms in the ICT sector (94%) and professional services sector (90%) being particularly focused towards other businesses as customers, compared to firms in tourism (100%), education (100%), and retail (67%) being more focused on consumers as customers.

Specific examples of the types of innovative products and services these firms are offering include: a vibration-based system for preventative maintenance of machines, an online notary service with tools for signature verification, production of biodegradable packaging, cloud-based software for the energy sector, sensors to help detect when manhole covers are removed designed for telecom firms, solar-powered smart benches for bus stops, using biometric technology and artificial intelligence to improve marketing campaigns, and digital tools for vineyard management.

The firms are small and medium-sized enterprises, with a mean (median) size of 17.6 (8) workers, and median annual sales between 100,000 and 250,000 euros. Their founders have a median age of 41, are typically well-educated (with 47% of firms having a founder with a Masters degree), and 86% speak English well. Most firms are already engaged in exporting, with 71% of the firms having export sales at the time of application, and 46% making sales to Europe. The main export challenge for most firms is thus expanding exports at the intensive margin, not starting to export. Likewise, most firms have some online presence, with a webpage and 75% having a business Facebook page. However, the majority have considerable scope to improve this digital presence. For example, only 25% are using search engine optimization (SEO), and many of the Facebook and webpages lack features that might help better attract customers such as customer stories and testimonials and calls to action for purchases, as well as content in English or other foreign languages that can help attract foreign customers. This lack of digital marketing practices was even true among firms in the ICT sector, with the majority being run by owners with scientific and engineering backgrounds who could solve tough technical problems, but who had less experience with marketing.

2.2. Details of the treatment and control offerings

The treatment and control groups were blinded to treatment status, with both being told that they had been selected for the PowerUP program, that they would receive some online training, and that there would be a chance to prepare a pitch deck and get selected for a chance to pitch to investors at the end. This was done to provide some form of opportunity to all participants, and to attempt to reduce the risk of the control group attriting and not providing any follow-up data. The difference between the two groups then lies in the intensity of the capacity-building support provided, and the amount of actionable advice that accompanied it.

The treatment group was offered a combination of virtual group-based training sessions by Zoom, as well as a series of virtual one-on-one meetings with a Deloitte consultant. These training sessions took place one or two days per month between December 2020 and May 2021, with homework set after each one, and the one-on-one follow-up sessions interspersed. In total, they were offered just over 30 h of group-training sessions, and 5 h of one-on-one consulting. The group-training was provided in English, with one standalone booster session offered for Albanian speakers due to lower English proficiency for some of this group. The trainers were subject matter experts from Deloitte, all of whom were experienced working with companies in the region. The training sessions involved lectures, with the use of Zoom features like polls, raise hand, and chat to enable some interactivity and the ability for firms to ask questions and share experiences.

The group-based training sessions were organized around one topic per month, with three of the six months focusing on topics related to customers, markets, and market expansion, including the use of digital tools to expand the number of customers. This began with an overview of the impacts of the COVID-19 pandemic's effects, and recommended firms adopt an agile approach to revamp firm demand by strengthening their marketing strategy to enlarge and consolidate the domestic and foreign customer base. Firms were given advice on strategies for evaluating and prioritizing growth opportunities abroad, including how to identify potential sales channels, find partners in new markets, and segment the foreign customer base to focus on high-margin customers.

Given the pandemic, firms were advised to focus more on their online selling capabilities, including improving their website, using search engine optimization (SEO) tools to drive traffic, and using data to track customer experiences and satisfaction. Instructors also discussed all the digital touchpoints along the lifecycle of the relationship with a customer (e.g., awareness, consideration, purchase, experience, advocacy, and loyalty) and specific actions participants can implement to leverage digital solutions. The training discussed how to use customer testimonials and feedback to diffuse information about their products, and specific strategies that could be used to strengthen relationships with customers. Homework involved experimenting with at least one new digital touchpoint, requesting customer feedback, analyzing it, and putting in practice SEO tips.

The other three months covered three additional topics. The first was financial management and profitability. This included using financial statements, doing cash flow analysis, and making forecasts of income and expenses upcoming. One component of this dealt with financing trade, with specific tips on how to improve trade payables management. The second topic was human resources management. Given the median firm had 8 workers, part of this discussed how to do HR management without a HR manager, with a focus on how to recruit and retain talent. The second day of this was heavily focused on sales, including how to incentivize and compensate sales staff, and tips for recruiting sales agents abroad to help with selling to foreign customers. The final topic covered was pitching skills, and how to prepare pitches of different types for investors.

One of the downsides of a fully online offering was fewer chances for the informal networking with other firms that can take place during in-person training. Firms could learn a little about the experiences of other firms in entering new markets through

some of the interactive discussion during training, and there were two short (15 min) zoom room discussions with other firms. Coupled with the firms being geographically dispersed and hesitation about in-person meet-ups during the pandemic, and we do not think networking with other firms was an important channel here. In our follow-up survey, only 4 firms report making a sale to another firm in the program. Table A6 shows a very low control mean for our measures of networking with other firms in the program and an insignificant treatment impact.

The one-on-one sessions offered a chance for firms to get firm-specific advice from consultants on applying these tools to their businesses. They were offered four sessions with Deloitte's consultants,[5] along with up to two additional sessions with international consultants provided by the World Bank. Firms would share the homework with the consultant and their focus questions for discussion before the meeting, and then receive tailored advice. For example, in our qualitative interviews, a Bosnian tourism firm noted how they had received suggestions for their digital marketing campaign and insights on their website and promotions. Two IT firms reported they had received advice around customer segmentation, with one noting that the consultant helped them to realize they were trying to do too many small margin activities that could not scale, with a suggestion to focus more on larger contracts. An Albanian firm that produced CRM software for the real estate sector received advice on how to enter other markets in Europe by using pre-existing relationships with large multinational companies as well as through using the diaspora. A Macedonian firm received specific advice on revamping its website in English, implementing SEO, and changing their social media posts so that they contain a call for action for users. After each session, the consultant would send the firm a summary and a list of specific actions to try out before the next meeting. E.g. within the next two weeks, obtain foreign customer testimonials and share them on your digital channels; within the next month devise the market penetration strategy with the Italian partner and which customer personas to target.

In contrast, the control group were just offered three webinar sessions of approximately one to 1.5 h each, in December 2020, February 2021, and May 2021, for a total of 4 h of training. One session was on customers and markets, one on finance and internal performance, and one on pitching. This content was at a conceptual or strategic level, with few specific actionable tips recommended, no homework, and no one-on-one sessions. All of the control content was also received by the treatment group as part of their training, but while the control training covered broad strategy and diagnostics, the treatment group would then receive advice on specific actions.

At the conclusion of training, firms in both treatment and control groups were provided with a pitch deck and a traction sheet (a simple sheet where they could provide some numbers on elements of market traction), which they were asked to submit to Deloitte by the end of May 2021. These were then scored by judges,[6] the top 20 proposals were then asked to submit video pitches, and the top 11 firms then selected to pitch live to a jury of regional venture capital investors at a finals event held on June, 29, 2021. The three best companies were selected to receive the Western Balkan Entrepreneurial Excellence Award, as well as receive additional, tailor-made, business advisory support by Deloitte.

2.3. Take-up

Fig. 1 shows the proportion of firms that attended each training session by treatment status. Initial take-up of training was high, with 93% of the treatment group and an estimated 94% of the control group attending the first session virtually. Most firms remained logged on for the full session, with 85% of the treated firms that attended staying for at least 6 h of the first day. Only 3 treated firms (2.7%) attended no sessions at all. However, we see attendance for both groups falling off over time, with attendance falling to only 42% of treated firms by the final session, and only 46% of control firms attending their third and final session. The mean (median) treated firm attended 5.5 (6) of the 9 sessions, while the mean (median) control firm attended 2.1(2) out of the 3 much shorter sessions for this group. Given that more of the sessions focused on customers and markets than other topics, and that these sessions came first, the result is that many more firms received training on customers and markets than other topics. The one-on-one consulting sessions for treated firms showed a similar pattern of initially high attendance, followed by drop-off: 87.6% attended the first meeting, 70.8% the second, 55.8% the third, and only 21.8% four or more meetings. 19.5% of treated firms used the additional consultants provided by the World Bank. A total of 136 firms (71 treatment, 65 control), submitted a pitch deck and traction sheet to be scored by judges and be eligible to be selected to pitch in the final competition. This constitutes 60.4% of the sample.

Three factors appear likely to account for this drop-off in training. The first is the difficulty of offering group-based training to a very heterogeneous mix of firms that differ in sector, size, life-cycle phase, sophistication, and nationality. We compared characteristics of the 64% of treated firms that attended five or more sessions versus those that attended fewer. Attendance rates varied substantially by nationality: 81% of the 26 Serbian firms in the treatment group attended five or more sessions, compared to 68% of the 34 Bosnian firms, 67% of the 24 Macedonian and Montenegrin firms, and only 41% of the 29 Albanian and Kosovar firms. One key reason for this is likely English proficiency, which was lower for the Albanian speakers, and led to a special booster

[5] Deloitte's consultants were all from the Balkans region. Table A5 shows the median consultant was 36, had a Masters' degree and 5 years of experience at Deloitte, and 56% said they had previously worked for a multinational or an exporter. These consultants are on average slightly younger and less experienced than the consultants, and similar to the outsourcing professionals, used in Nigeria in Anderson and McKenzie (2022). In Colombia, the consultants used in Iacovone et al. (2022) had an average age of 45, with 17 years experience and 74% having worked previously for a multinational or exporter. Deloitte's consultants are therefore not more experienced than is typical in government-run consulting programs.

[6] Each proposal was scored by three randomly assigned judges from a pool of 21 judges.

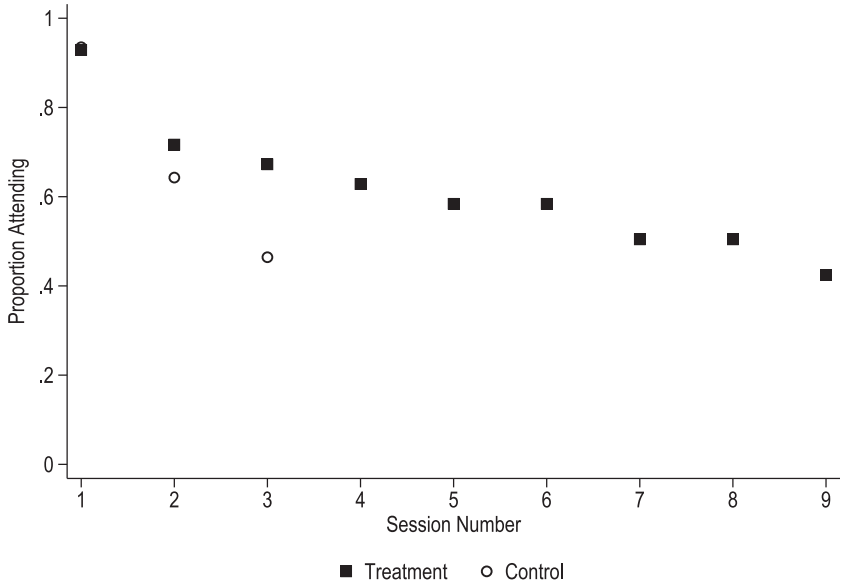

Fig. 1. Attendance rates per training session.
Notes: First session for Treatment was a full day, remaining sessions were half days. Session 8 and 9 cover attendance in either day of month 5 and month 6 trainings, since training was split by firm type on those days. Control group was only offered three short webinar sessions. Control group session 1 attendance data is estimate based on aggregate summary, since individual attendance data were erased before they could be downloaded.

session being offered for these firms. Firms were also less likely to attend the majority of sessions if they were in manufacturing compared to ICT, and if they had low revenues and profits at the time of application.

The second potential explanation for a drop-off in training attendance was firm fatigue and other issues associated with holding training sessions online. Holding sessions online was the only viable option during the COVID-19 pandemic, and indeed one reason entrepreneurs reported being unable to attend some sessions was because they were sick with COVID-19, or short-staffed because other employees were sick. It is unclear how much of a role holding training online played in determining attendance. On one hand, it can be difficult to pay attention to hours of video, online training tends to offer less chances for interaction with the instructors and other firms, and there can be a temptation to log off and address other business needs that arise. But this may be offset by two benefits of holding the sessions online. The first is that firms could have multiple people attending and watching the session. For example, in our qualitative interviews one entrepreneur noted how he used the sessions to bring in different people from his team to help introduce them to topics related to their role in the firm. The second is that for a six-country program of this sort, the counterfactual is likely to involve holding in-person training sessions that rotate among cities in the different countries. For example, in a precursor program in five Western Balkan countries, Cusolito et al. (2021) found that only 43.7% of firms attended at least one masterclass in-person, and then typically only the one held in their home country.

A final reason for the drop-off in training attendance is that it may have reflected some firms finding the training content less useful than others. Our own impressions from virtually attending the training sessions to observe were that they varied in the dynamism of the instructor, and in the extent to which they provided high-level strategic advice and broad overviews versus specific actionable steps. Some firms noted they either found some of the content too broad, not directly applicable to their business, or to be covering things they were already doing. However, even when content was somewhat familiar, several firms noted that it helped affirm that they were on the right track, or spurred them to get around to implementing practices they knew they should be doing but had not yet implemented. They also noted the one-on-one sessions had been helpful in addressing firm-specific issues. Firms showed most interest in improving customers and markets, and so were less likely to attend one-on-one sessions later in the program that were intended to focus on other business aspects such as financial performance. Only 9% of treated firms said the program had not been useful for their business, with 45% saying it was somewhat useful, and 46% very useful. Overall, 81% of treated firms said the program met or exceeded their expectations, and 90% would recommend it to other firms like theirs. The control group was less satisfied with the program, with only 27% saying it had been very useful to their firm, and 35% saying it did not meet their expectations.

3. Results

3.1. Data collection and attrition

Our baseline data come from the application form and diagnostic sheets submitted by all firms as they entered the program. Administrative data from the program provide the data on attendance used in our take-up analysis above. We then have three

sources of follow-up data to measure outcomes of participating in the program. The first are the scores from judges, based on the pitch decks and traction sheets submitted by firms at the end of May 2021. The purpose of these judge scores were to identify the firms that were most ready to represent the region in front of investors. They are only available for 136 out of the 225 firms (60.4%), and capture a measure of investment readiness of the firm seven months after the program started. The judges were a mix of investors and experts in small and medium enterprise growth in the region, and were blinded to treatment status. The judges gave firms a score out of 5 for seven different categories, with pre-assigned weights then used to arrive at a total score: unique value proposition (20%), technology (10%), business model (10%), team (10%), traction (30%), market expansion (10%) and finally pitch deck effectiveness (10%). The total scores of each of the three judges was then averaged to get a final score for each firm.

Our main follow-up data measure impacts approximately one year after the program started. We conducted a follow-up survey, which was answered by 180 firms (80%), with 87 control (77.7%) and 93 treated firms (82.3%) answering. This is supplemented by data from the traction sheets for an additional 14 firms, to give some follow-up data for 194 firms (86% of firms, 84% of control and 88% of treatment). Some of the firms only supplied partial follow-up data, resulting in item non-response on some of the financial measures, including our key outcome of export sales. However, Appendix Appendix B shows that we cannot reject equality of response rates by treatment status, and that the set of firms responding to the endline survey are balanced on baseline observables, as are those who supply export data.

This follow-up survey contains data on our main outcomes of interest for measuring whether the program enabled firms to gain more customers, increase revenues, and expand exports. We investigated whether administrative data sources could be used to provide more information on these outcomes. However, with firms from six different countries, it was not possible to link these firms to administrative tax data that could provide information on revenues. Export data, such as that included in the World Bank's Exporter Dynamics Database, is not currently available for these countries in our follow-up period, is provided in an anonymized format, and most importantly, comes from customs data that only captures goods crossing borders, and thus would miss the service exports that the majority of firms in our sample specialize in.[7] Hence survey data provide the only sales and export data available.

Our final source of follow-up data comes from a detailed scoring exercise we undertook on the digital presence of firms in December 2021, also one year after the program started (with a more basic version collected at baseline, and used as control variables). Improvement of digital presence to attract new customers was a key area emphasized by the program, and took on more importance during the COVID-19 pandemic as other opportunities to interact in-person with potential customers were limited. These data were collected by two independent consultants with linguistic knowledge of the prevalent languages in the region, who were blinded to treatment status. An additional advantage is that these data can be collected for all 225 firms in the experiment.

We measure three elements of digital presence, capturing the extent to which firms are employing modern marketing methods to be attractive to domestic and international customers. The first is based on Google searches, and reflects whether the firm has used search engine optimization techniques to make it easy for potential customers to find them. Our consultants score the firm on seven features of this search, including their page rank, whether there is a description in an English search, whether google provides a right menu address and contact information, and whether it is clear from the search result what the company does. The second measure examines the firm website, and scores in based on 18 features from whether the website is functioning, fast to load, available in multiple languages, and optimized for mobile viewing; to whether it includes marketing features such as customer testimonials, appealing aesthetics, a clear statement of what differentiates their product or service, and includes a clear call to action towards contacting the company or making a purchase. The final measure is based on the company's business Facebook page, and scores the company on 15 features including whether they actively post, provide clear descriptions of what the company does, include customer reviews, highlight new products or services, offer promotions, and include attractive images.

In Appendix Appendix B, we use these digital presence data to test whether the firms that did not answer the endline survey differ systematically in these follow-up outcomes from the firms that did. The attriting firms have similar Google, website, and Facebook scores, and, coupled with the balance on baseline observables for the responding sample, provide us with additional confidence that attrition is not likely to be causing systematic bias. In the next subsection we discuss two additional approaches we use to examine robustness to attrition.

3.2. Estimating treatment effects

To estimate the impact of the program on the outcomes of interest, we use the following regression for firm i in randomization stratum s:

$$Outcome_i = \alpha + \beta Treat_i + \sum_{s=1}^{S} \delta_s 1(i \in s) + \epsilon_i \quad (1)$$

[7] Indeed, the U.S. Bureau of Economic Analysis explicitly mentions that estimation of service trade for the U.S. is based on surveys, not administrative data.

This regression includes dummy variables for each of the randomization strata, implicitly controlling for the combination of baseline diagnostic score, country, and whether they have a digital product. We also include the baseline value of the outcome of interest for an Ancova specification for the outcomes for which baseline data were collected. Robust (Eicker-White) standard errors are used. The coefficient of interest β corresponds to the intention-to-treat effect (ITT), which is the effect of being offered the full intensive training and one-on-one consulting program, compared to just being offered the short webinars in the control group.

Our pre-analysis plan (PAP) noted that we had a simple three-step process in mind, in which training should lead firms to adopt new business practices and take specific actions recommended by the program, then the firm may pivot, innovate, and gain short-term traction with customers, and thirdly, these changes could lead to firms experiencing more demand for their products and achieving more customers, sales, and expansion into new markets. As we note above, the pandemic led to some changes in the program delivery, which coupled with the pattern of take-up, meant that there was even more focus on customers and markets and digital presence, and less on some of the other business practices. We focus then on the actions and outcomes most closely related to what was actually taught in the program, with an online populated PAP,[8] and appendix Appendix E providing details on a few other outcomes that we either were not able to measure well, or that no longer coincided with the changed focus of the program.

We use two approaches towards multiple hypothesis testing. The first is to reduce the number of outcomes through building aggregate indices of closely related outcomes within the same domain. For example, we build indices of different business practices and different aspects of improved digital presence on Google or Facebook. Second, we also provide sharpened q-values (Benjamini et al., 2006) which maintain the False Discovery Rate when making comparisons across multiple outcomes.

Obtaining data from these firms proved to be challenging, and the result is attrition for our survey measures. We have shown that the attrition rates do not differ significantly with treatment status, that baseline means remain balanced for the sample answering the survey, and that the follow-up digital presence measures (that have no attrition) do not differ significantly for attritors and non-attritors. Taken together, this suggests that an assumption of missing-at-random may be a reasonable approximation. Nevertheless, we also use two additional approaches to examine robustness to attrition. The first is to use the post-double-selection lasso (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). In practice we do not find any of our baseline covariates being selected for the treatment prediction equation, and our PDS lasso estimates are similar to our results using Ancova and strata fixed effects. Second, in Table A4, we show our main results are robust to using the bounding approach of Lee (2009).

3.3. Impacts on digital presence, business practices and pitching to judges

Table 2 tests whether firms implemented changes in their digital presence and business practices that had been recommended in the training, and whether they were able to pitch more effectively to judges. Columns 1 through 3 examine the impact on digital presence. We see positive and significant impacts on the Google score and Facebook score in columns 1 and 3, and a small and statistically insignificant impact on the website score. Table 3 unpacks these changes in Google score and Facebook score to see where the improvements are coming from. We see the improvement in the Google score is driven by firms having a higher page rank, and being 10 percentage points more likely to show up in the first page of search results (relative to a control mean of 59%). This makes it more likely that potential customers will encounter the firm in an online search, and reflects the emphasis on using search engine optimization. The improvement in our Facebook score comes from improvements in multiple dimensions, including at the extensive margin of being 6 percentage points more likely to have a dedicated business Facebook page, and then a series of changes that make the pages more likely to attract customers. The pages are 9 percentage points more likely to provide a clear description of what the business does, 11 percentage points more likely to include stories featuring how the company has helped customers, 10 percentage points more likely to include a special offer, and 16 percentage points more likely to feature content in English. These features are designed to help firms attract more customers, including those in international markets. These impacts on the Google Score and Facebook score are robust to using principal components as an alternative way of aggregating the different measures.

While firms significantly improved their independently evaluated digital presence, this was not accompanied by a broader improvement in self-reported other business practices. Column 4 of Table 2 shows a near-zero and statistically insignificant impact on an overall index measure of practices in customers and markets, human resources, technology adoption, finance and accounting, and networking that were ex ante areas the training was anticipated to cover. Appendix Table A6 shows this also is true for each sub-component index. This likely reflects a combination of firms dropping out of some of the later sessions of training, some of the advice provided being too high-level and not directly actionable, and firms needing to focus on boosting customers during the recovery from the pandemic. Column 5 shows we do see a significant 16.5 percentage point impacts on firms reporting they use customer personas to help target their customer focus, which corresponds with them focusing mostly on efforts to boost customers during this time.

Finally, Column 6 of Table 2 shows no significant impact on the scoring of submitted pitch decks by judges. The mean score is 2.89 out of 5, and the point estimate is very close to zero at −0.01 points. This scoring occurred only 6 months after firms had

[8] Available in the AEA RCT registry at https://www.socialscienceregistry.org/versions/180228/docs/version/document.

Table 2
Treatment Effects on Digital Presence, Business Practices and Judges.

	(1)	(2)	(3)	(4)	(5)	(6)
	Google Score	Web Score	Facebook Score	Business Practices	Customer Personas	Judges Scores
Panel A: Ancova + Strata Fixed Effects						
Assigned to Treatment	0.071**	−0.020	0.050**	0.011	0.165**	−0.007
	(0.035)	(0.023)	(0.022)	(0.026)	(0.077)	(0.104)
Lagged Variable Included	Yes	No	Yes	Yes	Yes	No
Mean of Control Group	0.563	0.531	0.431	0.438	0.541	2.891
Sample Size	225	225	225	157	153	136
Sharpened q-value	0.104	0.409	0.104	0.709	0.104	0.895
Panel B: Using PDS Lasso to Select Additional Controls						
Assigned to Treatment	0.071**	−0.016	0.050**	0.011	0.165**	−0.007
	(0.035)	(0.022)	(0.022)	(0.026)	(0.077)	(0.099)
# Vars Selected to Model Outcome	0	2	0	0	0	0
# Vars Selected to Model Treatment	0	0	0	0	0	0
Sample Size	225	225	224	157	153	136
Sharpened q-value	0.095	0.513	0.095	0.699	0.095	0.890

Notes: Regressions control for randomization strata fixed effects and lagged outcome where available. PDS Lasso controls for randomization strata fixed effects and partials out lagged outcome variable. Number of vars then denotes the number of additional variables (out of 32) selected in either the outcome or treatment equations. Robust standard errors in parentheses. Significance levels are denoted by: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. The digital presence measures Google Score, Web Score and Facebook Score score firms according to their features respectively on Google searches (such as English description, right menu, Google search page rank etc.), their website (18 features such as loading speed, call to action, product(s) presentation clarity etc.), and their Facebook profile (15 features such as testimonials stories, English posts, promotions etc. Business Practices is an index of the proportion of all practices in customers and markets, human resources, technology adoption, finance and accounting, and networking that the firm is employing. Customer Personas is a binary outcome for whether the firm has created personas for its target customers. Judges Scores is the score awarded by judges to the submitted pitch deck of the firm). The Google, Web, Facebook, and Business Practice scores are on a scale from 0 to 1, while the Judges scores are on a scale from 1 to 5.

Table 3
Detailed Treatment Effects on Google and Facebook Scores.

	Control	Treatment	Standard	Sharpened
	Mean	Effect	Error	q-value
Firm shows up in Google in the first page of search results	0.585	0.102**	(0.048)	0.094
(Inverse of) Rank at which firm shows up in the results	0.550	0.091**	(0.043)	0.094
Firm's website first on the page of results	0.580	0.103**	(0.048)	0.094
Google description in English	0.607	0.060	(0.042)	0.187
Google right menu shows a map/contact information	0.482	0.044	(0.040)	0.220
Clear google description	0.625	0.042	(0.042)	0.220
No modifier (doo/shpk) needed to find the firm on google	0.509	0.057	(0.045)	0.202
Google Score	0.563	0.071**	(0.035)	
PCA Google Score	−0.246	0.477**	(0.235)	
Has a Facebook Page	0.866	0.064*	(0.034)	0.173
Business Specific	0.871	0.059*	(0.034)	0.184
Posted in the past week	0.384	0.043	(0.052)	0.318
Clear description	0.723	0.093**	(0.044)	0.173
Clear value proposition	0.478	0.013	(0.052)	0.541
Easy to find contact	0.813	0.075*	(0.039)	0.173
Appealing photos	0.786	0.061	(0.041)	0.203
Customer reviews left in the past 3 months	0.045	0.018	(0.028)	0.389
Customer stories	0.259	0.112**	(0.055)	0.173
Stories highlighting progress in the last 3 month	0.286	−0.055	(0.048)	0.266
Special offers in the last 3 months	0.094	0.104***	(0.038)	0.060
English language version/English posts	0.344	0.155***	(0.057)	0.060
Serbo-Croatian language version/Serbo-Croatian posts	0.393	0.017	(0.050)	0.541
Albanian language version/Albanian posts	0.183	0.053	(0.045)	0.266
Online orders possible from Facebook shop	0.094	−0.009	(0.036)	0.541
Number of Facebook followers	3834	2340	(1494)	0.203
Facebook Score	0.431	0.050**	(0.022)	
PCA Facebook Score	−0.176	0.555**	(0.231)	
Number of Observations	225			

Notes: Data available for all firms, so sample size for each regression is 225. Regressions include controls for randomization strata and baseline value of overall index. PDS Lasso used to select additional baseline variables. Robust standard errors in parentheses. Significance levels are denoted by: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Scores are independently scored aspects of the firms digital presence. Google Score and Facebook Score are the overall index measures in Table 2 and weight each component equally. PCA Google Score and PCA Facebook Score take the first principal component as an alternative weighting scheme. Sharpened q-values are for the family of Google index components, and for the family of Facebook index component families respectively.

started the program, and firms are unlikely to have changed their team, technology, or fundamental value proposition for investors during this time, and may only just be starting to gain traction by expanding into new markets. Appendix Table A8 shows firms were no more likely to innovate or pivot, so are not changing their product mix. Higher scores from judges could arise if training taught firms how to better pitch themselves to investors, but fewer firms attended the last session of the training on pitching, and the content may not have been very new for firms given previous experience putting together pitch decks for investors or start-up competitions in the region. Moreover, during the pandemic, the supply of outside venture funding fell, as did interest in getting funded while firms focused on recovering or boosting sales. Appendix Table A8 shows no significant impact of treatment on interest in outside investment, nor in the likelihood of receiving investment. This reduced interest in immediate outside investment may also explain why only 60% of firms submitted pitch decks to be judged.

3.4. Impacts on customers, sales, and exports

We then examine whether the customer and market-focused training and one-on-one sessions were successful in helping firms boost their customers and expand their market. Column 1 of Table 4 estimates the impact of treatment on the number of customers in 2021 (winsorized at the 95th percentile). We see a large, positive, and significant impact of 159 customers, which represents an 85% increase relative to the control mean. Fig. 2 examines where in the distribution this improvement is taking place. We see that the baseline distributions of number of customers are extremely similar for treatment and control, and then the post-treatment shift is largely happening for the top 20% or so of the distribution. Column 2 shows a positive, but not statistically significant, impact of 12% on revenues, which is consistent with more customers bringing higher revenues for firms, but imprecisely measured. We had difficulty getting firms to report profitability, and the profits data are highly skewed and contain some negative and zero values. Appendix Table A8 shows positive but insignificant impacts on the inverse hyperbolic sine of profits, and a positive impact on winsorized profits that is marginally significant. However, we view these results on profits as at most suggestive given data quality issues.

The remaining columns of Table 4 then report the impacts on firm exports. Column 3 shows a positive, but not significant, impact on total exports in 2021. This unconditional effect includes many zeros for firms not exporting, along with impacts for those that are exporting. Recent work by Chen and Roth (2023) has highlighted that in such a setting with many zeros, the estimated treatment effects of inverse-hyperbolic sine transformations are scale-dependent, and reflect a combination of the extensive margin effect and intensive margin effect, with the weighting depending on the units of measurement used. We therefore look separately at the extensive and intensive margins.

Column 4 shows 67.6% of the control group were exporting in 2021, and that there is a statistically insignificant impact of 4.9 percentage points on this extensive margin. Table A7 also shows statistically insignificant impacts on other extensive margin impacts of market expansion, in terms of whether they sell in a new city or country, how many countries they sell to in the Western Balkans, and whether they sell in the European Union. An overall index of these measures has a statistically insignificant 0.09

Table 4
Treatment effects on Firm Performance and Market Expansion.

	(1) Total Customers	(2) I.H.S. Revenues	(3) I.H.S. Exports	(4) Exports at all (Extensive)	(5) Log Exports (Intensive)
Panel A: Ancova + Strata Fixed Effects					
Assigned to Treatment	159**	0.123	0.198	0.049	0.661**
	(77)	(0.379)	(0.643)	(0.071)	(0.296)
Lagged variable included	Yes	Yes	Yes	Yes	Yes
Mean of Control Group	186	12.294	7.162	0.676	9.894
Mean of Control in Winsorized Levels	186	466,020	67,891		113,727
Sample Size	152	143	139	139	97
Number of Zeros	0	4	42	42	0
Sharpened q-value	0.115	0.835	0.835	0.835	0.115
Panel B: Using PDS Lasso to Select Additional Controls					
Assigned to Treatment	159**	0.123	0.198	0.049	0.661**
	(73)	(0.358)	(0.591)	(0.067)	(0.265)
# Vars Selected to Model Outcome	0	0	0	0	0
# Vars Selected to Model Treatment	0	0	0	0	0
Sample Size	152	143	139	139	97
Sharpened q-value	0.067	0.793	0.793	0.793	0.067

Notes: Regressions control for randomization strata fixed effects and lagged outcome where available. PDS Lasso controls for randomization strata fixed effects and partials out lagged outcome variable. Number of vars then denotes the number of additional variables (out of 32) selected in either the outcome or treatment equations. Robust standard errors in parentheses. Significance levels are denoted by: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Number of Customers is winsorized at the 95% level and is the number of customers in 2021. IHS Revenues is the inverse hyperbolic sine (IHS) transformation of revenues in quarters 2 and 3 of 2021. For the outcome of export sales, we show IHS Export Sales, which refers to IHS transformation of export sales, the dummy Export Sales (Extensive) that is 1 if the export sales are positive, and Log Export Sales (Intensive), the intensive measure for those firms that do export.

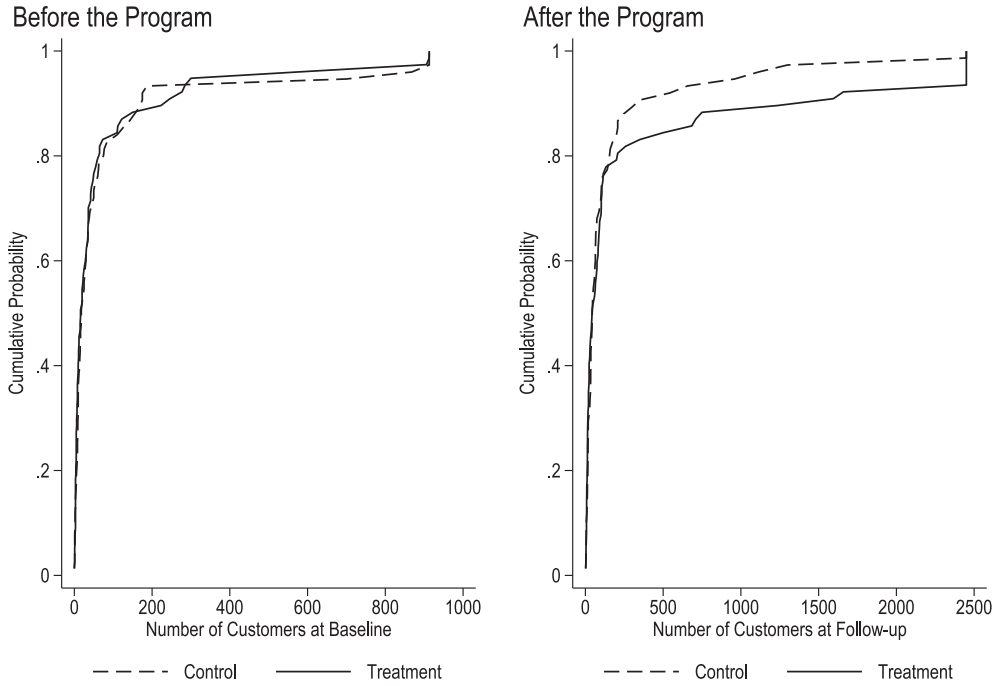

Fig. 2. The Distribution of Total Customers Shifted Right for Treated Firms.
Notes: Cumulative distributions by treatment group of the the number of customers before and after the program (winsorized at the 95th percentile).

standard deviation effect. So there may be a small positive increase at the extensive margin, but since the majority of firms were already exporting, the main room for improvement comes at the intensive margin. Moreover, some of the firms not exporting were in sectors where it would be difficult or not possible to do so - such as several firms in tourism, healthcare, and education services. The non-exporters tend to be smaller in size, with lower revenues and fewer employees than exporters.

The training may be more successful in helping firms that are exporting to export more by reaching new customers and additional markets. Column 5 of Table 4 examines the impact at the intensive margin for those firms that are exporting, and shows a large positive and significant impact.[9] The magnitude of 0.66 log points, is equivalent to an approximate 94% increase, although the standard error of 0.27 log points indicates considerable uncertainty as to the exact magnitude. Fig. 3 estimates impacts quarter by quarter for the quarters we have of reported exports. We see pre-treatment, in 2020, treated exporting firms slightly lagged control exporting firms in the volume of exports. In contrast, post-intervention, the difference is now positive in quarter two of 2021, and positive and significant by quarter three of 2021.

Fig. 4 compares the distribution of exports pre- and post-intervention. We see the mass of firms with zero exports, and similar pre-treatment distributions. The improvement in export sales then is seen most in the middle of the distribution, between the 30th and 80th percentiles.

3.5. Qualitative evidence

Our quantitative results show that firms were able to improve their digital presence, attract more customers, and improve exporting at the intensive margin. In order to ensure these improvements were genuine, and to get more insights on how the program had helped firms achieve these outcomes, we supplemented our quantitative results with in-depth qualitative interviews. We interviewed firms that had improved their exports or digital presence the most (a positive deviance approach). These case studies confirmed firms had improved, and provided some more detailed examples of how the program had helped firms to do this.

Some of the improvement in digital presence appears to have come through companies learning about new marketing methods that they did not know much about before, as well as prompting them to devote more attention to their social media efforts. For example, the manager of a Bosnian company selling specialized safety equipment said that he was not aware of SEO before the program, and had relied on reaching customers through email newsletters to existing customers only. They had now hired someone to help them implement this, and also were experimenting with online sales as a way to reach a broader customer base. They also noted that Facebook was prevalent for their customer base, and had invested more in using Facebook

[9] Conditioning on exporting could lead to biased results if treatment affected the characteristics of which firms export. If anything, we would expect this to lead to a downward bias, if treatment induced smaller treated firms to start exporting. Our PDS Lasso estimates do not select any variables to explain treatment, which coupled with the lack of significant impact on the extensive margin of exporting, suggests that any such bias is small in our case.

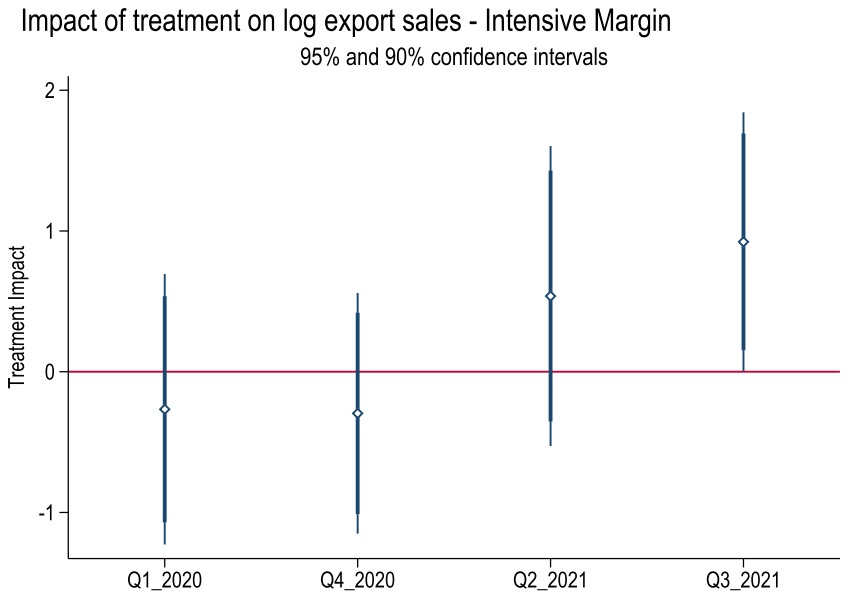

Fig. 3. The Intensive Margin Impact on Exports Over Time.
Notes: Impact of treatment on log export sales by the reported quarters (first and fourth quarter of 2020, second and third quarter of 2021).

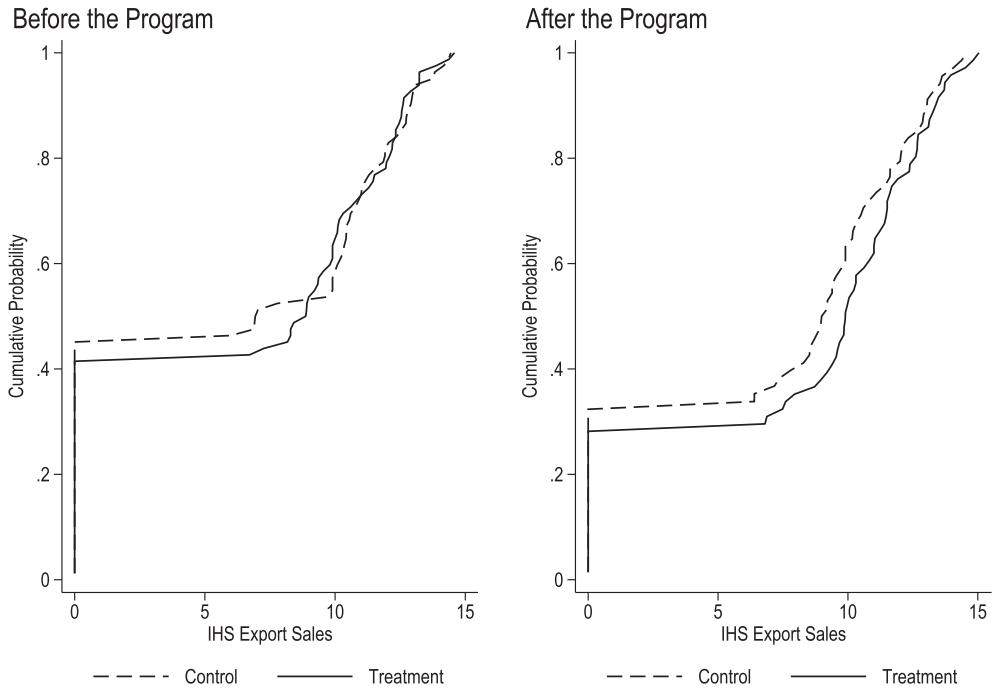

Fig. 4. The Distribution of Exports Shifted Right for Treated Firms.
Notes: Cumulative distributions by treatment group of the inverse hyperbolic sine (IHS) transformation of export sales before and after the program, split by treatment group.

to attract customers. A Bosnian company making baby-carriers noted that the training had caused them to start Google SEO, and focus more on their social media. A Serbian firm producing smart agricultural devices also noted how the program had prompted them to really focus on improving their digital presence, and they used in-house staff to implement SEO, increased their presence on LinkedIn, and revamped their website landing page.

The CEO of Blinking, a Serbian company specializing in digital identity verification, which placed third in the final competition, stated that the program helped them to become better at addressing customers' needs and marketing, with “the most important

part related to growth beyond the Western Balkans region."[10] Interviews with other companies provided more specifics on how this export expansion occurred, with two of the main channels being a combination of specific advice that was received in the one-on-one sessions, coupled with encouragement giving them more confidence and impetus to act on pre-existing intentions to expand markets. For example, the Bosnian specialized safety equipment firm noted they received specific advice from the Deloitte consultant about the terms and conditions needed for entering into regional DIY retail chains, information about the size of the potential market, and help in defining what types of customers they should focus their expansion on. This helped them become more decisive in expanding their regional business. The Bosnian baby carrier firm noted the consultant had recommended that they open a new storage location in Croatia, test the new market there, and then consider that a logistical gate to Europe. Following that advice, they realized that they needed to upgrade their product with better fabric and hand-made details to differentiate from competitors, and they were able to increase export sales by 90% as a result.

These channels stressed in the qualitative interviews also help explain why not all firms may have been able to benefit as much from the program. A Serbian mobile-development firm owner noted that the program likely pushed firms in many sectors to renovate themselves and improve digital presence, whereas for innovative start-ups in the B2C digital space (which most of our firms were not), this would have been less relevant. However, it is notable that many ICT firms still did benefit from help on their digital marketing: often founders with engineering or computer science backgrounds were technically strong on developing tailored ICT service solutions, but had not had training or experience in modern digital marketing. Another reason that came up for some firms not benefiting as much was that the heterogeneity of industries and products meant that the Deloitte consultants had more specific knowledge on how to expand exports in some industries compared to others.

3.6. Cost-effectiveness and cost-benefit

The treatment of training and consulting had an average cost of approximately $3350 per firm assigned to treatment. This includes fixed costs of recruitment and content development, and the marginal cost of just holding the training and consulting was approximately $2140 per firm assigned to treatment. Most of the costs of the control offering were the fixed costs of recruitment and content development, so the average cost of $690 per firm are much larger than the marginal cost of just holding the webinars, of $81 per firm. We then consider two approaches to understanding whether such a program is worth its cost from a policymakers' perspective: cost-effectiveness and cost-benefit.

Cost-effectiveness is the easier of the two to discuss. Policymakers in many countries have a goal of increasing firm exports and building the supply of firms that may eventually be attractive to outside investors. As our introduction notes, cheap and light-touch approaches such as information interventions and one-day seminars have not been successful in improving exports. Atkin et al. (2017) were successful after several years of effort in spurring exports in Egypt, but they specifically note that given the difficulties they faced in generating orders and the implicit labor costs of their time, it was unclear whether their program would pass a cost-benefit test, and they do not document the total costs. In Colombia, the consulting services used by Iacovone et al. (2023) to attempt to improve productivity and exports cost $13,800 per firm and did not increase exports. Given the dearth of programs demonstrated to increase exports, PowerUP appears relatively cost-effective at achieving this goal, with the online delivery of content helping to keep costs low. It is also cheaper than the in-person prior investment readiness program, Pioneers of the Balkans, which cost $4065 per treated firm (Cusolito et al., 2021).

A full cost-benefit analysis would require more detailed information on firm profit impacts over the long-term, as well as whether there were any demonstration effects that got other firms to export, or spillover benefits to workers in terms of higher pay. Our revenue and profits data are highly skewed and noisy, with imprecise estimates. However, as a rough approximation, if we take the point estimate of the increase in revenues over six months of 12%, this equates to a €52,000 increase in revenues at the mean, and €17,000 increase at the median. Assuming a conservative profit margin of 10% would yield an approximate gain in profits of €1700 to €5200 over six months, or approximately the same magnitude as treatment costs. This calculation has large uncertainty associated with it, but at least suggests that it is plausible that the intervention pays for itself within the first year[11] However, this lack of precision and uncertainty over impacts can make it hard for individual firms to know whether such an investment would be worth it, as Lewis and Rao (2015) show is the case for advertising in general.

4. Conclusions

An all-online capacity building program that combined group-based training and short one-on-one consulting sessions was able to help SMEs to boost their digital presence, gain more customers, and sell more exports. The main mechanisms for this improvement appear to be through helping make some firms aware of the importance of modern digital marketing practices, providing specialized advice on acquiring customers in the European market, and giving entrepreneurs confidence that they were on the right track, which encouraged them to strengthen their efforts to attract more customers. This may help give the firms the traction they need to be more attractive to outside investors over a longer run, but in the short-term, firms were not any more likely to be interested in outside investment, or to have their pitch decks scored any better by outside judges.

[10] We name this company since they gave permission for their story to be publicly featured: https://webalkans.eu/en/stories/who-are-you-serbias-new-identity-verification-technology/.

[11] Using the noisy profit gain in Table A8 would show an even faster recovery of the costs of training.

Given the challenges of rapidly pivoting the format and content of the capacity building program during the COVID-19 pandemic, we view these results as promising. However, we also see several areas where improvement could be made for future programs. The first is in the targeting and selection of firms. The heterogeneity of firms across sector and lifecycle stage makes it more difficult to provide training content that meets the needs of all firms. The one-on-one sessions did provide some opportunities for tailored advice, but having a more homogenous pool of firms would allow for more specific training content. Second, and related, firms noted that parts of the training content were too general and high-level, and while they offered a general understanding, did not lead to prompts for specific actions that they could employ in their firms. Rather than trying to cover many topics at a strategic or general level, it may be better to concentrate on really focused activities that firms can immediately use in their firms. We see improvements in some of the areas that firms received specific instructions and homework on, such as using SEO.

Third, while the health concerns of the COVID-19 pandemic required all training and consulting to take place online, we believe that a mixed modality may be more effective in future programs. Our experiment shows that online live sessions do offer some advantages and can deliver some improvements. But complimenting these with some in-person sessions would allow more opportunities for firms to network with one another and to have more interactive and tailored discussions with trainers.

Finally, we note that these types of innovative SMEs are very reluctant to share their data, making tracking and measurement of firm outcomes difficult. Firms were required to sign a Memorandum of Understanding (MOU) that committed them to contribute their time to attending and to submitting data sheets. But in practice firms needed to be contacted very many times to obtain data from them, and the MOU acted more as a soft commitment device than something that they felt legal obligation to adhere to. The predominance of service sector firms mean that even if firms were able to be linked to administrative data on exports, this would not help since service exports do not get captured by border transaction data. From a methodological standpoint, in order to measure impacts of these programs in the future, and over longer time periods, it would be useful to require firms at the time of application to both legally commit to providing three years of annual financial data, along with written permission to have their sales and tax data released to the program implementers by the corresponding tax authorities. Offering firms some annual feedback or benchmarking based on the data they and other firms provide could help make this also a beneficial activity for firms.

Data availability

https://data.mendeley.com/datasets/xkrpv4khnj/1 (Original data) (Mendeley Data)

Declaration of Competing Interest

The authors declare that they have no competing interests to report, and that all funding sources have been reported in the paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jinteco.2023.103794.

References

Allen, T., 2014. Information frictions in trade. Econometrica 82 (6), 2041–2083.
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 (1), 157–209.
Anderson, J., van Wincoop, E., 2004. Trade costs. J. Econ. Lit. 42 (3), 691–751.
Arkolakis, C., 2008. Market penetration costs and the new consumer margin in international trade. J. Polit. Econ. 118, 1151–1199.
Atkin, D., Khandelwal, A., Osman, A., 2017. Exporting and firm performance: evidence from a randomized experiment. Q. J. Econ. 132 (2), 551–615.
Bakos, Y., 2001. The emerging landscape for retail e-commerce. J. Econ. Perspect. 15 (1), 69–80.
Belloni, A., Chernozhukov, V., Hansen, C., 2014. Inference on treatment effects after selection among high-dimensional controls. Rev. Econ. Stud. 81, 608–650.
Benjamini, Y., Krieger, A., Yekutieli, D., 2006. Adaptive linear step-up procedures that control the false discovery rate. Biometrika 93, 491–507.
Bergquist, L., McIntosh, C., Startz, M., 2022. Search Costs, Intermediation, and Trade: Experimental Evidence from Ugandan Agricultural Markets Working Paper.
Breinlich, H., Donaldson, D., Nolen, P., Wright, G., 2017. Information, Perceptions and Exporting – Evidence from a Randomized Controlled Trial Working Paper.
Cadot, O., Fernandes, A., Gourdon, J., Mattoo, A., 2015. Are the benefits of export support durable? Evidence from Tunisia. J. Int. Econ. 97 (2), 310–324.
Carballo, J., Chatruc, M.R., Santa, C.S., Martincus, C.V., 2022. Online business platforms and international trade. J. Int. Econ. 137 (103599).
Chen, J., Roth, J., 2023. Log-Like? Identified Ates Defined with Zero-Valued Outcomes Are (Arbitrarily) Scale-Dependent Working Paper.
Couture, V., Faber, B., Gu, Y., Liu, L., 2021. Connecting the countryside via e-commerce: evidence from China. Am. Econ. Rev. Insights 3 (1), 35–50.
Cusolito, A., Dautovic, E., McKenzie, D., 2021. Can government intervention make firms more investment-ready? A randomized experiment in the western balkans. Rev. Econ. Stat. 103 (3), 428–442.
Frankel, J.A., Romer, D.H., 1999. Does trade cause growth? Am. Econ. Rev. 89 (3), 379–399.
Goldberg, P.K., Reed, T., 2021. Demand-Side Constraints in Development: The Role of Market Size, Trade, and (in)Equality Working Paper.
Goldfarb, A., Tucker, C., 2019. Digital economics. J. Econ. Lit. 57 (1), 3–43.
Görg, H., Henry, M., Strobl, E., 2008. Grant support and exporting activity. Rev. Econ. Stat. 90 (1), 168–174.
Greenstein, H., 2011. How to Show Market Traction Inc.com.
Iacovone, L., Maloney, W., McKenzie, D., 2022. Improving management with individual and group-based consulting: results from a randomized experiment in Colombia. Rev. Econ. Stud. 89, 346–371.
Iacovone, L., McKenzie, D., Meager, R., 2023. Bayesian Impact Evaluation with Informative Priors: An Application to a Colombian Management and Export Improvement Program Working Paper.

Jin, Y., Sun, Z., 2021. Lifting Growth Barriers for New Firms: Evidence from an Entrepreneur Training Experiment with Two Million Online Businesses Working Paper.
Kim, Y.R., Todo, Y., Shimamoto, D., Matous, P., 2018. Are seminars on export promotion effective? Evidence from a randomised controlled trial. World Econ. 41 (11), 2954–2982.
Lee, D., 2009. Training, wages, and sample selection: estimating sharp bounds on treatment effects. Rev. Econ. Stud. 76, 1071–1102.
Lewis, R., Rao, J., 2015. The unfavorable economics of measuring the returns to advertising. Q. J. Econ. 130, 1941–1973.
Martincus, C.V., Carballo, J., 2008. Is export promotion effective in developing countries? Firm-level evidence on the intensive and the extensive margins of exports. J. Int. Econ. 76 (1), 89–106.
McKenzie, D., Woodruff, C., Bjorvatn, K., Bruhn, M., Cai, J., Gonzalez-Uribe, J., Quinn, S., Sonobe, T., Valdivia, M., 2021. Training entrepreneurs. VoxDevLit 1 (2).
Munch, J., Schaur, G., 2018. The effect of export promotion on firm-level performance. Am. Econ. J. Econ. Pol. 10 (1), 357–387.
Srhoj, S., Vitezić, V., Wagner, J., 2020. Export Boosting Policies and Firm Behaviour: Review of Empirical Evidence around the World. Working Paper395.
Verhoogen, E., 2022. Firm-level upgrading in developing countries. J. Econ. Lit. forthcoming. https://www.aeaweb.org/articles?id=10.1257/jel.20221633&from=f.