ELSEVIER
www.elsevier.com/locate/worlddev

World Development Vol. 39, No. 9, pp. 1592–1610, 2011

0305-750X/$ - see front matter

doi:10.1016/j.worlddev.2011.04.021

Can Entrepreneurial Activity be Taught? Quasi-Experimental Evidence from Central America

BAILEY KLINGER
Harvard University, Cambridge, Massachusetts, USA

and

MATTHIAS SCHÜNDELN *
Goethe University Frankfurt, Germany

Summary. — Business training is a widely used development tool, yet little is known about its impact. We study the effects of such a business training program held in Central America. To deal with endogenous selection into the training program, we use a regression discontinuity design, exploiting the fact that a fixed number of applicants are taken into the training program based on a pre-training score. Business training significantly increases the probability that an applicant to the workshop starts a business or expands an existing business. Results also suggest gender heterogeneity as well as the presence of financial constraints.

Key words — entrepreneurship, business training, regression discontinuity design, Central America, El Salvador

1. INTRODUCTION

Entrepreneurial activity is important for development (e.g., Baumol, 1968): Entrepreneurs implement new business ideas, or adopt profitable ideas from others to local circumstances to start new businesses, or they experiment with new materials and processes to expand their business. Inevitably some fail. A process of creative destruction, as described famously by Schumpeter (1911), leads to progress. Recognizing this, development organizations increasingly use business training as a development tool. On the other hand, many classic theories of entrepreneurial activity treat entrepreneurial ability as exogenous (e.g., Lucas, 1978). Thus it is unclear if a business training program that trains individuals that do not have a business (and may, therefore, not have the entrepreneurial ability) can even have the most immediate intended effect, namely start-up of a business.

In this paper we exploit the specifics of the training program that we study to implement a quasi-experimental research design (a regression discontinuity design) to study whether entrepreneurial activity can be taught, in particular whether business training can lead to increasing numbers of businesses started or expanded. We will study this question by analyzing the results of business training programs that the NGO TechnoServe held in Central America during 2002–05. The program is intended for both individuals who wish to start a business, as well as for those who already have a business. Those individuals with existing businesses have about 10 employees on average. Thus, unlike some other programs, this program targets businesses of a size beyond that of household enterprises.

It is now recognized that small and medium sized enterprises (SMEs) can play a significant role in the development process. A particularly well-known proponent of entrepreneurship through different phases of development has become the Global Entrepreneurship Monitor (GEM) consortium, which documents and analyzes the relationship between entrepreneurship and development (e.g., Bosma & Levie, 2010). This consortium works with a model (Bosma, Acs, Autio, Coduras, & Levie, 2009) that considers how the importance of various environmental conditions for entrepreneurship changes at different phases of economic development. At the same time, according to this particular model, "the relative importance of entrepreneurship education and training increases as economies develop economically" (Coduras Martínez, Levie, Kelley, Sæmundsson, & Schøtt, 2010, p. 12).

Obstacles that entrepreneurs face may be external and apply to all entrepreneurs in an economy, such as regulations (e.g., Djankov, La Porta, Lopez-de-Silanes, & Shleifer, 2002), infrastructure (e.g., Datta, 2008), or macroeconomic uncertainty and property rights (e.g., Svensson, 1998). In these cases, action at the macro level might be necessary. However, many constraints are idiosyncratic to the (potential) entrepreneur. For example, individual entrepreneurs may not be willing to take risks. Some may simply not know how to navigate regulations and how to deal with other formal aspects of running a business, or how to deal with banks to finance their new business or expand their existing business. "Entrepreneurial spirit" appears often to be seen as inherent in individuals, yet even for those who possess this entrepreneurial drive, they may not actually start a business or expand an existing business because of the myriad of potential obstacles. However, some of the idiosyncratic obstacles could potentially be overcome through business-specific education of (potential) entrepreneurs. Theoretically, any of the central aspects of

* We have received very useful comments from Guido Imbens, Nicola Fuchs-Schündeln, Dan Levy, seminar participants at the World Bank and anonymous referees. We are grateful for research assistance from Michael Durán, and support from many individuals at TechnoServe, in particular Carolina Cely, Sandra Jetten, Steve Londner, and Bruce McNamer. Final revision accepted: January 20, 2011.

entrepreneurship—the GEM model for example identifies three main components, namely attitudes, activity and aspirations—could be affected by training and lead to business launch or expansion of an existing business. This paper's goal is to investigate whether this is possible.

We study the effect that a specific training program implemented by an NGO in a large number of countries has on entrepreneurial activity, namely on business start-up and business expansion. In addition, another feature of the program, the quasi-experimental injection of substantial amounts of capital into some businesses, also allows us to investigate whether entrepreneurs are financially constrained, and the extent to which financing constraints hinder investment. Finally, we also investigate whether the effect of training and financing varies by gender. Although these questions will be studied in the context of one particular organization, the broader goal of this paper is to shed some light on the larger question of whether entrepreneurial ability is exogenous and inherent to a firm owner or to what extent and how it can be shaped by training.

TechnoServe supports SMEs in the form of the so-called "Business Plan Competitions", which are intended to help individuals who are interested in setting up a new business to gain the necessary skills, and to help existing small and medium-scale entrepreneurs who wish to significantly expand their businesses improve their skills and entrepreneurial abilities. In each competition, there is first a preliminary screening process. A fixed number of applicants are selected for the program, which then receive the training. At the end of this phase, applicants submit a first draft of their business plan.

Based upon a review by a small panel of "judges", that is individuals who are knowledgeable about business, which may include for example business owners or business consultants, a smaller group is selected to continue to the final phase, where they receive additional support and business development services to complete their business plan. Note that the judges typically come from large companies so that they are not potential competitors to business plan competition participants (who are seeking to launch/expand smaller firms) and that the judges are not involved in any of the training before the presentations. The finalized plans are then again evaluated by judges, and a fixed number of the top plans receive a financial prize of between US$6,000 and US$15,000 (depending on country and year), receipt of which is conditional on investment in the business, plus some additional business development services.

Existing evaluations of similar programs in the context of developing countries cannot answer the key question that is at the core of any program evaluation: How would the entrepreneur have done if he or she had not gone through the training workshop? The usual approaches of finding comparison groups are based on observable characteristics. However, we can expect self-selection of entrepreneurs into a program based on unobservable characteristics, as for example unobserved business opportunities. If these same (unobserved) characteristics also determine the future success and self-selection is not taken into account in the research design, then estimates of the treatment effects will be biased. Similarly, the purposeful assignment of individuals to a training program, for example of unemployed individuals, would lead to biased estimates if it is not properly controlled for.

To overcome the econometric problems of self-selection and purposeful program placement, we take advantage of one particular feature of the program under study. Applicants have to undergo a standardized evaluation process in which a score is determined that is supposed to characterize the potential entrepreneurial ability of an applicant. The sole decision criterion for access to the workshop is this score. If the score is above a cutoff, the applicant is accepted to the workshop; otherwise he/she is rejected. This feature can be used to study the effect of program participation on outcomes based on a regression discontinuity design, by comparing rejected and accepted applicants.

We focus on the first-order impact that a business development program of the kind that we study attempts to achieve, namely the question whether training actually induces participants to start entrepreneurial activities, that is to start up a business or significantly expand an existing one. We view these outcomes as key to the above cited literature. If entrepreneurial skills are rather fixed, then there should be no significant effect of a training program on start-up or expansion of businesses and consequently there cannot be a significant effect on secondary outcomes, such as sales, production, job creation etc. Thus, establishing whether business development programs that aim at creating new or larger businesses are successful in this primary goal is an important first step in this research agenda.

Our findings indicate that the program is successful in promoting entrepreneurial activity. We find economically important and statistically significant changes in the probability that individuals open a new business or expand an existing business that is due to participation in the full training program. Looking at the different stages of the program, we find that the first round seminar-based training seems to affect the expansion of businesses more than the launching of new businesses. On the other hand, the second round, in which the business plan is developed more fully with more one-on-one assistance, affects more the launching of businesses. Similarly, the last round treatment, which is the receipt of prize money, has significant effects on launching, but smaller and less significant effects on expansion of businesses.

The research is of immediate practical relevance. The development community is increasingly focusing on promoting Private Sector Development, recognizing that entrepreneurship is a key building block of sustained poverty-reducing economic growth.[1] Yet, despite an emerging body of work that studies the effects of business training, for example Coduras Martínez et al. (2010), Bosma and Levie (2010), and Karlan and Valdivia (in press), there is still much to be learned about what works in scalable enterprise-based solutions to poverty.

The paper proceeds as follows. We first provide background on the program that we study and the data that we use. We then investigate whether the discontinuity in program participation that we wish to exploit holds. After that, we introduce the regression discontinuity design in the context of our application. Section 5 then presents the main results, as well as robustness checks. In that section we also discuss the role of gender. The final section concludes.

2. THE TECHNOSERVE BUSINESS PLAN COMPETITIONS

This section provides some background about the business plan competitions that the NGO TechnoServe runs. A more detailed description is given in the appendix, which also includes a figure to illustrate the timeline of a competition (Table 11). Since 2002 TechnoServe organizes Business Plan Competitions to promote entrepreneurship. These competitions provide training to both nascent entrepreneurs seeking to start a new business as well as to entrepreneurs with existing businesses that hope to undertake a significant expansion into a new product or market.

The competitions function as follows. The first phase consists of the organization and its local partners publicizing the competition and collecting applications. Applicants give basic personal information as well as a summary of their business idea. Each application is scored, and the top applicants in each sector are accepted into the program. It is important to note that the number of applicants that are admitted into the program, and the number of participants that subsequently progress to each stage, is fixed before the competition begins. This creates an exogenous cutoff which we will exploit to evaluate the program. It is also important to note that the scoring at the conclusion of phases 1, 2 and 3 is performed using standardized guidelines, but the panel of judges is often different for different sectors. Therefore, the numerical scores may not be strictly comparable across sectors, years and countries, and for this reason we standardize scores at the country-year-sector level (see the data section below).

The accepted applicants are admitted into phase 2, which consists of an entrepreneurial training program (the "training stage 1"). In the case of El Salvador, this training is provided first by the UNCTAD's Empretec program over seven working days, which focuses on developing both technical business skills and "core entrepreneurial behaviors" (for details see the appendix), followed by a TechnoServe workshop on business plan preparation. In the other countries TechnoServe provides the entire training program in three to four workshops, covering such topics as entrepreneurial orientation and attitudes, strategic planning, financial projections, marketing, and how to write a business plan. This stage also provides country-specific information on starting a business, such as the relevant government departments and programs, different legal forms, and potential sources of capital. The participants then must prepare a formal business plan, which is submitted in writing and evaluated by a panel of judges.

The top scoring business plans at the conclusion of phase 2 ("training stage 1") proceed to phase 3 ("training stage 2"), in which the participants further refine their business plans and receive more one-on-one assistance with mentors and consultants. These refined plans are then presented to a jury of the above mentioned "judges", which gives each plan a detailed score. The top scoring business plans receive a monetary reward of approximately US$9,000 (between US$6,000 and US$15,000)[2] that is to be invested in a business. TechnoServe controls the disbursement to suppliers or pays directly for capital expenses to insure the funds are invested in the proposed business.

In the case of El Salvador, the training consists of the 7-day UNCTAD Empretec course, followed by two four-hour sessions by TechnoServe on how to compose a business plan using the methodology described above. The participants then have approximately 8 weeks before the business plans are due. In the case of Nicaragua and Guatemala, there is no Empretec training, and instead the classes based on the business plan methodology described above are done in more detail. They are divided up into approximately three full-day sessions, one per weekend. The participants have approximately one month to then compose and turn in their business plans. For all competitions, the plans advancing to phase 3 are revised over a period of 1–2 weeks.

It should be stressed that each stage of the training contains a number of different elements and we will not be able to disentangle precisely, which aspect of the training is responsible for any effect that we see, or whether it is indeed the combination of all aspects that leads to the results. In particular, it should be noted that the competitive aspect of the program, that is the fact that individuals receive feedback on their projects that are compared to other projects, implies a difficulty for the identification of the specific aspect of the program that leads to the effects that we observe: The competitive aspect of the program might be an important aspect in terms of leading to results, not just as a selection process to narrow down the group of individuals, but because it provides some form of external validation of the quality of the proposed business idea. As with the other program aspects that are lumped into the different stages of the training, unfortunately, we will not be able to separate the effect of this feedback that the entrepreneur receives from the effect of other aspects of the program.

In sum, the set-up is a multi-phased competition where the participants are scored at each phase and an exogenous cutoff is used to determine who proceeds to the subsequent phase. This allows for a more robust impact evaluation of the different stages of the program than in the case of programs where the cutoffs are often endogenous.

3. DATA

As part of its monitoring and evaluation efforts, TechnoServe attempts to survey all competition participants who were accepted into phase 2, about one year after the competition. They determine whether or not the participant launched the new business or carried out the expansion that they proposed in the competition. TechnoServe also asks for sales and employment figures for the business. The data from these surveys were provided to the authors by TechnoServe. Importantly, TechnoServe did not survey rejected applicants. Therefore, we collected additional, comparable data for observations below the phase 1 cutoff. In addition, we followed-up with competition participants who were accepted into the program but that TechnoServe was not able to reach in their own follow-up surveys to further complete the dataset. In addition to outcomes data, for each participant we have their status prior to entering the program, which is available from the application forms that each individual submitted: whether or not they were already operating a business, and if so, the number of employees and sales. Finally, we have demographic information, as well as the scores for each of the phases before the participants were cut from the competition.

We study business plan competitions in three different countries (El Salvador, Guatemala, and Nicaragua) over four years (with two competitions in Nicaragua and three in El Salvador), which had between 38 and 163 participants that were trained in at least one stage. Our sample consists of 655 accepted and rejected applicants to these entrepreneurial training workshops, 377 of these had received at least some training and 278 were rejected applicants who did not receive any training.

Because we had to follow-up with individuals several years after the program started, response rates and recall errors are potential reasons for concern. First, note that TechnoServe itself followed up with participants in the first round of training a year after the program. Out of these individuals, we have data on post-program status for 92.4% of those who participated in the first stage of the training program, and 97.3% for those who participated in the second stage of the training program. So recall and response rates are not an important concern for the later rounds of the training. Second, our central analysis is about business start-up and expansion, which are likely recalled with less error than if we were asking retrospectively about continuous variables such as sales or profits. Third, most of our interview attempts were with individuals

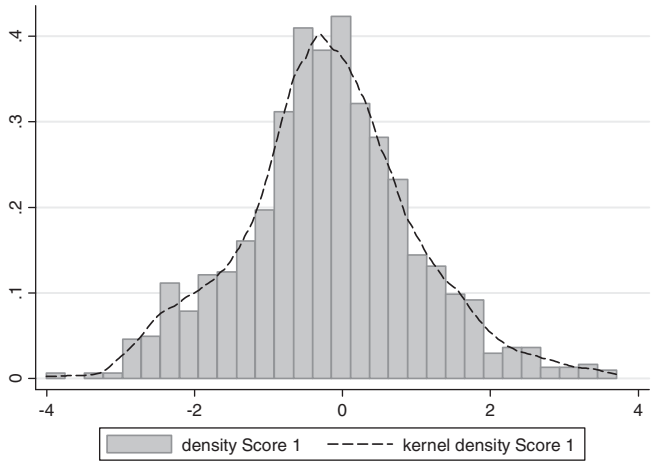

Figure 1. *Is there a discontinuity in first-stage scores around the cutoff?*

who were rejected in the initial application stage. In the end, out of the full population of 907 individuals for which we do have useful records from the application stage, in particular including a phase 1 score, we were able to obtain post-program information for 655 of them, that is for 72.2% of the relevant population.[3] To investigate further whether non-response is indeed a concern, we compare pre-program characteristics of those for which we have obtained post-program data and those for which we were not able to obtain data through our own interview attempts. More specifically, we perform two-sided *t*-tests of the difference in the baseline characteristics that are available to us for all individuals who submitted an application, namely the phase 1 score (the score that was assigned to the initial application to the program), and demographics that were available on the application form, namely age and gender. With respect to the baseline characteristics listed on the application, we do not find statistically significant differences between those for which we have post-program data and those for which we do not have post-program data, making us confident that non-response is not a major concern.

Unfortunately, we have incomplete data on business size (employment and sales) for businesses before the program. After the program we have this information only if the business is a new start-up or if it significantly expanded its business. Therefore, our analysis will focus on the variables that summarize the key goal of the business plan competitions, namely that individuals become more entrepreneurial, that is we focus on starting a business or significantly expanding a business (which are both measured as binary variables). These also have the advantage that they are less prone to measurement error than outcomes such as profits (see de Mel, McKenzie, & Woodruff, 2009). The mean number of employees of businesses after the program is about 10. Therefore, the businesses that we study are of a size well beyond household enterprises, which are somewhat more studied.

(a) *Did the discontinuity hold?*

Our methodology presumes a discontinuity in program participation. Therefore, before proceeding, we first investigate whether program participation was following a strict rule, that is whether there is indeed a discontinuity in program participation based on the score that an individual received in the application process. We standardize the scores of phase 1, 2, and 3 by subtracting the cutoff and dividing by the standard deviation of each individual competition, such that the cutoff is always equal to 0 and scores have the same variance. In Figure 1 we show a histogram and the kernel density of phase 1 scores. This figure shows that scores just above the cutoff are not more likely, thus, there is no evidence that individuals were moved up so that they just fall above a threshold score.[4]

The second question is whether indeed individuals above a cutoff are more likely to be trained than those below the cutoff. To gain insights into this, consider the histograms in Figure 2 (overlaid with kernel density estimates). On the left side we have individuals that were not trained, and on the right there are those that were trained. We clearly see that the cutoff was obeyed to almost perfectly. Thus, there is a significant discontinuity in the raw training data: Applicants below a threshold are much less likely (highly significant in regressions not shown) to enter the program than applicants above a threshold.

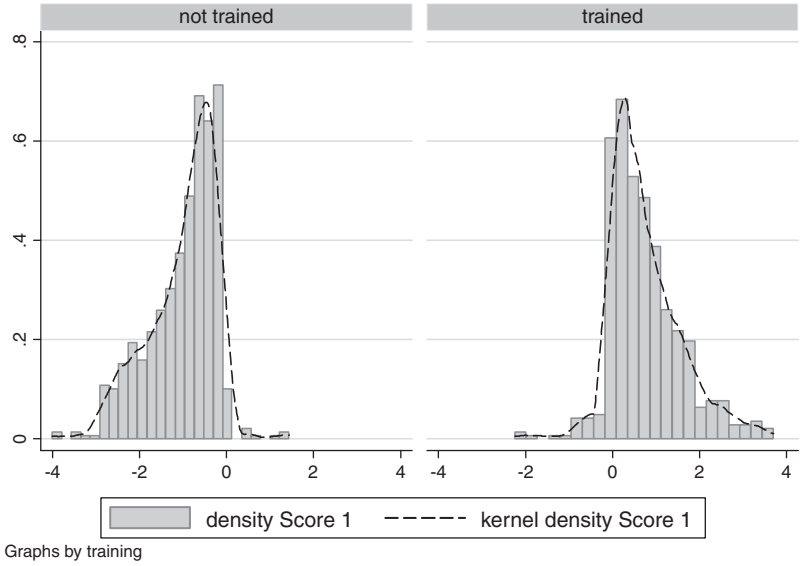

Figure 2. *Is there a discontinuity in training participation?*

Similarly, we confirm that the discontinuity exists at the stage where trained individuals potentially can progress to the second stage of training. Finally, we investigate whether the discontinuity exists at the winner stage. It turns out that at the winner stage, the cutoff was strictly enforced, that is at this stage the discontinuity is sharp. This is no surprise given what is immediately at stake, namely prize money of at least US$6,000 for the winners. So there is no concern about accepted applicants not taking the prize, and apparently there was also no one receiving a prize who did not make the cutoff. Prizes were awarded to individuals with the highest scores in the evaluation at the last stage of the competition.

4. METHODOLOGY: REGRESSION DISCONTINUITY DESIGN

To establish causality, we resort to a quasi-experimental approach, namely a regression discontinuity research design.[5] Our setup is somewhat non-standard because it involves a number of sequential stages, with each involving a separate discontinuity that we will exploit. To illustrate the econometric approach, however, we will first discuss the standard setup with one discontinuity. Consider the following structural relationship between training participation and outcomes:

$$y_i = \alpha + \gamma \cdot 1(\textit{training participant})_i + \delta \overline{X}_i + u_i \quad (1)$$

Here, $1(\textit{training participant})$ is an indicator variable that is equal to one if the individual participated in the training program, and γ is the parameter of interest. $\overline{X}_i$ is a vector of observable characteristics. However, as is well known, the difficulty with getting at causality is that in non-experimental settings training participation, $1(\textit{training participant})$, is most likely correlated with unobserved characteristics u_i.[6]

The fact that applicants are assigned a score, based on which program participation eligibility is determined, can be exploited for a regression discontinuity design. We have demonstrated that indeed a significant discontinuity exists in the applicants' scores at the end of phase 1; applicants with a score above a certain cutoff have a significantly higher probability of getting into the program. We exploit this for identification of the effect of the training. Similarly, a discontinuity exists at the scores for the second and third round. We use these to estimate the effect of additional training, and the effect of the prize money that the "winners" of the business plan competitions receive.

In a regression discontinuity research design the general idea is that if for individuals around the cutoff unobservable characteristics do not vary discontinuously, then program participation can be thought of as essentially randomly assigned to individuals around the cutoff. This can be exploited by estimating the baseline regression discontinuity equation (using the discontinuity created by phase 1 score $S1$ as an example to illustrate the approach):

$$y_{i,after} = \alpha + \gamma \cdot 1(S1_i \geqslant \overline{S1}) + \delta \overline{X}_i + f(S1_i) + u_i \quad (2)$$

where $y_{i,after}$ is an outcome of interest for applicant i, $S1$ is applicant i's phase 1 score, $\overline{S1}$ is the cutoff score for being admitted into the training program, $1(S1_i \geqslant \overline{S1})$ is an indicator function which is equal to one if applicant i's phase 1 score is above the cutoff, $f(\mathrm{S1_i})$ is a polynomial in applicant i's phase 1 score, and $\overline{X}_i$ is a vector of other exogenous controls.[7] The parameter γ in this case would be an estimate of the average treatment effect. $f(S1_i)$ is included because the regression discontinuity design relies on an assumption that holds only in the limit, and, therefore, in practical terms holds only in a small neighborhood around the cutoff. Especially if the variable that was used to assign treatment (here the score $S1$) is related to outcomes, using all available observations and thus increasing the interval around the cutoff score that is used for estimation will likely introduce a bias in γ. Because we know the value of the variable $S1$ hat was used to assign treatment we can include this in our regression framework (which controls for the effect of this variable), and increase the area around the cutoff from which we draw observations, thus increasing the number of observations available for estimation purposes. While there is in general no definite guidance about the window around the cutoff which should be used in linear regressions, nor about the polynomial $f(\mathrm{S1_i})$ that should be used if the whole range of data is used, Lee and Lemieux (2010, p. 318) suggest that "[...] it is essential to explore how RD estimates are robust to the inclusion of higher order polynomial terms [...] and to changes in the window width around the cutoff point [...]." Accordingly, we will perform the relevant robustness checks in our analyses below.

The procedures described above rely on a "sharp" discontinuity, in which program participation is perfectly predicted by the program score, thus $1(S1_i \geqslant \overline{S1}) = 1(\textit{training participant})_i$. However, more frequently some individuals above the cutoff will not participate in the program, while some below will participate. We have shown above (in Figure 2) that this is true for the present data. This is what is termed a "fuzzy" regression discontinuity design. Here, the regression of interest is

$$y_{i,\mathrm{after}} = \alpha + \gamma \cdot T_i + \delta \overline{X}_i + f(S1_i) + u_i \quad (3)$$

where T is an indicator function that is equal to one if an individual participated in the training, and zero otherwise. The problem in the case of a fuzzy design is that the characteristics that determine whether the individual participates in the program or not are unobserved to the econometrician but likely there is a correlation between those unobservables and outcomes of interest, introducing a correlation between T and u_i. Thus, estimating Eqn (3) with OLS no longer gives consistent estimates. To deal with this, in a fuzzy regression discontinuity design the existence of the cutoff score can be exploited as an instrument for program participation.[8]

To show robustness, we also estimate the average intent-to-treat effect γ^{ITT} with the following specification:

$$y_{i,\mathrm{after}} = \alpha + \gamma^{ITT} \cdot 1(S1_i \geqslant \overline{S1}) + \delta \overline{X}_i + f(S1_i) + u_i \quad (4)$$

In our setup there is one non-standard feature, though: the treatment is not homogenous, with treatment determined by a sequence of assignment variables (i.e., the scores assigned after three different phases of the training program).[9] The treatment heterogeneity results from the fact that out of the set of trained individuals only a subset is selected based on further scores (score 2, *S2*, and score 3, *S3*) after the first round of training. It is important to keep this in mind when we interpret the results. Given our setup it is not possible to estimate the effect of training separately from the effect of the financial treatment. Instead, we identify three different treatment effects: (1) We first estimate the effect of being in the first training program; this includes some individuals who go on to the second stage of the training program and a few that eventually win the competition and receive a monetary prize. (2) We can then estimate the effect of the additional training, conditional on having been in the first training. (3) We can estimate the effect of winning the competition, conditional on having been in the first training *and* having participated in the second round of training.

We do have non-perfect discontinuities at $\overline{S1}$ and $\overline{S2}$, that is a "*fuzzy*" research design. Therefore, we need to instrument T_i and $T2_i$ which we do with the indicator variables $1(S1_i \geqslant \overline{S1})$ and $1(S2_i \geqslant \overline{S2})$, respectively. However, we do have a sharp discontinuity in the last stage and, therefore, we do not need to instrument for W, the winning indicator. We generally include linear and squared terms of the score variables $S1$, $S2$, and $S3$. We also run specifications in which we interact the functions $f(S1_i)$, $f'(S2_i)$, and $f''(S3_i)$ with the indicator variables T, $T2$, and W, respectively, to allow for the effect of the various stages of the business plan competition to vary with the scores. Note that in this latter case we can only identify $\gamma(\overline{S})$, the local treatment effect at the cutoff $\overline{S}$.

(a) *Are outcomes smooth around cutoff before the program?*

The identification assumption is that $E[u|S]$ is continuous at the score S. This implies that outcomes in the absence of an intervention are smooth around the cutoff (e.g., Lee & Lemieux, 2010, p. 296). To support this assumption, we will first check outcomes before the individuals participated in the training program. Consider the distribution of applicants with and without a business. Figure 3 shows results from two locally weighted regressions of an indicator variable that is equal to one if an individual owned a business before the program on the score $S1$: one locally weighted regression for individuals with scores below the cutoff, and one with scores above the

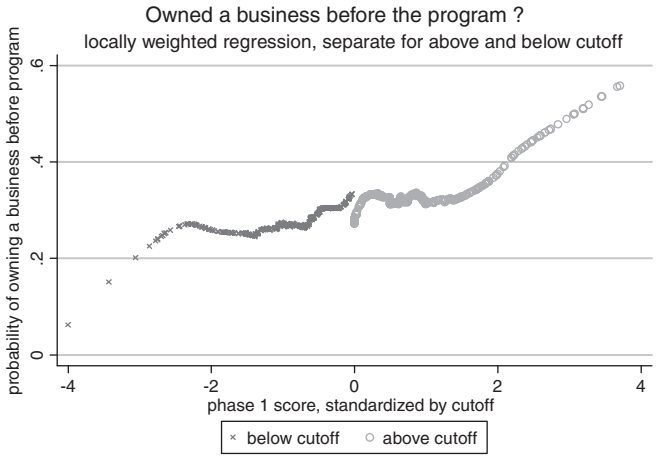

Figure 3. *Is there a discontinuity before the program?*

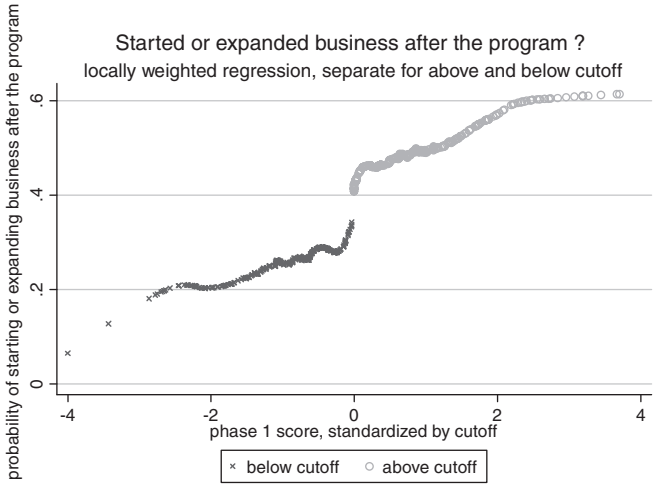

Figure 4. *Is there a discontinuity after the program?*

Table 1. *Are outcomes smooth around the cutoff before the program?*

Dependent variable =	All observations				Restrict window to <1 std. dev.			
	Existing business?	Sales before	Employees before	Age	Existing business?	Sales before	Employees before	Age
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intent to treat 1 (ITT 1)	0.072	−58.162	0.092	−1.031	0.038	−95.185	−4.499	−0.252
	(0.054)	(41.801)	(2.996)	(1.177)	(0.085)	(70.291)	(4.155)	(1.795)
$S1$ (phase 1score)	0.052	63.227	1.518	0.913	0.135	130.409	10.512	−2.975
	(0.036)	(33.810)*	(2.092)	(0.653)	(0.129)	(116.134)	(6.417)	(2.552)
ITT 1 * $S1$	−0.007	−53.667	0.872	−0.537	−0.091	−135.929	−10.222	6.747
	(0.047)	(39.371)	(2.655)	(0.970)	(0.165)	(143.012)	(8.991)	(3.435)**
Nicaragua	0.012	−126.201	3.719	−0.300	−0.007	−132.177	2.836	0.153
	(0.057)	(40.949)***	(2.786)	(1.305)	(0.066)	(51.202)**	(2.844)	(1.458)
El Salvador	0.088	−139.763	−7.970	−4.126	0.146	−141.138	−7.371	−3.704
	(0.084)	(54.879)**	(3.693)**	(1.534)***	(0.095)	(62.770)**	(3.524)**	(1.751)**
Year = 2003	0.170	43.097	14.240	−0.003	0.073	35.825	0.000	−0.785
	(0.070)**	(67.270)	(6.537)**	(1.655)	(0.093)	(83.628)	(0.000)	(2.146)
Year = 2004	0.265	−19.591	0.000	−3.389	0.258	−16.971	−11.966	−3.893
	(0.120)**	(90.257)	(0.000)	(2.579)	(0.145)*	(112.077)	(6.797)*	(3.097)
Year = 2005	0.389	−15.837	9.109	−1.768	0.347	−23.267	−5.610	−1.587
	(0.100)***	(76.265)	(3.133)***	(2.039)	(0.120)***	(90.899)	(5.802)	(2.484)
Constant	0.006	219.141	0.099	39.823	0.083	265.117	18.958	37.759
	(0.112)	(86.906)**	(4.081)	(2.267)***	(0.144)	(112.556)**	(6.799)***	(2.921)***
Observations	746	228	195	907	497	163	135	567
R-squared	0.05	0.08	0.09	0.01	0.04	0.09	0.09	0.02

Notes: Standard errors in parentheses; omitted country is Guatemala, omitted year is 2002; for further explanation see Section 4.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

cutoff. This figure illustrates that there is no visible jump in business ownership at the discontinuity before the program.

We can also perform the analysis more formally, namely through a regression of an indicator variable that is equal to one if an individual owned a business before the program on an indicator that is one if an individual's score falls above the program's cutoff $1(S1_i \geqslant \overline{S1})$, and the score itself. In an OLS regression in which we further allow the correlation of the score with the various outcomes to be different below and above the discontinuity, and also include country and year fixed effects, the indicator variable is highly insignificant. Thus, there is no indication of a pre-program discontinuity in businesses owned around the cutoff. In a similar fashion we investigate sales and employment of existing businesses in the year before the program. Again we find no discontinuity around the cutoff. As a check for discontinuities in demographic variables, we also run the same regression with age as the dependent variable, and find no discontinuity. These results are summarized in Table 1. In sum, this section provides strong evidence for the absence of a discontinuity before the program starts, and thus is suggestive evidence that is consistent with the underlying smoothness assumption around the cutoff.

(b) *Discontinuities around cutoff after the program?*

In the previous section we have checked the assumption of the regression discontinuity design that there are no discontinuities at the cutoff in the absence of the intervention. The second step of this research design then looks for a break around the discontinuity after the program. Thus, as the next step of our analysis we first simply provide a graphical analysis, that is plot outcomes against the program score, without controlling

Table 2. *The effect of training on business launch or expansion*

	Dependent variable = 1 if new business launched or existing business expanded							
	Assignment to treatment				Received training			
	Baseline	Restrict window to <1 std. dev.	Restrict window to <1 std. dev., treatment varies with score	Include 4th order polynomial	Baseline	Restrict window to <1 std. dev.	Restrict window to <1 std. dev., treatment varies with score	Include 4th order polynomial
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intent to treat 1 (ITT 1)	0.168 (0.057)***	0.188 (0.086)**	0.221 (0.094)**	0.180 (0.067)***				
Received training (T)					0.250 (0.056)***	0.285 (0.076)***	0.355 (0.089)***	0.280 (0.063)***
$S1$ (phase 1 score)	0.041 (0.025)	0.017 (0.082)	−0.088 (0.148)	0.032 (0.044)	0.023 (0.024)	−0.033 (0.070)	−0.220 (0.142)	−0.003 (0.040)
ITT 1 * $S1$			0.154 (0.180)					
$T * S1$							0.251 (0.165)	
$S1^2$				0.023 (0.019)				0.029 (0.019)
$S1^3$				0.001 (0.005)				0.004 (0.005)
$S1^4$				−0.002 (0.002)				−0.003 (0.002)
Male	−0.003 (0.043)	0.058 (0.056)	0.056 (0.056)	−0.001 (0.043)	−0.011 (0.042)	0.050 (0.056)	0.047 (0.056)	−0.007 (0.042)
Age	0.032 (0.010)***	0.042 (0.014)***	0.042 (0.014)***	0.033 (0.010)***	0.032 (0.010)***	0.041 (0.013)***	0.042 (0.013)***	0.032 (0.010)***
Age^2	−0.000 (0.000)***	−0.000 (0.000)***	−0.000 (0.000)***	−0.000 (0.000)***	−0.000 (0.000)***	−0.000 (0.000)***	−0.000 (0.000)***	−0.000 (0.000)***
Nicaragua	−0.018 (0.061)	−0.052 (0.071)	−0.042 (0.072)	−0.006 (0.062)	−0.031 (0.061)	−0.064 (0.070)	−0.048 (0.071)	−0.017 (0.061)
El Salvador	0.084 (0.076)	0.157 (0.089)*	0.157 (0.089)*	0.098 (0.078)	0.052 (0.076)	0.120 (0.089)	0.134 (0.090)	0.072 (0.077)
Year = 2003	0.032 (0.071)	−0.029 (0.095)	−0.031 (0.095)	0.037 (0.072)	0.039 (0.071)	−0.009 (0.094)	−0.013 (0.094)	0.045 (0.071)
Year = 2004	0.141 (0.119)	0.203 (0.147)	0.192 (0.147)	0.146 (0.119)	0.139 (0.117)	0.205 (0.145)	0.202 (0.145)	0.147 (0.118)
Year = 2005	−0.060 (0.093)	−0.012 (0.116)	−0.015 (0.116)	−0.046 (0.094)	−0.054 (0.092)	−0.002 (0.115)	0.008 (0.115)	−0.037 (0.093)
Constant	−0.346 (0.221)	−0.604 (0.293)**	−0.659 (0.300)**	−0.402 (0.227)*	−0.368 (0.218)*	−0.634 (0.289)**	−0.752 (0.299)**	−0.444 (0.223)**
Observations	655	435	435	655	655	435	435	655
R-squared	0.14	0.11	0.12	0.14	0.16	0.13	0.14	0.16

Notes: Standard errors in parentheses; omitted country is Guatemala, omitted year is 2002; for further explanation see Section 4.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

Table 3. *The effect of training on business launch or expansion: Instrumental variables results and robustness to different windows around the cutoff*

	Dependent variable = 1 if new business launched or existing business expanded							
	Instrumental variables				Alternative windows around cutoff			
							Instrumental variables	
	Baseline	Restrict window to <1 std. dev.	Restrict window to <1 std. dev., treatment varies with score	Include 4th order polynomial	Restrict window to <2 std. dev.	Restrict window to <0.5 std. dev.	Restrict window to <2 std. dev.; Instr. Variables	Restrict window to <0.5 std. dev.; Instr. Variables
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Received training (T)	0.192 (0.065)***	0.219 (0.099)**	0.259 (0.110)**	0.208 (0.077)***	0.248 (0.061)***	0.334 (0.096)***	0.184 (0.074)**	0.209 (0.127)*
$S1$ (phase 1 score)	0.038 (0.026)	0.013 (0.083)	−0.121 (0.158)	0.028 (0.045)	0.020 (0.035)	0.001 (0.156)	0.045 (0.038)	0.150 (0.185)
$T * S1$			0.192 (0.185)					
$S1^2$				0.026 (0.019)				
$S1^3$				0.001 (0.005)				
$S1^4$				−0.002 (0.002)				
Male	−0.008 (0.042)	0.053 (0.056)	0.050 (0.056)	−0.005 (0.043)	−0.001 (0.045)	0.091 (0.074)	0.001 (0.045)	0.099 (0.074)
Age	0.032 (0.010)***	0.041 (0.013)***	0.042 (0.013)***	0.032 (0.010)***	0.041 (0.011)***	0.051 (0.019)***	0.041 (0.011)***	0.051 (0.019)***
Age^2	−0.000 (0.000)***	−0.000 (0.000)***	−0.000 (0.000)***	−0.000 (0.000)***	−0.000 (0.000)***	−0.001 (0.000)**	−0.000 (0.000)***	−0.001 (0.000)**
Nicaragua	−0.024 (0.061)	−0.056 (0.071)	−0.043 (0.072)	−0.012 (0.062)	−0.005 (0.062)	−0.042 (0.088)	0.001 (0.063)	−0.023 (0.089)
El Salvador	0.065 (0.077)	0.131 (0.090)	0.144 (0.091)	0.080 (0.077)	0.082 (0.079)	0.173 (0.118)	0.092 (0.079)	0.190 (0.119)
Year = 2003	0.037 (0.071)	−0.015 (0.095)	−0.019 (0.095)	0.043 (0.071)	0.041 (0.075)	−0.149 (0.117)	0.040 (0.075)	−0.155 (0.118)
Year = 2004	0.140 (0.118)	0.203 (0.145)	0.201 (0.145)	0.146 (0.118)	0.141 (0.123)	0.134 (0.179)	0.143 (0.123)	0.127 (0.180)
Year = 2005	−0.058 (0.092)	−0.008 (0.115)	−0.001 (0.115)	−0.042 (0.093)	−0.032 (0.096)	−0.016 (0.148)	−0.034 (0.096)	−0.030 (0.148)
Constant	−0.342 (0.219)	−0.606 (0.291)**	−0.691 (0.303)**	−0.405 (0.225)*	−0.573 (0.235)**	−0.875 (0.382)**	−0.545 (0.236)**	−0.811 (0.385)**
Observations	655	435	435	655	591	260	591	260
R-squared	0.16	0.13	0.14	0.16	0.15	0.18	0.15	0.17

Notes: Standard errors in parentheses; omitted country is Guatemala, omitted year is 2002; instruments are intent to treat indicator (in columns 1, 2, 3, 4, 7 and 8) and intent to treat indicator plus intent to treat indicator interacted with score 1 (column 3); for further explanation see Section 4.

* Significant at 10%.

** Significant at 5%.

*** Significant at 1%.

for any other variables. Specifically, we look at whether individuals started a business, or expanded a business and consider locally weighted regression (Figure 4). Using separate locally weighted regressions for observations below and above the cutoff, we note a pronounced change in the outcome variable around the cutoff, suggesting a jump at the discontinuity. Of course, a more refined version of this kind of empirical analysis will be performed using the regression discontinuity design, which is reported in the next section.

5. RESULTS

This section presents the main results. We estimate linear probability models, and the instrumental variables regressions are two-stage least squares regressions. All regressions include country and year fixed effects. We use the following variable names: T is an indicator variable that is equal to one if an aspiring entrepreneur was trained. Similarly, $T2$ is an indicator which is equal to one if the individual received the additional training in training stage 2. While W is an indicator that is equal to one if the individual is a winner of the business plan competition. "Intent to treat 1" (or short "ITT1") is equal to one if $S1 \geqslant \overline{S1}$; similarly, "Intent to treat 2" (short "ITT2") is equal to one if $S2 \geqslant \overline{S2}$. No such distinction is necessary in the final cutoff stage of the competition, as in the third (winning) stage there is a sharp discontinuity, and everyone above the threshold is actually winning the prize money.

Throughout, we investigate the effect of the program on two types of outcomes related to entrepreneurial activity. We look at the start-up of new businesses and the expansion of existing businesses. To this end, we have to restrict our samples as follows: In the analysis in which start-up is the outcome of interest, we restrict the sample to individuals that did not own a

Table 4. *The effect of training on business launch or expansion, separate regressions*

	All instrumental variables regressions							
	Dependent variable = 1 if new business launched				Dependent variable = 1 if existing business expanded			
	Baseline	Restrict window to <1 std. dev.	Restrict window to <1 std. dev., treatment varies with score	Include 4th order polynomial	Baseline	Restrict window to <1 std. dev.	Restrict window to <1 std. dev., treatment varies with score	Include 4th order polynomial
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Received training (T)	0.094	0.041	0.037	0.044	0.249	0.424	0.556	0.284
	(0.073)	(0.110)	(0.121)	(0.085)	(0.091)***	(0.135)***	(0.145)***	(0.123)**
$S1$ (phase 1 score)	0.035	0.075	0.088	0.088	0.018	−0.189	−0.688	−0.014
	(0.029)	(0.094)	(0.174)	(0.051)*	(0.035)	(0.118)	(0.216)***	(0.089)
$T * S1$			−0.019				0.728	
			(0.207)				(0.251)***	
$S1^2$				0.024				0.015
				(0.022)				(0.037)
$S1^3$				−0.008				0.005
				(0.006)				(0.015)
$S1^4$				−0.003				−0.002
				(0.002)				(0.005)
Male	0.011	0.117	0.118	0.013	−0.049	−0.092	−0.091	−0.044
	(0.048)	(0.064)*	(0.064)*	(0.048)	(0.063)	(0.083)	(0.081)	(0.064)
Age	0.024	0.031	0.031	0.025	−0.007	0.004	0.004	−0.007
	(0.011)**	(0.015)**	(0.015)**	(0.011)**	(0.015)	(0.021)	(0.020)	(0.016)
Age^2	−0.000	−0.000	−0.000	−0.000	0.000	−0.000	−0.000	0.000
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)	(0.000)	(0.000)	(0.000)
Nicaragua	−0.031	−0.048	−0.049	−0.009	−0.045	−0.125	−0.063	−0.040
	(0.071)	(0.081)	(0.082)	(0.073)	(0.084)	(0.097)	(0.097)	(0.085)
El Salvador	−0.101	−0.047	−0.047	−0.103	−0.556	−0.576	−0.515	−0.541
	(0.105)	(0.129)	(0.130)	(0.107)	(0.116)***	(0.127)***	(0.126)***	(0.119)***
Year = 2003	−0.111	−0.112	−0.112	−0.093	−0.057	−0.070	−0.095	−0.058
	(0.076)	(0.099)	(0.100)	(0.077)	(0.125)	(0.154)	(0.151)	(0.128)
Year = 2004	−0.186	−0.113	−0.113	−0.195	−0.508	−0.470	−0.471	−0.499
	(0.149)	(0.182)	(0.183)	(0.150)	(0.182)***	(0.217)**	(0.213)**	(0.185)***
Year = 2005	−0.316	−0.252	−0.252	−0.304	−0.854	−0.882	−0.848	−0.842
	(0.122)**	(0.153)	(0.154)	(0.124)**	(0.153)***	(0.182)***	(0.179)***	(0.156)***
Constant	−0.005	−0.226	−0.218	−0.018	1.469	1.322	1.041	1.425
	(0.247)	(0.334)	(0.346)	(0.255)	(0.362)***	(0.454)***	(0.457)**	(0.376)***
Observations	392	252	252	392	249	171	171	249
R-squared	0.13	0.11	0.11	0.13	0.35	0.36	0.39	0.35

Notes: Standard errors in parentheses; omitted country is Guatemala, omitted year is 2002; instruments are intent to treat indicator (in columns 1, 2, 4, 5, 6 and 8) and intent to treat indicator plus intent to treat indicator interacted with score 1 (in columns 3 and 7); for further explanation see Section 4.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

business before the program. Analogously, in the analysis of expansion, we restrict the sample to entrepreneurs that already had a business before the training program. To be able to use the full sample, which increases the precision of our estimates, we also consider an outcome variable in which we pool both outcomes. This variable is equal to one if an individual with an existing business significantly expanded her business, or, for individuals without a business at the time of application for the program, launched a new business.

(a) *Total effect of the training program*

In this part we exploit the discontinuity that is created by the application procedure that interested applicants have to undergo. Thus, we will estimate the total effect of the training program, not just the training in the first stage, and the estimated effect will include the effect of, potentially, going on to later stages of the program. [10]

Our main result is shown in Tables 2 and 3: Here, the dependent variable is equal to one if an individual with an existing business significantly expanded her business, or, for individuals without a business at the time of application for the program, launched a new business. In all four specifications of column 1–4 of Table 2 we find strong effects of the business plan competition program on entrepreneurial activity, that is business start-up/launch or expansion of existing businesses. The parameter estimates imply a 17 to 22 percentage points increase in the probability of opening or expanding a business as a consequence of the program. Column 1 shows the baseline results. In column 2 the analysis is restricted to include only data from individuals who scored within a window of one standard deviation of the phase 1 score around the cutoff. The results are qualitatively unchanged. Column 3 shows that the results are robust to including an interaction term of the intent to treat variable with the score, that is if we let the effect of the program vary with the score. The results are also robust to inclusion of a fourth

order polynomial in the phase 1 score (column 4). Columns 5–8 show analogous results for the variable "received training" as opposed to the intent to treat variable.

In columns 1–4 of Table 3 we take into account that the design is fuzzy in the first stage of the training program, that is that not everyone who was assigned to training actually participated in the program, and vice versa. These columns show the instrumental variables (IV) results. Again we find effects of the program of the same order of magnitude. Finally, columns 5–8 of Table 3 show that the results are robust to using different windows around the cutoff, namely a window of two standard deviations, and a window of 0.5 standard deviations (both for the OLS specification, columns 5 and 6, and the instrumental variables specification in columns 7 and 8).

Looking at launch and expansion separately (Table 4), we find strong and significant effects on business expansion, while we do not find a statistically significant effect of the program on launching of businesses. The (statistically insignificant) parameter estimates imply a four to nine percentage points higher probability of opening a business (for individuals without a business before the start of the program) in the treatment group, and a (statistically significant) increase of 25 to 56 percentage points in the probability of expanding a business (for individuals with an existing business before the program) in the treatment group.

Regarding other variables of interest, we note that there is no consistent picture regarding the difference between male and female applicants in the baseline probability that a business is launched or expanded.[11] On the other hand, we find strong age effects for launching a business: older individuals are much more likely to launch a business than younger individuals.

(b) *Effect of second stage training, conditional on first stage training*

Next, we study the effect of entering the second stage training on business outcomes, conditional on the first training stage. Again, we first look at the pooled outcome variable, which is equal to one if either an existing business was expanded or a new business was launched. Table 5 shows the results. We restrict our analysis to the instrumental variables

Table 5. *The effect of the second stage of training on business launch or expansion*

	Dependent variable = 1 if new business launched or existing business expanded All instrumental variables regressions			
	Baseline	Restrict window to <1 std. dev.	Restrict window to <1 std. dev., treatment varies with score	Include 4th order polynomial
	(1)	(2)	(3)	(4)
Received training	0.167	0.394	0.416	0.248
($T2$)	(0.098)*	(0.143)***	(0.144)***	(0.117)**
$S2$ (phase 2 score)	0.023	−0.210	−0.386	−0.044
	(0.039)	(0.126)*	(0.183)**	(0.067)
$T2 * S2$			0.340	
			(0.244)	
$S2^2$				−0.000
				(0.029)
$S2^3$				0.008
				(0.007)
$S2^4$				−0.001
				(0.003)
Male	0.008	−0.095	−0.101	0.015
	(0.065)	(0.084)	(0.084)	(0.065)
Age	0.057	0.073	0.072	0.060
	(0.015)***	(0.019)***	(0.019)***	(0.015)***
Age^2	−0.001	−0.001	−0.001	−0.001
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Nicaragua	0.042	0.076	0.064	0.045
	(0.098)	(0.122)	(0.122)	(0.098)
El Salvador	0.379	0.355	0.341	0.381
	(0.108)***	(0.148)**	(0.148)**	(0.108)***
Year = 2003	0.025	−0.002	0.023	0.038
	(0.090)	(0.127)	(0.127)	(0.091)
Year = 2004	0.354	0.360	0.375	0.356
	(0.154)**	(0.192)*	(0.191)*	(0.155)**
Year = 2005	0.060	0.007	0.006	0.062
	(0.112)	(0.156)	(0.155)	(0.113)
Constant	−0.927	−1.218	−1.271	−1.028
	(0.336)***	(0.441)***	(0.441)***	(0.345)***
Observations	272	156	156	272
R-squared	0.18	0.22	0.23	0.19

Notes: Standard errors in parentheses; omitted country is Guatemala, omitted year is 2002; instruments are intent to treat indicator (in columns 1, 2, and 4) and intent to treat indicator plus intent to treat indicator interacted with score 2 (column 3); for further explanation see Section 4.

* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

estimates, which estimate the effect of participating in the program (as opposed to the intent to treat effect), taking into account the fact that some are assigned to treatment but do not participate. Overall, we find a strong effect of the second stage training on this combined outcome variable. Looking at launch and expansion separately (Table 6), we see that the significant combined effect is mainly due to a significant effect of this second stage of the training program on launching a new business, rather than on expanding an existing business.

(c) *Effect of winning, conditional on second stage training*

The final winning stage is different for two reasons. First, as we have seen above, the discontinuity at this stage is sharp, therefore, we do not have to instrument the indicator variable of interest (W, which is equal to one if the individual won the competition). Second, we also argue that at the last stage individuals are fairly similar to each other in terms of their unobservables, because there are only a few of them, which all have gone through two rounds of selections already, and that consequently they are all "close" to the cutoff, which would allow us to exclude $f''(S3_i)$. This matters in practical terms, because at this last stage we have relatively few observations, so increasing the degrees of freedom matters for our econometric analysis. Thus, our preferred specification does not condition on phase 3 score (Greenstone, Hornbeck, & Moretti, 2010, use a similar approach in the context of location decisions of industrial plants). However, we also show specifications that include the score. In addition, unlike before, in this stage it is also easier to name the treatment more specifically. In the earlier training stages the training is composed of several parts and it is not possible to identify separately which training component or which group of components is responsible for results. Here, however, the most significant part of the additional treatment is the prize money.

We find economically significant changes in the probability of starting or expanding a business due to winning the competition (Table 7), and the coefficient on the *winner* variable is

Table 6. *The effect of the second stage of training on business launch or expansion, separate regressions*

	All instrumental variables regressions							
	Dependent variable = 1 if new business launched				Dependent variable = 1 if existing business expanded			
	Baseline	Restrict window to <1 std. dev.	Restrict window to <1 std. dev., treatment varies with score	Include 4th order polynomial	Baseline	Restrict window to <1 std. dev.	Restrict window to <1 std. dev., treatment varies with score	Include 4th order polynomial
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Received training (T2)	0.425 (0.139)***	0.724 (0.215)***	0.726 (0.213)***	0.465 (0.184)**	0.068 (0.121)	0.222 (0.162)	0.201 (0.166)	0.087 (0.144)
S2 (phase 2 score)	−0.006 (0.055)	−0.315 (0.178)*	−0.591 (0.255)**	−0.025 (0.133)	−0.022 (0.047)	−0.215 (0.140)	−0.092 (0.213)	−0.034 (0.076)
T2 * S2			0.537 (0.340)				−0.230 (0.298)	
$S2^2$				0.013 (0.048)				0.022 (0.034)
$S2^3$				−0.004 (0.025)				0.002 (0.007)
$S2^4$				−0.004 (0.008)				−0.001 (0.003)
Male	0.049 (0.084)	−0.017 (0.119)	−0.023 (0.118)	0.052 (0.087)	0.032 (0.083)	0.076 (0.098)	0.085 (0.099)	0.021 (0.086)
Age	0.028 (0.019)	0.031 (0.024)	0.036 (0.024)	0.030 (0.019)	0.020 (0.021)	0.045 (0.029)	0.048 (0.029)	0.019 (0.021)
Age^2	−0.000 (0.000)	−0.000 (0.000)*	−0.001 (0.000)*	−0.000 (0.000)*	−0.000 (0.000)	−0.001 (0.000)	−0.001 (0.000)	−0.000 (0.000)
Nicaragua	−0.005 (0.131)	0.020 (0.158)	−0.014 (0.157)	−0.009 (0.133)	−0.003 (0.113)	0.090 (0.144)	0.085 (0.145)	0.002 (0.114)
El Salvador	0.196 (0.242)	0.315 (0.354)	0.196 (0.358)	0.189 (0.245)	−0.460 (0.131)***	−0.541 (0.185)***	−0.530 (0.186)***	−0.469 (0.134)***
Year = 2003	−0.203 (0.106)*	−0.269 (0.159)*	−0.205 (0.161)	−0.189 (0.108)*	0.025 (0.141)	0.046 (0.178)	0.029 (0.180)	0.012 (0.146)
Year = 2004	0.086 (0.279)	0.275 (0.384)	0.207 (0.382)	0.083 (0.282)	−0.439 (0.205)**	−0.596 (0.248)**	−0.605 (0.249)**	−0.450 (0.209)**
Year = 2005	−0.190 (0.241)	−0.121 (0.356)	−0.213 (0.357)	−0.192 (0.244)	−0.799 (0.168)***	−0.878 (0.219)***	−0.879 (0.220)***	−0.818 (0.173)***
Constant	−0.266 (0.453)	−0.458 (0.595)	−0.564 (0.593)	−0.315 (0.480)	0.998 (0.488)**	0.443 (0.683)	0.426 (0.687)	0.996 (0.505)*
Observations	150	89	89	150	108	58	58	108
R-squared	0.26	0.29	0.31	0.27	0.46	0.57	0.58	0.46

Notes: Standard errors in parentheses; omitted country is Guatemala, omitted year is 2002; instruments are intent to treat indicator (in columns 1, 2, 4, 5, 6 and 8) and intent to treat indicator plus intent to treat indicator interacted with score 1 (in columns 3 and 7); for further explanation see Section 4.

* Significant at 10%.

** Significant at 5%.

*** Significant at 1%.

Table 7. *The effect of winning the business plan competition on business launch or expansion*

	Dependent variable = 1 if new business launched or existing business expanded						
	Baseline	Include score	Restrict window to <1 std. dev.	Restrict window to <1 std. dev., treatment varies with score	Include 4th order polynomial	Treatment varies with score, incl. log(prize-money), restrict window	Treatment varies with score, incl. prize-money, restrict window
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Winner (W)	0.323	0.222	0.300	0.239	0.361		
	(0.087)***	(0.127)*	(0.172)*	(0.194)	(0.166)**		
$S3$ (phase 3 score)		0.061	0.065	0.222	−0.182	0.035	0.402
		(0.058)	(0.199)	(0.303)	(0.153)	(0.206)	(0.270)
Winner * $S3$				−0.258			
				(0.373)			
$S3^2$					0.002		
					(0.052)		
$S3^3$					0.105		
					(0.053)**		
$S3^4$					0.024		
					(0.011)**		
Log(prizemoney + 0.001)						0.024	
						(0.021)	
Log(prizemoney + 0.001) * S3						−0.030	
						(0.040)	
Prizemoney/1,000							0.011
							(0.015)
Prizemoney * $S3$							−0.041
							(0.034)
Male	0.058	0.065	−0.015	−0.009	0.053	−0.007	0.005
	(0.091)	(0.092)	(0.121)	(0.122)	(0.091)	(0.122)	(0.122)
Age	0.050	0.044	0.057	0.058	0.046	0.058	0.058
	(0.018)***	(0.019)**	(0.040)	(0.040)	(0.019)**	(0.040)	(0.040)
Age^2	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001
	(0.000)***	(0.000)**	(0.001)	(0.001)	(0.000)**	(0.001)	(0.001)
Nicaragua	0.188	0.211	0.373	0.355	0.233	0.353	0.340
	(0.132)	(0.133)	(0.165)**	(0.168)**	(0.133)*	(0.168)**	(0.168)**
El Salvador	0.137	0.150	0.360	0.367	0.170	0.375	0.408
	(0.175)	(0.177)	(0.237)	(0.239)	(0.175)	(0.239)	(0.239)*
Year = 2003	−0.176	−0.153	−0.233	−0.228	−0.170	−0.225	−0.212
	(0.124)	(0.127)	(0.164)	(0.165)	(0.126)	(0.165)	(0.164)
Year = 2004	−0.253	−0.252	−0.199	−0.165	−0.252	−0.150	−0.085
	(0.214)	(0.217)	(0.285)	(0.291)	(0.216)	(0.289)	(0.281)
Year = 2005	−0.472	−0.446	−0.385	−0.349	−0.439	−0.335	−0.268
	(0.172)***	(0.175)**	(0.233)	(0.240)	(0.174)**	(0.238)	(0.227)
Constant	−0.185	−0.008	−0.505	−0.448	−0.185	−0.280	−0.391
	(0.438)	(0.458)	(0.811)	(0.820)	(0.464)	(0.805)	(0.839)
Observations	109	108	56	56	108	56	56
R-squared	0.32	0.32	0.38	0.38	0.36	0.38	0.38

Notes: Standard errors in parentheses; omitted country is Guatemala, omitted year is 2002; for further explanation see Section 4.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

mostly statistically significant, and statistically highly significant in the preferred specification that does not include the score S3. [12] Looking at the launch and expansion of businesses separately (Table 8), we find that the point estimate in the preferred specification (columns 1 and 5) is larger for launching a business. However, the finding regarding the effect of the program at this last stage is less consistent across specifications that include the score. Looking at launch and expansion separately, we do find a statistically significant effect in the first specification for business launch. For business expansion there is a positive, though statistically insignificant, association between winning the competition and expanding the business in the preferred specification (column 5). The statistical imprecision and the lower robustness of the results at this stage of the competition may well be explained by the relatively small number of observations in this last stage of the program. [13]

(d) *Investigating country-specific effects*

In this section, we investigate the question whether the training has different effects across the three countries under consideration. Entrepreneurship and entrepreneurial skills might be at different stages in different countries, and, therefore, the effect of training might also differ. In addition, recall that in El Salvador, TechnoServe puts participants through UNCTAD's Empretec program, whereas in the other countries the

Table 8. *The effect of winning the business plan competition on business launch or expansion, separate regressions*

	All OLS regressions							
	Dependent variable = 1 if new business launched				Dependent variable = 1 if existing business expanded			
	Baseline	Restrict window to <1 std. dev.	Restrict window to <1 std. dev., treatment varies with score	Include 4th order polynomial	Baseline	Restrict window to <1 std. dev.	Restrict window to <1 std. dev., treatment varies with score	Include 4th order polynomial
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Winner (W)	0.339	0.161	0.159	0.028	0.184	0.537	0.788	0.350
	(0.126)***	(0.275)	(0.291)	(0.267)	(0.124)	(0.293)*	(0.404)*	(0.228)
$S3$ (phase 3 score)		0.266	0.273	0.474		−0.300	−0.825	−0.295
		(0.294)	(0.425)	(0.330)		(0.383)	(0.692)	(0.215)
Winner ∗ $S3$			−0.014				0.729	
			(0.616)				(0.798)	
$S3^2$				0.172				0.009
				(0.127)				(0.062)
$S3^3$				−0.237				0.107
				(0.176)				(0.061)*
$S3^4$				−0.115				0.022
				(0.065)*				(0.013)*
Male	0.044	−0.142	−0.142	0.069	0.073	0.055	0.125	0.004
	(0.128)	(0.189)	(0.196)	(0.126)	(0.147)	(0.238)	(0.252)	(0.155)
Age	0.036	0.040	0.041	0.033	0.053	0.031	0.029	0.016
	(0.023)	(0.055)	(0.058)	(0.024)	(0.040)	(0.112)	(0.113)	(0.046)
Age^2	−0.000	−0.001	−0.001	−0.000	−0.001	−0.000	−0.000	−0.000
	(0.000)*	(0.001)	(0.001)	(0.000)*	(0.001)	(0.002)	(0.002)	(0.001)
Nicaragua	0.171	0.379	0.380	0.217	−0.043	0.243	0.344	−0.021
	(0.204)	(0.317)	(0.329)	(0.202)	(0.180)	(0.319)	(0.339)	(0.199)
El Salvador	0.385	0.714	0.712	0.495	−0.606	−0.458	−0.363	−0.742
	(0.279)	(0.245)***	(0.260)**	(0.280)*	(0.277)**	(0.545)	(0.559)	(0.304)**
Year = 2003	−0.304	−0.206	−0.206	−0.268	0.015	−0.325	−0.395	−0.024
	(0.159)*	(0.207)	(0.213)	(0.158)*	(0.188)	(0.403)	(0.413)	(0.213)
Year = 2004	0.000	0.296	0.293	0.000	−0.671	−1.016	−1.052	−0.852
	(0.000)	(0.358)	(0.380)	(0.000)	(0.318)**	(0.494)*	(0.499)*	(0.330)**
Year = 2005	−0.329	0.000	0.000	−0.250	−0.916	−1.165	−1.140	−1.071
	(0.214)	(0.000)	(0.000)	(0.218)	(0.287)***	(0.466)**	(0.470)**	(0.299)***
Constant	−0.133	−0.340	−0.343	−0.010	0.587	0.615	0.234	1.296
	(0.577)	(1.067)	(1.101)	(0.624)	(0.858)	(2.324)	(2.375)	(1.015)
Observations	57	29	29	57	46	24	24	45
R-squared	0.37	0.43	0.43	0.45	0.50	0.58	0.61	0.54

Notes: Standard errors in parentheses; omitted country is Guatemala, omitted year is 2002; instruments are intent to treat indicator (in columns 1, 2, 4, 5, 6 and 8) and intent to treat indicator plus intent to treat indicator interacted with score 1 (in columns 3 and 7); for further explanation see Section 4.

* Significant at 10%.

** Significant at 5%.

*** Significant at 1%.

training is entirely conducted by TechnoServe. For this reason, we investigate the robustness of the results to allowing the effect of the training to vary by country, by interacting the training variables with a Nicaragua and an El Salvador indicator variable. The key results are reported in Table 9. Columns 1 and 2 of Table 9 suggest that the treatment effect of the full training program is significantly larger in El Salvador than in Guatemala, the omitted country, and has a larger point estimate in El Salvador than in Nicaragua. For the analogous analysis of the later rounds of the competition we do not find significantly different treatment effects, including for the difference between El Salvador and Nicaragua. The findings are consistent with a larger effect of the program in El Salvador. Because the setup of the program is slightly different in El Salvador (the only country where the Empretec program is used) it is unclear whether the difference in results is due to differences in the samples of entrepreneurs or whether the difference is due to program differences. In any case, this section confirms that the key findings are not simply due to pooling the data from these different countries.

(e) *The role of gender*

Empirical work as well as policy makers' decisions suggests that the effects of the program may vary by gender. In particular, in the policy world, micro finance organizations often target women, partly based on the idea that women may have less access to credit. On the other hand, recent empirical work by de Mel, McKenzie, and Woodruff (2007) suggests that there may be differences in the returns to capital between male and female entrepreneurs. [14]

In this section we, therefore, study whether treatment effects of the training program vary by the participant's gender. To this end, we interact the treatment indicator variables for the various stages of the business plan competition (training of first stage, training of second stage and winning) with the

Table 9. *The effect of the training on business launch or expansion: is El Salvador different?*

	Dependent variable = 1 if new business launched or existing business expanded				
	Effect of full training		Effect of second stage of training		Effect of winning
	Instrumental variables; interaction with El Salvador	Instrumental variables; interaction with El Salvador; restrict window to <1 std. dev.	Instrumental variables	Instrumental variables; restrict window to <1 std. dev.	OLS
	(1)	(2)	(3)	(4)	(5)
Received training	0.014	0.028			
	(0.083)	(0.118)			
T ∗ Nicaragua	0.233	0.210			
	(0.108)**	(0.126)*			
T ∗ El Salvador	0.338	0.421			
	(0.119)***	(0.146)***			
*S*1 (phase 1 score)	0.042	0.010			
	(0.025)*	(0.086)			
Received training2			0.149	0.091	
(*T*2)			(0.152)	(0.173)	
T2 ∗ Nicaragua			0.079	0.245	
			(0.171)	(0.200)	
*T*2 ∗ El Salvador			−0.049	−0.053	
			(0.160)	(0.188)	
*S*2 (phase 2 score)			0.033	0.055	
			(0.041)	(0.050)	
Winner					0.171
					(0.228)
Winner					0.125
∗ Nicaragua					(0.244)
Winner					−0.006
∗ El Salvador					(0.248)
*S*3 (phase 3 score)					0.062
					(0.059)
Male	−0.004	0.049	0.007	0.076	0.069
	(0.042)	(0.055)	(0.065)	(0.083)	(0.093)
Age	0.032	0.039	0.057	0.051	0.044
	(0.010)***	(0.013)***	(0.015)***	(0.020)**	(0.019)**
Age^2	−0.000	−0.000	−0.001	−0.001	−0.001
	(0.000)***	(0.000)***	(0.000)***	(0.000)**	(0.000)**
Nicaragua	−0.107	−0.138	−0.004	−0.140	0.158
	(0.081)	(0.091)	(0.122)	(0.142)	(0.189)
El Salvador	−0.078	−0.096	0.392	0.435	0.133
	(0.096)	(0.124)	(0.118)***	(0.126)***	(0.218)
Year = 2003	0.068	0.025	0.029	0.006	−0.154
	(0.072)	(0.095)	(0.090)	(0.120)	(0.128)
Year = 2004	0.194	0.267	0.356	0.399	−0.267
	(0.121)	(0.148)*	(0.154)**	(0.190)**	(0.219)
Year = 2005	0.005	0.057	0.056	0.120	−0.443
	(0.095)	(0.118)	(0.113)	(0.140)	(0.177)**
Constant	−0.368	−0.557	−0.903	−0.847	0.023
	(0.219)*	(0.288)*	(0.338)***	(0.415)**	(0.470)
Observations	655	435	272	196	108
R-squared	0.18	0.16	0.19	0.21	0.32

Notes: Standard errors in parentheses; omitted country is Guatemala, omitted year is 2002; instruments are the intent to treat indicator and intent to treat indicator interacted with El Salvador dummy variable (in columns 1–4); for further explanation see Section 4.

* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

indicator variable that is one if an individual is male, and zero otherwise. Because here the focus is on gender, we also include an indicator variable that is equal to one if the program under consideration was open only for women (the "all female program" variable in the table). At the final/winning stage there are only few observations. Thus, there is no variation in this variable any more once launch and expansion are treated separately, which is why it is omitted in columns 8 and 9.

The results indicate that the effect of the full training program on business start-up or expansion is larger for male participants. In the first three columns of Table 10 we see that the coefficient on the trained ∗ male interaction term is positive, and this coefficient is statistically significant in the last specification (column 3). Columns 4–6 of Table 10 then show that the more positive effect of training on males seems to be mainly due to difference in the effect of the first stage of train-

Table 10. *The effect of training on business launch or expansion: the interaction with gender*

Dependent variable = 1 if …	Effect of full training (IV)			Effect of second stage of training (IV)			Effect of winning		
	Launch or expand (1)	Launch (2)	Expand (3)	Launch or expand (4)	Launch (5)	Expand (6)	Launch or expand (7)	Launch (8)	Expand (9)
Received training	0.159	−0.019	0.230						
	(0.130)	(0.147)	(0.169)						
T ∗ male	0.091	0.090	0.294						
	(0.118)	(0.134)	(0.158)*						
$S1$ (phase 1 score)	0.005	0.067	−0.211						
	(0.084)	(0.095)	(0.120)*						
Received training2 ($T2$)				0.269	0.572	−0.012			
				(0.166)	(0.270)**	(0.214)			
$T2$ ∗ male				−0.100	−0.235	0.151			
				(0.168)	(0.252)	(0.214)			
$S2$ (phase 2 score)				0.027	0.036	−0.041			
				(0.047)	(0.066)	(0.060)			
Winner							0.557	0.546	0.316
							(0.160)***	(0.325)	(0.208)
Winner ∗ male							−0.318	−0.244	−0.191
							(0.184)*	(0.353)	(0.240)
Male	0.007	0.069	−0.260	0.146	0.195	0.069	0.169	0.083	0.164
	(0.091)	(0.103)	(0.122)**	(0.103)	(0.130)	(0.160)	(0.103)	(0.140)	(0.187)
All female program	0.092	0.095	−0.005	0.276	0.270	−0.013	0.508		
	(0.156)	(0.244)	(0.244)	(0.181)	(0.511)	(0.254)	(0.262)*		
Age	0.040	0.029	−0.001	0.050	0.017	0.016	0.052	0.038	0.048
	(0.014)***	(0.015)*	(0.021)	(0.020)**	(0.029)	(0.027)	(0.018)***	(0.024)	(0.040)
Age^2	−0.000	−0.000	−0.000	−0.001	−0.000	−0.000	−0.001	−0.001	−0.001
	(0.000)***	(0.000)*	(0.000)	(0.000)**	(0.000)	(0.000)	(0.000)***	(0.000)*	(0.001)
Nicaragua	−0.060	−0.055	−0.135	−0.009	−0.020	0.035	0.171	0.192	−0.072
	(0.071)	(0.082)	(0.096)	(0.114)	(0.145)	(0.153)	(0.130)	(0.207)	(0.184)
El Salvador	0.080	−0.111	−0.562	0.303	0.098	−0.375	−0.158	0.387	−0.618
	(0.126)	(0.193)	(0.216)**	(0.145)**	(0.456)	(0.208)*	(0.220)	(0.280)	(0.279)**
Year = 2003	−0.012	−0.112	−0.039	0.002	−0.190	0.052	−0.152	−0.279	0.028
	(0.095)	(0.100)	(0.154)	(0.121)	(0.141)	(0.198)	(0.122)	(0.164)*	(0.189)
Year = 2004	0.161	−0.169	−0.419	0.264	−0.030	−0.383	−0.522	0.000	−0.638
	(0.169)	(0.232)	(0.275)	(0.210)	(0.487)	(0.325)	(0.251)**	(0.000)	(0.322)*
Year = 2005	−0.051	−0.312	−0.831	0.009	−0.271	−0.714	−0.744	−0.323	−0.876
	(0.143)	(0.211)	(0.247)***	(0.164)	(0.460)	(0.266)***	(0.217)***	(0.215)	(0.293)***
Constant	−0.503	−0.093	1.473	−0.826	−0.062	0.932	−0.004	−0.228	0.598
	(0.309)	(0.382)	(0.472)***	(0.422)*	(0.752)	(0.619)	(0.441)	(0.596)	(0.862)
Observations	435	252	171	196	107	77	109	57	46
R-squared	0.13	0.11	0.38	0.20	0.29	0.45	0.36	0.38	0.51

Notes: Standard errors in parentheses; omitted country is Guatemala, omitted year is 2002; instruments are the intent to treat indicator and intent to treat indicator interacted with male (in columns 1–6); for further explanation see Section 4.

* Significant at 10%.

** Significant at 5%.

*** Significant at 1%.

ing: We no longer see a consistent picture of the T2 ∗ male interaction term, which flips between positive and negative signs, none of them significant.

Finally, in columns 7–9 of Table 10 the winner ∗ male interaction term is now negative throughout, with coefficients being statistically significant in the specifications that pool both outcomes of interest, that is where the dependent variable is equal to one if a business either expanded or was newly established. Thus, we find a significantly larger effect, albeit only at the 10% significance level, for starting-up/expanding a business for women in the winning stage. [15] Because of the small samples in the later stages of the program, special care must be taken in the interpretation of the coefficients of the winning stage. In particular, insignificant results for the interaction terms do not imply that there is no gender difference. Overall, the findings in this sub-section are consistent with the hypothesis that women are more constrained than men to come up with funding for business start-up or expansions and as such also in line with non-experimental findings for transition economies reported in Muravyev, Talavera, and Schäfer (2009).

6. CONCLUSION

We investigate whether training and business development programs in developing countries can help improve entrepreneurial skills and foster entrepreneurial activities, such as the creation and expansion of businesses, which at the micro-level is of immediate relevance for the livelihoods of individuals and at the macro-level essential for long-term growth. From a theoretical point of view, our study is related to work on entrepreneurship in diverse areas of economics, such as economic

growth, finance, and industrial organization in which entrepreneurial skills are often taken as exogenous (e.g., Lucas, 1978; Gollin, 2008). If this assumption is true, programs such as the one that we study may have very limited effects on new firm creation. We, therefore, focus on the first-order effects, especially on start-up of new firms. From a policy point of view, the present paper helps to fill a gap because programs such as the one that we study are widespread around the developing world but their impact is generally not yet well understood.

We exploit a quasi-experimental setup in the program under study to estimate the causal impact of this particular training and business development program using data from three Central American countries. The findings are based on the sample of individuals who applied to the program. Generalizing the results beyond such a selected group of individuals is not possible. But arguably the questions asked in this paper are most relevant for precisely those individuals who consider participating in a business development program and not for the general population. Another caveat is that the paper is concerned with the short-term goal of the typical business development program, namely to start or expand a business, and we are unable to say anything about long-term consequences, such as whether businesses fail and what the long-run incomes generated by the entrepreneur are, and how those compare to the entrepreneur's outside opportunities.

It should also be noted that entrepreneurs may be uncertain about their ability, and that remaining in the competition (being moved up to the next round) serves as a signal about this ability to the entrepreneur. This element of the program, which provides external validation of the intended entrepreneurial activity, is specific to the business plan competitions and is not generally provided in business training programs. Thus, the competitive aspect of the program might be important in terms of leading to results, beyond its function as a selection process to narrow down the group of individuals. This should be kept in mind when using the results to guide policy. If this aspect of the program was indeed important as a signal of the quality of the entrepreneur's ideas, the training could still have the intended effect, namely increasing the number of business start-ups, but through an unforeseen channel that may have little to do with the actual content of the training workshop.

Our findings indicate that overall the program is successful in inducing the creation of new businesses and the expansion of existing businesses. Our baseline parameter estimates imply a four to nine percentage points higher probability of opening a business (for individuals without a business before the start of the program) in the treatment group and a 25 to 56 percentage points higher probability of expanding a business (for individuals with an existing business before the program) in the group of treated entrepreneurs. We distinguish between the effect of different stages of the program. The findings suggest that not all parts of the program are equally successful, and we also find differential impacts of different parts of the program on the start-up of new business and the expansion of existing businesses.

For policy purposes, the results suggest that if expansion of existing businesses is the goal, the activities in stage 1 of the training, that is broadly speaking general business training and skills necessary for writing a business plan, seem to be the important ones. On the other hand, if the creation of new businesses is the goal of an intervention, the more specific business plan support activities of the second stages of the training, which are specifically targeted at individuals' business ideas, seem to be the relevant ones. In addition, exploiting the fact that in the last stage the most successful participants of the program (the "winners") receive substantial monetary prizes, we find some evidence for financial constraints. In particular, we do find both statistically as well as economically significant effects of winning on the probability of launching a new business. The findings further suggest that financial constraints are more important for women who wish to start or expand a business than for men.

NOTES

1. Other prominent programs of this type are run by the United Nations Conference on Trade and Development (UNCTAD), the ILO and a number of NGOs such as Enablis and Endeavor. The one program with the biggest scope may be the one run by UNCTAD: UNCTAD supports SMEs in the form of the so-called "Empretec" enterprise development program, which has been implemented in at least 27 countries and has run workshops that have trained more than 70,000 current and aspiring entrepreneurs. Importantly, one of the TechnoServe business plan competitions actually uses the Empretec training program.

2. The exact amounts and number of prizes handed out in our data are as follows: El Salvador 2002: four prizes of US$15,000. El Salvador 2003: five prizes of US$12,000. El Salvador 2005: two prizes of US$6,000. Guatemala 2005: 11 prizes of US$10,000. Nicaragua 2004: nine prizes of US$10,000. Nicaragua 2005: six prizes of US$10,000.

3. Because the identification of the effect depends on estimating the discontinuity precisely, we had put special efforts into our data collection within the neighborhood around the cutoff. Thus, in the one-standard deviation neighborhood around the cutoff, we have data on 435 out of 567 individuals, i.e. for 76.7%, in the 0.5 standard deviation neighborhood, this number increases to 79.5% (260 out of 327 individuals were interviewed).

4. One could imagine that a central office sets a standard above which all those who achieve the standard are allowed into the training and a local office that then moves up scores, just lifting them above the threshold, so that many more individuals can participate.

5. For a theoretical treatment see van der Klaauw, Hahn, and Todd (2001) or Lee and Lemieux (2010). For other prominent examples of this approach, see for example Angrist and Lavy (1999), van der Klaauw (2002), Jacob and Lefgren (2004), or Ludwig and Miller (2007).

6. For example, it is likely that individuals who are interested in a training program such as the one offered by TechnoServe have some ideas about a potential business that they are planning to start. So comparing individuals in the program with individuals from the general population in a simple OLS regression framework in which we regress business outcomes on a training-participation dummy variable would give biased estimates. In this particular example, we expect a positive correlation between u_i and program participation, and OLS would overestimate the true effect of the program. However, one could also imagine that people enter the training workshop because they lack ideas, do not gain anything from the workshop and never start a business or never expand their business, even with training, while those who do have business ideas, and would also gain from the workshop, skip the workshop because they are busy starting-up or expanding their businesses. This would imply a negative correlation

between unobservables and training participation and thus a downward bias on the OLS estimate. It follows that the direction of the bias is indetermined.

7. The other controls serve to reduce the noise and thus to increase precision of the estimates, but they are not necessary for this approach to give consistent estimates.

8. In a first stage, program participation is predicted based on the value of the assignment variable (score $S1$) and an indicator whether the value of the assignment variable falls above the cutoff for program participation. To estimate the parameters of Eqn (3), the estimated participation probability from this first stage is then used in the second stage instead of the treatment indicator variable. If the assignment variable is assumed to enter with the same functional form in the first and the second stage, standard two-stage procedures (such as two-stage least squares) can be used for this.

9. Some other authors have dealt with non-homogenous treatments (e.g. different amounts of financial aid offers, as in van der Klaauw, 2002), however, those are determined by only one assignment variable, and different cutoffs within this one variable determine the treatment amounts.

10. However, we are also able to get some idea about whether training is useful by itself by excluding those individuals who ended up as winners, i.e. also won some monetary prize. When we do so, the main results – not reported here – are qualitatively unchanged (for full results see the working paper version Klinger & Schündeln, 2007). This suggests that the total training effect is not simply due to the financial aspects of the training.

11. We will further investigate the differences in the training effect between male and female individuals in a section below.

12. At the winning stage we have a relatively large number of right hand side variables for a relatively small number of observations, when we use the baseline specifications that includes demographics, year and country dummies. To investigate robustness, we have also estimated regressions that exclude some of the controls. The results, not shown here, indicate that the key findings regarding significance and magnitude of the winner variable in column 1 of Table 7 do not change in important ways when we exclude either demographics, year and country dummies, or both at the same time.

13. It may be surprising that there are individuals who win in this competition but do not open a business or do not expand an existing business significantly. One plausible hypothesis is that the prize money is not sufficient for some more ambitious plans of some entrepreneurs and that they are not able to secure additional funding, or at least not yet by the time of the data collection.

14. Note that in more developed economies the evidence is mixed. For example, Blanchflower, Levine, and Zimmerman (2003) do not find statistically significant differences between male and female owned businesses in the US in their ability to obtain small business finance.

15. In results that are not shown, we find that the difference in the magnitude of the coefficients is even larger if we run the regression separately, thus allowing all coefficients to vary with gender. However, because the number of observations is reduced by running regressions separately, results are less precise and the coefficients not statistically different.

REFERENCES

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics, 114*(2), 533–575.

Baumol, W. (1968). Entrepreneurship in economic theory. *American Economic Review, Papers and Proceedings, 58*(2), 64–71.

Blanchflower, D. G., Levine, P. B., & Zimmerman, D. J. (2003). Discrimination in the small-business credit market. *Review of Economics and Statistics, 85*(4), 930–943.

Bosma, N. S., Acs, Z. J., Autio, E., Coduras, A., & Levie, J. (2009). *Global entrepreneurship monitor 2008 executive report*. Babson Park, MA, US: Babson College, Santiago, Chile: Universidad del Desarollo and London, UK: London Business School.

Bosma, N., & Levie, J. (2010). *Global entrepreneurship monitor 2009 executive report*. Babson Park, MA, Babson College, Santiago, Chile: Universidad del Desarollo, Reykjavik, Iceland: Háskólinn Reykjavík University and London, UK: London Business School.

Coduras Martínez, A., Levie, J., Kelley, D. J., Sæmundsson, R. J., & Schøtt, T. (2010). *Global Entrepreneurship Monitor special report: A global perspective on entrepreneurship education and training*. Babson Park, MA, Babson College, Santiago, Chile: Universidad del Desarollo and Reykjavik, Iceland: Háskólinn Reykjavík University.

Datta, S. (2008). *The Impact of Improved Highways on Indian Firms*. Working paper.

de Mel, S., McKenzie, D., & Woodruff, C. (2007). *Who does Microfinance Fail to Reach? Experimental Evidence on Gender and Microenterprise Returns*. BREAD Working Paper No. 157, Bureau for Research and Economic Analysis of Development.

de Mel, S., McKenzie, D., & Woodruff, C. (2009). Measuring microenterprise profits: Must we ask how the sausage is made?. *Journal of Development Economics, 88*(1), 19–31.

Djankov, S., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2002). The regulation of entry. *The Quarterly Journal of Economics, 117*(1), 1–37.

Gollin, D. (2008). Nobody's business but my own: Self-employment and small enterprise in economic development. *Journal of Monetary Economics, 55*(2), 219–233.

Greenstone, M., Hornbeck, R., & Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy, 118*(3), 536–598.

Jacob, B. A., & Lefgren, L. (2004). Remedial education and student achievement: A regression–discontinuity analysis. *Review of Economics and Statistics, 86*(1), 226–244.

Karlan, D., & Valdivia, M. (in press). Teaching entrepreneurship: Impact of business training on microfinance clients and institutions. *Review of Economics and Statistics.*

Klinger, B., & Schündeln, M. (2007). *Can entrepreneurial activity be taught? Quasi-experimental evidence from Central America*. Working paper no. 153. Center for International Development at Harvard University.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature, 48*(2), 281–355.

Lucas, R. (1978). On the size distribution of business firms. *Bell Journal of Economics, 9*(2), 508–523.

Ludwig, J., & Miller, D. L. (2007). Does head start improve children's life chances? Evidence from a regression discontinuity design. *The Quarterly Journal of Economics, 122*(1), 159–208.

Muravyev, A., Talavera, O., & Schäfer, D. (2009). Entrepreneurs' gender and financial constraints: Evidence from international data. *Journal of Comparative Economics, 37*(2), 270–286.

Schumpeter, J. A. (1911). *Theorie der wirtschaftlichen Entwicklung*. Leipzig: Duncker & Humblot.

Svensson, J. (1998). Investment, property rights and political instability: Theory and evidence. *European Economic Review, 42*(7), 1317–1341.

van der Klaauw, W., Hahn, J., & Todd, P. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica, 69*(1), 201–209.

van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression-discontinuity approach. *International Economic Review, 43*(4), 1249–1287.

APPENDIX A

A.1 *Content of the business training program*

This appendix provides a more detailed description of the content of TechnoServe's "business plan competitions". This appendix draws heavily on material from TechnoServe.

Phase 1 Interested (potential) entrepreneurs submit applications. These are scored to identify promising applicants. Scores are solely based on the observable information that is submitted by the applicant, including the demographics, prior business experience, and the business idea that is described at the time of application. A fixed number of individuals are admitted to the program, which determines a cutoff, and individuals with scores above the cutoff are allowed to enroll in the training program. In our data set approximately 43% of all applicants are accepted into the training workshop.

Phase 2 The training program in the business plan competitions is meant to provide participants with all of the tools and knowledge necessary to take their business idea and compose a complete and detailed business plan. It is organized around the key sections of a business plan, with a class-based lecture with theory and examples, along with some limited breakout discussions among participants. This process forces the participant to consider all of the facets of their business, such as production, finance, marketing, sales, and logistics, in a systematic way. The process is organized in layers, with each successive layer providing greater detail on the previous one. The first layer consists of two units, the second layer of three units, and the third layer of six units.

The first unit is the business description, in which participants will have to describe the business in general terms, compose the firm's mission and vision statements, define the firm's goals, and perform a structured analysis of strengths, weaknesses, opportunities, and threats related to the business idea. Examples from other businesses are presented. The second unit covers the description of the business concept. Its goal is to teach participants to define the concept in terms of how it satisfies market needs, instead of defining the concept from the business owner's own point of view.

The next three units cover industry analysis, market analysis, and descriptions of the product/service. These are meant to dig one level deeper in their description of the business concept. In the section on industry analysis, participants are taught how to define industries in terms of employment and value added, as well as market trends. The trainers also discuss describing the industry in terms of geographic and temporal production patterns, and finally suggest possible sources of data in the country of the competition, such as relevant government agency websites, industry association websites, and general information sources. The unit on market analysis describes how one can precisely identify potential target markets through segmentation by geography, demographics, incomes, and so on, and then evaluate the potential of each market segment to determine the most suitable ones to target. The unit on description of the product or service is meant to re-orient participants' conceptualization of their product to that of the customer rather than producer by encouraging them to describe it in terms of its principal functions, characteristics, and uses to consumers.

The base layer of the business plan training consists of six units: competitors, marketing plan, sales projections, operations, organization and administration, and financial projections. In the competitors unit, the trainers teach participants the dimensions on which they need to characterize the competition and compare it to their own capabilities. They are also taught about the importance of barriers to entry for future competitors. The market plan module attempts to cover issues, such as how to differentiate a product/service, how to set prices, and talks about creating and protecting brand names, distribution chains, and general marketing strategies, such as advertising campaigns, strategic alliances, and free samples. The sales projection unit suggests how to make monthly sales projections, including sources of estimates, and how to then plug them into the financial projections. The operations module includes lectures on lean production, standardization of inputs, employment options, quality control, and inventory control. The organization and administration unit teaches participants how to make organizational charts and job descriptions, as well as design organizations around customer service. The final and longest module is on financial projections, including investment estimates, operational cost estimates, financing requirements (including potential sources of investment), financial statements (balance sheet, income statement, and cash flow statement), and finally overall profit/loss projections.

In addition to these topics, there are also lectures on general entrepreneurial attitudes and orientation, as well as country-specific information on starting a business, such as the relevant government departments and programs, different legal forms, and potential sources of capital.

At the conclusion of this phase, participants have been presented all of the components of a business plan. They prepare a draft business plan, which is presented to a panel of judges, who are asked to judge those draft business plans in a number of dimensions, for example, with respect to the size and development of the expected market of the product/service that is proposed, with respect to the marketing and sales strategy, or whether financial projections seem reasonable. A particularly large weight is on the (final) question whether the judge would invest his/her own money into the proposed business. Those participants with the top phase 2 scores continue to phase 3. In our data set approximately 37% of all individuals who underwent this evaluation process are accepted into phase 3.

Phase 3 In phase 3, participants are given detailed feedback in order to revise and improve their business plan. They are also given one-on-one time with consultants to help refine their business plan. At the conclusion of this phase, they have a longer and more detailed business plan that has benefited from review and advice by both consultants and the judges. These plans are scored, and the top ones win the financial prize. In our data set approximately 32% of all phase 3 participants win a prize.

Timing In the case of El Salvador, the training consists of the 7-day UNCTAD Empretec course, followed by two four-hour sessions by TechnoServe on how to compose a business plan using the methodology described above. The participants then have approximately 8 weeks before the business plans are due. In the case of Nicaragua and Guatemala, there is no Empretec training, and instead the classes based on the business plan methodology described above are done in more detail. They are divided up into approximately three full-day sessions, one per weekend. The participants have approximately one month to then compose and turn in their business plans. For all competitions, the plans advancing to phase 3 are revised over a period of 1–2 weeks.

The various phases of the program are summarized in Table 11.

A.2 *Summary statistics*

See Table 12.

Table 11. *Timeline and nomenclature*

Phase	Condition	Activity	Score at end of phase
Phase 1	None, all can apply	Application	Score 1 ($S1$)
Phase 2	If score 1 is above a cutoff	Training ("training stage 1", T)	Score 2 ($S2$)
Phase 3	If score 2 is above a cutoff	Additional training ("training stage 2", $T2$); finalizing a business plan	Score 3 ($S3$)
Final stage	If score 3 is among the highest	Winner of competition (W); receipt of a prize money to be invested in a business	

Table 12. *Summary statistics for all applicants and those who participated in the first round of training*

Variable	Obs	Mean	Std. Dev.	Min	Max
All applicants					
Age	655	36.13	11.16	18	78
Gender (1 = male)	655	0.72	0.45	0	1
Owned business before start of program? (1 = yes)	641	0.39	0.49	0	1
If owning a business: sales before program (in US$ 1000)	214	83	208	0	2,137
If owning a business: employees before program	180	8.1	13.7	0	96
All participants of first stage training					
Age	271	36.03	11.01	20	78
Gender (1 = male)	271	0.69	0.46	0	1
Owned business before program? (1 = yes)	257	0.42	0.49	0	1
if owning a business: sales before program (in US$ 1000)	102	70	132	0	759
If owning a business: employees before program	68	9.1	17.0	0	96