Hunting and Nurturing Gazelles: Evidence from Business Plan Competitions in Ethiopia

A Dissertation

Submitted to the National Graduate Institute for Policy Studies

(GRIPS)

in Partial Fulfillment of the Requirements for the Degree of

PhD in Development Economics

by

Abebe Ambachew Ayana

August 2022

Hunting and nurturing gazelles: Evidence from business plan competitions in Ethiopia

A Dissertation

Submitted to the National Graduate Institute for Policy Studies (GRIPS)

in Partial Fulfillment of the Requirements for the Degree of

PhD in Development Economics

by

Abebe Ambachew Ayana

August 2022

Abstract

Business plan competition has been considered as an innovative support policy instrument to spur entrepreneurship and small business development in developing countries. Government, NGOs, and international development organizations invest huge resources for the implementation of this support program in many countries around the world. However, empirical evidence on its effectiveness is scant.

This dissertation evaluates the effectiveness of two business plan competitions (Bruh and EDC) ran by the federal government of Ethiopia in identifying high-growth potential enterprises (potential gazelles) and nurturing them through the interventions designed as part of the program.

To this end, I compiled administrative data from the competition records and conducted a follow-up survey on the universe of about 500 applicants to measure actual business outcomes a year after the application.

The first analytical chapter evaluates the causal effect of the training intervention of the business plan competitions on business entry and expansion using a fuzzy regression discontinuity design by exploiting business plan scores and exogenous cut-off points. The result revealed that in any measure of business success the training beneficiaries were not different from their rejected counterparts because the rejected applicants (control group) had also similar training in other similar programs. Though the study is not informative about the effectiveness of the program, the substantial take-up of the control group in substitute program documented in this study could be helpful to explain the modest or negligible impacts the entrepreneurship training programs reported in previous studies.

The second analytical chapter examines if a business plan competition can be a successful policy option to identify potential gazelles through its rigorous screening procedure. In general, the results show that the business plan score is a significant predictor of entrepreneurial success. Judges were more effective in predicting enterprise growth at the bottom and top of the distribution, implying that the most promising projects and the non-serious ones are relatively easier to identify. However, I found heterogeneities in prediction success between the two competitions despite their implementation in the same setting. The results helped us provide preliminary explanation for the mixed results of the previous literature and draw conditions under which the experts' prediction accuracy could be improved.

Overall, the study suggests that a properly designed and implemented business plan competition is helpful at least to differentiate firms based on their growth potential which is a key to tailored policy and proper targeting.

Acknowledgement

First and foremost, no word would suffice to express my deep gratitude for my main advisor Prof. Litschig Stephan for his extraordinary support at every-stage of the research. The completion of this study would not have been possible without his continuous support and persistent guidance. He has been a source of my strength and inspiration. I feel blessed and lucky to have him as my advisor.

I am also extremely grateful for my sub-advisors, Prof. Tetsushi Sonobe and Prof. Akio Yamazaki for their supports, insightful comments, and guidance since the very design of this study. I have also learned a lot from and got inspired by them right from the course works of the PhD journey and they have been supportive for me in every aspect.

I would like to express my heartfelt love and gratitude for all my family members including my mother, my wife, kids, my siblings, and in-laws. My special appreciation go to my wife Mekdes Endale for shouldering the responsibility of raising kids and letting me focus on my study; and for my kids Yordanos Abebe and Barkot Abebe for their understanding and patiently waiting for three years without getting the hug of their father.

I am also deeply grateful for my friends and colleagues for their support and encouragement in various ways during the entire journey of this PhD program, particularly for Dr. Mulu Gebreeyesus, Dr. Tigabu Getahun, Dr. Girum Abebe, Dr. Tsegaye Gebrekidan, Dr. Million Gebreyes, Dr. Dawit Kelemework, Dr. Berihu Assefa, Dr. Jemal Mohammed, Dr. Mulugeta Getu, Mr. Yalew Mekonnen, Mr. Melese Sitotaw, Mr. Arega Kumligne, Mr. Chane Ejigu, Dr. Hiwot Mekonen, Dr. Yohannis Mulu, Mr. Workneh Ayenew, Mr. Fisseha Asmare, Mr. Abay Agidew, Ms. Yichalem, and Mr. Degnet Shiferaw to name a few.

I want to thank various organization and their leaders who contributes for the successful completion of my study in one way or another. These include the Entrepreneurship Development Center (EDC Ethiopia) and its supportive staff particularly Wondwossen Tsegaye; the Jobs Creation Commission of Ethiopia and its staff members, especially Robel Zewdie and Yoseph Gizachew; the Federal Small and Medium Manufacturing Promotion Authority and its director general Mr. Asfaw Abebe; and my survey team members and respondents for their excellent support in the data collection.

Last, but not least, I would like to express my sincere appreciation for the Japanese government (MEXT) for providing me with a generous scholarship and allowing me to study in Japan. My special thanks goes to my university, the National Graduate Institute for Policy Studies (GRIPS), for having me as a student and for partly funding the research through its PhD research grant. In this respect, I am so grateful for the GRIPS research support team especially Mr. Kinase for his kind support. I am also indebted to the Policy Analysis program director Prof. Alistar Munro for his support and follow-up, and for the program coordinator Aiko Hashimoto for her kind support and cooperation during my stay in GRIPS. Arigato!!

DEDICATION

To my mother, Tiftie Alemu Abtew

Table of contents

Abstract i

Acknowledgement iii

Table of contents vi

CHAPTER 1 1

1. Introduction 1

1.1. Overview 1

1.2. Organization of the dissertation 4

2. Context and Data 5

2.1. The macro-Context 5

2.2. Description of the program 6

2.2.1. Overview of the program 6

2.2.2. Bruh Entrepreneurship Competition 8

2.2.3. EDC Startups' competition 13

2.3. The Data 15

2.3.1. Data collection methods 15

2.3.2. Description of the Data 20

CHAPTER 3 26

3. Nurturing Startups: The role of trainings of the business plan competitions 26

3.1. Introduction 26

3.2. Literature Review 32

3.3. Empirical strategy 41

3.3.1. Specification of the model 41

3.3.2. Test of the basic identification assumption: Manipulation test of score 44

3.3.2.1. Falsification test on pre-determined covariates 44

3.3.2.2. The Density Test 46

3.4. Estimation results of the models 49

3.4.1. The First-Stage Estimates 49

3.4.1.1. The Nominal First-Stage (NFS) 49

3.4.1.2. The Effective First-Stage (EFS) 52

3.4.1.3. Additional Robustness Checks on the effective First-Stage 63

3.4.2. Summarizing the findings of the first-stage 66

3.4.3. Estimation of the Reduced form equation 67
3.4.3.1. Reduced form results for business entry and survival.. 68
3.4.3.2. Reduced form results for performance indicators 74
3.5. Discussion and Concluding remarks 80
CHAPTER 4.. 82
4. Predicting entrepreneurial success and identifying gazelles using business plan competitions 82
4.1. Introduction 82
4.2. Review of related Literature.. 86
4.2.1. The rationale for bifurcation of firms and targeting 86
4.2.2. Methods to identify gazelles... 88
4.2.3. Evidence on using experts' judgment to predict entrepreneurial success 89
4.3. The Empirical Strategy.. 95
4.3.1. The Model 95
4.3.2. Definition the outcome of variables 96
4.4. Results and discussions 99
4.4.1. Predicting firm entry and survival 99
4.4.2. Predicting employment.. 110
4.4.3. Predicting Sales 115
4.4.4. Predicting profit.. 119
4.4.5. Predicting Aggregate growth 124
4.4.6. Heterogeneities in predicting success 129
4.4.6.1. Results disaggregated by competition type (Bruh and EDC) 129
4.4.6.2. Possible explanations for heterogeneity in Bruh versus EDC results 132
4.4.6.3. Prediction for existing Versus new firms 139
4.4.7. Robustness checks 141
4.5. Concluding remarks 143
CHAPTER 5 145
5. Conclusions and policy Implications 145
5.1. Conclusions 145
5.2. Policy Implications.. 146
5.3. Limitations and future research.. 149
References.. 150

APPENDICES 156
Appendix 2. Additional materials for chapter 2........ 156
Appendix 2.1. Evaluation criteria used to screen applicants........ 156
Appendix 2.2: Timeline and steps of the project 159
Appendix 2.3. Survey Questionnaire 161
Appendix 2.4. Some descriptive results about training........ 172
Appendices 3........ 173
Appendix 3.1. Results for additional robustness checks for the effective first-stage........ 173
Appendix 3.2. Reduced form results for *Bruh* sub-sample 185
Appendix 3.3. Reduced form results for EDC sub-sample 189
Appendix 4: Additional results of the fourth chapter 193
Appendix 4.1. Results of linear Versus cubic specifications 193
Appendix 4.2: Marginal effects from the Probit and Logit regression models........ 194
Appendix 4.3: Regression tables for Bruh and EDC sub-samples........ 197
Appendix 4.4. Graphical representations of the Quintile analyses 208

List of Figures

Figure 2.1 Distribution of standardized score of the business plan. 21
Figure 3. 1 Density Test for manipulation of scores 48
Figure 3. 2 Nominal First-stage (NFS) result of the probability of attending training of the program. 51
Figure 3. 3 Effective First-stage (EFS) result of the probability of attending entrepreneurship training from any program 53
Figure 3. 4 Effective First-stage result disaggregated by contents of the training covered 59
Figure 3. 5 Effective first-stage for covering major training modules of the program separately 62
Figure 3. 6 Reduced form results on probability of owning a business (self-employment) for the full sample 69
Figure 3. 7 Reduced form results on probability of own employment for the full sample 70
Figure 3. 8 Reduced form results on owning licensed (or formal) business 71
Figure 3. 9 Reduced form results on access to finance 73
Figure 3. 10 Reduced form results on dichotomous measures of firm performance (Full sample) 76
Figure 3. 11 Reduced form results on continuous measures of firm performance (Full sample) 78
Figure 4. 1 Plots of score versus probability of operating a firm for the full sample 102
Figure 4. 2 Plots of score versus probability of operating a firm for non-trained and trained sub-samples separately 103
Figure 4. 3 Binned scatter plots of score Versus level of employment for full sample 111
Figure 4. 4 Binned scattered plots of score versus level of employment for trained and non-trained sub-samples 112
Figure 4. 5 Binned scatter plots of score versus the IHS transformation of sales for full sample 116
Figure 4. 6 Binned scattered plots of score versus the IHS transformation of sales for non-trained and trained applicants 117
Figure 4. 7 Binned scattered plots of score versus the IHS transformation of profit for full sample 121
Figure 4. 8 Binned scattered plots of score versus the IHS transformation of profit for non-trained and trained applicants 122
Figure 4. 9 Binned scattered plots of score against aggregate growth for the full sample 124
Figure 4. 10 Binned scattered plots of score against aggregate growth for non-trained and trained applicants 125
Figure 4. 11 Aggregate growth index by score quintile (full sample) 127
Figure 4. 12 Estimated coefficients of score in comparison with Bruh versus EDC 131

List of Tables

Table 2. 1 Summary statistics of selected variables used in this study 23
Table 3. 1 Summary of literature on impact of business plan competitions, accelerators, and related interventions 36
Table 3. 2 Falsification test results of the running variable based on pre-determined covariates 46
Table 3. 3 Nominal First-Stage Estimates of the Effect of Scoring above the Cutoff on Training Attendance 54
Table 3. 4 Effective First-Stage Estimates of the Effect of Scoring above the Cutoff on Training Attendance 54
Table 3. 5 Reduced form Estimates of the effect of scoring above the cutoff on business entry for the full sample 74
Table 4. 1 Simulating the effects of changes in standardized scores on probability of operating a firm for cubic specification in comparison with the liner model 105
Table 4. 2 Main results about the prediction of firm ownership and survival by panel of experts 109
Table 4. 3 Prediction of firm level of employment by panel of experts 114
Table 4. 4 Prediction of firm sales by panel of experts 118
Table 4. 5 Prediction of firm profit by panel of experts 123
Table 4. 6 Prediction of firm aggregated growth by panel of experts 128
Table 4. 7 Prediction of success by score from panel of experts disaggregated by firm type (without controls) 140

CHAPTER 1

1. Introduction

1.1. Overview

Developing countries design and implement various policies and strategies to achieve industrial development as it is a key deriver of structural transformation (Ohno et al., 2022). Currently, in many countries, the Small and Medium Enterprise (SME) policy is regarded as part of the industrial policy and a proper SME and entrepreneurship development policy crucial for industrial development.

It has been more than seven decades since the notion of small and medium enterprises (SMEs) and entrepreneurship development was introduced in the growth and development agenda of countries with the introduction of targeted policies and establishment of SME support agency. For instance, government funded SME agencies were instituted in 1948 in Japan, 1953 in USA, 1954 in India, 1966 in Tanzania, and 1976 in Turkey (OECD, 2004). Despite such a long history of development efforts, in some countries SMEs were perceived rather as a synthetic construction mainly of "social and political" importance for long periods (Hallberg, 2000). Private sector development policies and strategies were skewed towards the need of large businesses even if SME constitute most part of the private sector in developing countries. Later, the all-round contribution of SMEs to the development of any emerging economy has been well acknowledged. Thus, nurturing SMEs becomes the central focus of current development policies and plan of most developing countries.

Several policies and programs have been experimented by governments, NGOs, and other actors to foster entrepreneurship. Businesses plan competition is one of these policies that attracted the attention of donor and huge resource is channeled for its implementation in many countries around the world. Business plan competitions commonly have dual purposes: selecting high-growth potential business through rigorous screening procedure and directly provide some supports (like training or grant) for part of the contenders (McKenzie, 2017). However, there is a dearth of studies regarding the effectiveness of business plan competitions in meeting their dual purposes.

This study intends to examine both objectives of a business plan competition by taking cases of two national business plan competitions conducted in Ethiopia using a quasi-experimental design method. In the first main analytical chapter of the dissertation, I evaluated the short-term impact of the non-monetary support (or training) component of the business plan intervention using a fuzzy regression discontinuity design by comparing applicant just below and above the business plan scores' cutoff. To this end, a year after they applied for the program, I conducted a follow-up survey on all participants of the competitions (about 500) to measure their business outcome and trace any possible treatments by substitute program.

The follow-up survey revealed that substantial numbers of rejected applicants (control groups) received the same types of training in substitute programs running in the market. This situation makes the study to have indefinite answer about the effectiveness of the program being evaluated (Bruh and EDC). Though it was not possible to provide conclusive answer for the effectiveness of the program in nurturing startups, the study uncovered that business outcomes of selected applicants of Bruh and EDC who passed the first screening are not statistically different from that of their rejected counterparts at least partly due to the contaminated controls.

In the second analytical chapter, I address whether business plan competition is a successful policy option to identify potential gazelles through accurately predicting growth potentials of motivated entrepreneurs who applied to the competitions. The study revealed that, in general, business plan competitions in Ethiopia succeeded in predicting entrepreneurial success, as measured by business entry and survival, level of employment, sales, profit, and an aggregate growth index. This implies that business plan competition is a successful policy option to identify enterprises with a good growth potential. However, the success in prediction of future business outcomes is profound only in case of EDC.

The major contributions of this dissertation are summarized as follows.

- It provides empirical evidence on the policy debate whether business plan competition is a successful option to foster industrial development in developing countries.
- It provides a plausible explanations why previous evaluations of entrepreneurship training failed to get a remarkable positive impact: it is possibly because control groups got the treatment in substitute programs.
- The study also draws a preliminary condition under which the *ex ante* identification of growth potentials of businesses by experts' judgment could be improved to achieve better accuracy in prediction.
- This dissertation also contributes for future research design of any impact study by demonstrating the importance of checking subjects after the baseline for any possible treatments by substitute programs, which is a key ensure the validity of the counterfactual.

In sum, by focusing on various aspects of the program, this study delivered concrete empirical evidence on the role of business competition in spurring entrepreneurship in developing countries.

1.2. Organization of the dissertation

This dissertation is organized into five chapters including this introduction chapter and the remaining four. The second chapter presents context and data. I start with a brief highlight of the macroeconomic context of Ethiopia to inform under what environment the program of interest was implemented, and then details of Bruh and EDC business plan competition are described. The final part of this chapter is about the data. The data and program description discussed in this chapter are commonly used for both analytical chapters.

The first main analytical chapter of the dissertation is about the causal effect of the training program, and it is presented in the third chapter. The fourth chapter is the second main analytical chapter of the dissertation, and it is about the role of business plan in identifying gazelles. Both the third and fourth chapters are organized to be stand-alone chapters with the inclusion of information in chapter 2. The fifth chapter concludes and draw some policy implications.

CHAPTER 2

2. Context and Data

2.1. The macro-Context

Ethiopia is one of the old independent states which is located in the horn of Africa. This landlocked country is the second most populous country in Africa next to Nigeria with an estimated population of more than 115 million. About 80% of the population lives in rural area where agriculture is the main livelihood. Agriculture, industry and services contributed about 23.5%, 29.3%, and 39.6% of the Ethiopian GDP, respectively (NBE, 2022). Ethiopian economy is one of the fast-growing economy with an average growth rate of 9.5% over the last 15 years(World Bank, n.d.) The informal sector provides more than 60% of the urban employment. The country's labor force increases by 2 million every year and the absorptive capacity of the labor market is being challenged. As a result, youth unemployment particularly in urban area are high(CSA (Central Statistical Agency, 2021).

The government of Ethiopia is known for its active industrial policies including construction of large industry parks with the aim to be the African manufacturing hub. To promote self-employment and local industrialization, the government hugely supports the development of Micro, Small, and Medium Enterprises (MSMEs). The first full-fledged micro and small enterprises development strategy was introduced 1997 and then revised in 2011 with a clear support framework for the sector. Since then, several government and NGO programs aiming at enterprise development have been implemented. These programs focus on skill development, technical training, kaizen, facilitation of market linages, microfinancing, development of working premises like industry clusters and working shades, among others(Gebreeyesus et al., 2018).

Despite these efforts, the entrepreneurship landscape is not enabling for startups and small businesses. Access to finance is a top problem to do businesses due to high collateral requirement and complicated procedures to get bank loan (World Bank, 2015). Though training opportunities seem to be relatively easily accessible due to the expansion of TVET colleges and universities, the skill mismatch is raised as a serious concern by industrialist.

In short, Ethiopia is a country with untapped opportunities for business including the huge size of the local market, strategic location of the country to access markets of Europe and middle east but some challenges to do business including the current political instability.

2.2. Description of the program

2.2.1. Overview of the program

Our program of interest is broadly the entrepreneurship support program for startups which comprises of two nation-wide business plan competitions called *Bruh* by JCC and *EDC start-ups' incubation* by EDC and targets young entrepreneurs with innovative business ideas or startups businesses. By tackling their constraints, the program aims to encourage young entrepreneurs with a good growth potential to start their own business, accelerate business growth, and expand their level of operation. Though the program is implemented as two independent projects, the fact that both are broadly similar in term of objective, target group, geographical coverage, type of intervention, and timeline gave us the opportunity to consider both cases for this study. Both combine incubation and competition (thus, *incupetion*) to select high-growth potential enterprises (gazelles) and provide them with grants of about 5000 USD for final winners and business development supports like training for applicants who can pass the first-round screening..

The first call of both competitions attracted about 640 applicants. After removing illegible and duplicated applicants within each competition, a total of 545 contenders eligible business plans got scored by a panel of judges or experts assigned by the competition organizers, which is the first-round screening. Based on their average scores, top 248 of the applicants passed the first screening and they were offered for the slots for the training intervention and also advance to the next round competition. These applicants are those who scored above the cutoff in their respective competition and considered as treatment group. Whereas the remaining applicants who scored below the cutoff were rejected and eliminated from the competition; and this group is control group.[1] Among the 248 offered applicants, 168 have attended the training provided by the competitions designed as part of the program. This implies that in the actual implementation this intervention, there were some cases of no-shows while there were not any crossovers, which has an implication for model selection. In the next rounds of the competitions, contestants were provided more customized supports to further develop their business idea and business plan. The competitions were concluded by final pitch competition based on which top 26 startups (20 in case of *Bruh* and 6 in EDC) were selected as final winners who were entitled to get the prize money (cash grant). The first editions of both competitions have come to end in June 2021. The complete timeline, process of the program, and key activities of the program including this impact study are summarized in Appendix 2.1, Figure 2.A.1. While this is a general overview of the program, further details of both competitions are presented as follows.

[1] In this study, applicants who scored above the cutoff are referred as 'offered' group which is also called in the literature as accepted applicants, successful applicants, winners, and qualified applicants, among other names. On the other hand, for applicants below the cutoff, I will use the term 'rejected' applicants, which is synonymous with terminologies such as runners-up, losers, unsuccessful, non-winners, and non-qualified applicants.

2.2.2. Bruh Entrepreneurship Competition

Bruh is a nation-wide business plan competition program for startups designed and implemented by the federal government of Ethiopia, specifically by the Jobs Creation Commission (JCC), in collaboration with the Master Card Foundation and Target Business Consultants. JCC is the organ of the federal government which was established in 2018 as per proclamation No.1097/2018 and the regulation No. 435/2018 with a mandate to 'drive the job creation agenda including the coordinating and supporting of job creation efforts made by various stakeholders and ensuring the adoption and implementation of pro-employment policies to achieve a sustainable jobs ecosystem'. Currently, it has been restructured as Ministry of Labor and Skills.

This competition targets young potential entrepreneurs all over the country, aged 15-29 years, or new businesses with not more than two years of operation. With its step-wise process, the program intends to select innovative businesses/business ideas and provide technical and financial supports which could help them establish and run successful business enterprises with high potentials of job creation. It has been envisaged that the Bruh will serve as a benchmark for other organizations that are interested to perform similar activities in Ethiopia.

This program was launched in January 2021 and will have a life span of three years. Over this period, one or two business plan competitions will be conducted per year in order to meet its three years' targets set to train about 500 selected startups, provide financial award (grant) for top 300, and provide additional business development support for the top 200 startups through their six months long acceleration program. Train, award, and accelerate are the three major pillars of the program's intervention. The details of the actual implementation of the program have been discussed below for each round. The discussion presented below is limited to the first edition of the *Bruh* business plan competition.

The first edition of Bruh business plan competition was conducted from January 2021 to May 2021. Having made sufficient preparations, the call for Bruh first round competition was announced in various medias and remained open from 1st January 2021 to 7th February 2021. Applicants had two options to apply depending on their convenience. Those who have internet access can apply online from the program's website while others can submit their application in-person through its regional counter parts, the urban job creation offices which operates in all districts across the county. The competition attracted a total of 345 applicants in this round. After doing the necessary pre-screening, 277 businesses which consists of 730 founding members (individuals) were found to be eligible for the competition.[2] The proposed businesses are in different sectors including IT based startups, agriculture and agro-processing, manufacturing, construction, and other services.

For the first screening process, a panel of 5 experts (judges) have been formed to individually score the proposed business ideas based on 5 pre-determined criteria set to evaluate the viability of the business. The criteria used at various stages of the evaluation is presented in Appendix 2.3. Scores of all the judges were averaged to selected the top 70 businesses (which have 140 founding members) for the next stage from the 277 eligible applicants.[3] The program implementers had to invite as many as 103 applicants whose scores are equal or greater than 63.05 per cent (which is the exogenous cut-off point) in order to fill the confirmed lists of 70 startups (pre-determined capacity) that are available and eligible to enter to what they call it 'bootcamp' for intensive training. Finally, 61 startups, represented by 112 founding members, showed-up for the bootcamp, showing that the program reached out about 87% of its target.   In this process, there is no any

[2] One applicant did not have complete information and another one was a duplicate, which make the scored list to be 275.
[3] The actual invitees are 71 startups, but two startups have been merged after first screening which makes the count to be 70..

cases of crossovers while there are about 41 non-compliers to the right of the cut-off (14 of them were not willing to commit the required time for the bootcamp, 12 had won other similar grants and were excluded by JCC, 6 were not traced, and 9 had confirmed to attend but actually didn't show up for various personal reasons).

The bootcamp is aimed to bring all contenders who passed the first screening together in the same place for intensive training which is designed to further enrich their ideas, develop their business plans, and prototyping before going for the final pitch competition. The bootcamp had a total duration of about one month, March 22-April 24/2021, in the premises of Ethiopian Management Training Institute in Bishoftu town, Ethiopia. The organizers of the competition had to cover all the costs including meal, accommodation, daily allowances, and other running expenses for trainees, trainers, and facilitators to stay together all the days and nights during the entire period of the bootcamp. That allows the competition participants to interact each other, closely work in groups, expand their networks, influence one another, and reshape their ideas. Needless to say, this could be excellent complements to the formal training sessions of the program and can be considered as an incubation by itself.

The whole period of the bootcamp was designed to have a formal training divided in to six major themes. These are:

- Entrepreneurship Competency
- Holistic Business Idea Development
- Visual Prototype and Product Development
- Legal Business Setup
- Market Research and Unique Selling Proposition (USP), and
- Business Plan Preparation

The training was offered by certified trainers hired for this purpose and duration of each training program is summarized in Appendix 2.2, Table 2.A.1. The training program was further supported by experience sharing events by guest speakers which involves talks on success and failure stories, industry visit, field trips, and fun games played in groups.

The bootcamp could be used not just to provide trainings for the competition participants but it also gives the golden opportunity for the organizers to closely look at the strengths and weaknesses of the proposed business idea, the entrepreneurial traits and other personal behaviors of each participant, as to how quickly each can adopt new skills gained in the process to make step-by-step improvements on their original idea and gather any important information about each participant. This significantly helps reduce the problem of asymmetric information in the endeavors of selecting high potential, genuine, and constrained entrepreneurs for further support.

In line with this view, two additional screenings have been conducted during the bootcamp before the final pitch competition. The second-round screening (which is the first within bootcamp screening) was made two weeks after the start of the bootcamp and immediately after the completion of the first two training themes listed above. At this stage, the top 40 businesses were selected to stay in the bootcamp, and thus continue to the next step of the competition, while the remaining 21 were eliminated from the competition. Again, using the same procedure, the third-round screening was undertaken after two more training programs to reduce the final pitch competition participants to the best 35 businesses. In both cases, groups of judges were formed from trainers themselves to evaluate viability of businesses based on a set of 10 criteria after each team made about 15 minutes (in the second round) and 10 minutes (in the third round) presentation.

Having attended the final training programs on legal business setup and business plan preparation and made sufficient preparations, these 35 businesses contested for the final pitch competition by presented their final business plan in front of a group of 4 judges.[4] This final competition has a pre-determined and clear objective of selecting the 20 winners. Based on the 5 criteria listed in Appendix 2.3, each judge scored every business and the average scores were used to determine the 20 winners of *Bruh* first edition business plan competition. In a colorful closing ceremony held on 27 May 2021 in Addis Ababa, the winners have received 200,000 Ethiopian Birr (equivalent to about USD 5000) cash each as a grant for their businesses.

According to JCC, top 10 of the winners set to receive additional business development supports in their 6 months-long acceleration program which helps the startups to be fully operational businesses. Further, the process of the competition since the start of the was recorded and televised in a national TV channel (Fana TV) and social medias in 6 episodes, each last for about 30 minutes.

Someone could raise a concern about the possibility that broadcasting the competition as a TV show could affect the scoring process some applicants with a catchy character could be preferred to appear in the show regardless of their business potential. However, this concern is not valid in the context of this competition for several reasons. First, the concern would have been valid if I had used scores after the second screening. Because it is only the activities after entering to the bootcamp and mainly the last 35 (finalists) were part of the program. But in the whole study, I use scores from the initial screening which was made even before meeting and getting to know with applicants in-person.

Second, even for the scores after the first-screening the concern is not valid because the TV show and the competition are owned by two different organizations. JCC is a federal government office

[4] Looking deep into the profile of judges indicate that one from a government agency, one from a big private bank, two from renowned large companies. The composition of judges clearly shows that there is no any potential conflict of interest between the judges and the contestants.

that bothers about macro level policies and programs to create more jobs. Their main concern is to select the best candidates who can grow fast and create jobs. There is no reason to care about Fana TV to have a good show, get more advertisement, and get more viewers. They just wanted to televise the program as a biproduct of their activities for other youth to learn something about entrepreneurship from this process. In addition, I watched the 6 episodes of the program and the entrepreneurs were not the central actors of the show to avoid the possibility that their ideas being imitated by the viewers. Thus, the scoring process is unrelated with televising the show.

2.2.3. EDC Startups' competition

This program is known by the name EDC startups' incubation since the program combines competition with incubation. The program was designed in 2020 by EDC head office in Addis Ababa with a plan to conduct it every year. The first version of this competition was implemented in 2021 by the Entrepreneurship Development Center (EDC), an organization formed in 2013 by the Ethiopian government in collaboration with UNDP to implement the entrepreneurship development programs in the Country.

The program aims to hunt innovative business ideas and start-up companies through business plan competition and help them establish, survive, and expand their businesses by providing various business development supports, seed money, and facilitating linkages. The call for the first version of this competition was announced by various media throughout the country and applicants had the option to apply either online or paper-based modalities through EDC's regional offices located in Addis Ababa, Amhara, Oromia, Tigray, and SNNP which constitutes more than 90% of the population in Ethiopia. The same application form, a brief baseline questionnaire, and evaluation criteria were prepared by the head office and used by all regional offices. While the competition was conducted on cluster or region basis, every procedure and interventions are standardized.

The first version of the competition attracted more than 295 applicants in total (143 in Addis Ababa, 65 in Amhara, 21 in SNNP, 26 in Tigray, and more than 40 in Oromia). The first version of the competition was conducted in 4 of the regions (or competition centers) by excluding Tigray due to the ongoing conflict in the region. Using 5 major pre-determined set of criteria consisting of 20 subcomponents, the experts of the regional offices scored the business plans of business ideas of applicants as the first screening of the competition. The criteria used in various stages are listed in appendix 2.1. Based on scores, 145 of the applicants were selected to the next stage and offered with the training slot, of which 107(73.7%) showed-up to the training. Like that of Bruh, there are no crossovers here while there are some causes of no-shows, which implies the possibility of sharp regression discontinuity design is ruled out and the fuzzy RD model becomes appropriate candidate.

These entrepreneurs were offered with a standard intensive Entrepreneurship Training Workshop (ETW) for 6 days (48 hours). The training covers a wide range of topics in entrepreneurship competency, business model/plan development, basic financial principles, and other related issues. Having made further screenings and offered contestants with additional training on business plan development, coaching, and counselling, a final pitch competition was held in each cluster to select regional winner. Finally, the regional winners came together in Addis Ababa and a national level pitch competition was held to select 6 grand winners. The grand winners were awarded a seed money amounting from 100,000 to 225,000 Birr (equivalent to 2500 -5000 USD) each depending on their needs and rank. In addition, the winners are entitled for one-on-one business development supports required beyond the competition period for the well-functioning and acceleration of the businesses.

2.3. The Data

This dissertation uses the same dataset for both main chapters and thus the data described in this section is commonly applicable for the entire study. In this section, I outlined the nature of the data, the data acquisition procedures with a special emphasis on the implementation of the follow-up survey, and fundamental results of descriptive statistics which are relevant for the understanding of the main results presented in chapter 3 and chapter 4.

2.3.1. **Data collection methods**

This study utilized administrative and survey data from first editions of *Bruh* and *EDC startups* business plan competitions. The data collection task was started early 2021, while the competitions were ongoing, from the administrative data by collecting and reviewing of administrative records of both competitions. By doing so, important information including profiles of all applicants, completed application forms, business plan of each contestant, the rules of the competitions, the scores given by judges for each business plan at various stages of the competition, the cut-off points of the scores used to select winners in each round, information about judges, status of each contestant in the competition (offered Vs rejected), the types of interventions each contestant got, if any, and the take-up rate the training intervention among those offered slot.

Using these records, I constructed administrative dataset consisting of treatment indicator, score (the running variable), baseline covariates, and other variables about the entrepreneurs' characteristics and their (proposed) businesses at the start of the program. In addition, I have also conducted personal interviews with program owners by visiting JCC and EDC offices in-person at the start of this project and then continuous virtual meeting with coordinators and consultants of the program. Through such frequent interaction, I learned the details of the implementation of the competitions and other necessary qualitative information required for this study.

Therefore, at that stage, the missing information to answer the research questions of this study was the outcome variables, treatment status by other substitute programs, and some variables to explore the transmission mechanisms. This is where conducting my own follow-up survey was required and to this end, I prepared the sampling frame using the entire lists of applicants of both competitions. From the total of 545 eligible applicants who got scored by the panel of judges, I found that 29 applicants were duplicates due to the fact some applicants applied to the competition in more than one project and some other had applied for both Bruh and EDC. After cleaning the list for the duplicates, my final sampling frame remains with a clean list of 516 eligible (potential) entrepreneurs and a census of these applicants were considered for this study.

Following Fafchamps & Quinn (2017) who argued that short-term impact is more appropriate for small interventions and microentrepreneurs like my case, I decided to collect the follow-up data just a year after the contestants applied for the business plan competition or equivalently about 8 months after the intervention of interest (training) of the program was completed. Hence, the follow-up survey was fielded from January to February 2022, within a month, using Computer Assisted Telephone Interview (CATI) method.

This method of data collection is appropriate in this context for at least four reasons. First, about 70% of the target groups had only business ideas, not operational firms, at the time of application and I was aware about this situation from the administrative data. It is also expected that at least some of them will remain in the same status even after a year (during the follow-up survey). Thus, it would be completely infeasible to try to physically trace these potential entrepreneurs who are scattered throughout the country to do a face-to-face interview, the other alternative method. Second, the volume of data (number of variables) required to answer the research questions are small and their measurements are simple to manage using about a 15-20 minutes-long phone survey. I also used a direct simple question to elicit information on profit and other outcomes as suggested by de Mel et al. (2009). Third, phone survey helped us avoid physical contact between enumerators and respondents amid the COVID-19 crises so that everyone stayed safe and the survey was completed smoothly. Finally, I had already secured mobile numbers of all the potential respondents from their respective application documents and at that moment mobile networks are available in all areas of the country except the conflict area of Tigray region where no applicant was part of the program in the first place.

Before embarking on the implement the survey, adequate pre-survey preparation was made. The survey questionnaire (intended for about 15-20 minutes-long phone survey) was carefully designed for the data collection process to be program-blind. I avoided to disclose for both the data collectors and respondent that I intend to evaluate Bruh and EDC business plan competitions using this data. As it can be seen in the introduction section of the survey questionnaire in Appendix 2.3, it was administrated as the usual general-purpose survey. Doing so is quite important to ensure the independence of the evaluation and credibility of the result by eliciting neutral information from the respondent about their current business status. Otherwise, respondent would associate their responses with the specific experience they had in the competitions being evaluated. Whatever feeling they had about the program, it would have likely affected their response had it been disclosed that the questionnaire is all about Bruh or EDC. In this regard, the neutrality of impact studies of many programs is highly suspicious as the data collection process is likely to be contaminated by program-specific experience of participants.

After the finalization of the questionnaire, it was translated into Amharic language to ensure its understandability for enumerators and facilitate proper paraphrasing while asked in local languages, programmed into SurveyCTO (a data collection platform subscribed for this purpose), installed to computer tablets, and pre-tested on some of the target groups. Then a survey team consisting of a programmer and survey coordinator, and 5 enumerators who are experienced in computer assisted firm surveys and proficient in local languages were formed. The survey team was provided with a 2-days intensive training on the questionnaire and survey protocols. Having made the necessary adjustment following the pilot survey and deploying the required logistics, the main survey was started in January 2022 and ended after a month in February 2022.

During the implementation of the survey, all the necessary monitoring and follow-ups were made to ensure the quality of data and reduce attrition. Every completed questionnaire is uploaded to the survey as soon as it is completed, and I had to check the collected data every day and provide immediate feedback when any omissions and errors are found. In some situations where respondents were not reached out using the available phone numbers, I had to contact them through email (they had filled email in the application form) and manage to trace many applicants who changed the phone number since application and able to get new contact information to provide for the survey team to do the interview. Though such coordinated effort, we managed to trace 494 entrepreneurs and collected high quality data. This makes the response rate to be 95.7%, a year after last contact, which is among the highest response rate ever achieved in small business follow up surveys.[5]

[5] When I compare it with other similar studies in a comparable time span for relevant group in African context, I found that response rate in McKenzie (2017) was 69.9% for new and 72.3% for existing businesses in round 1 for training sample; in Blattman & Dercon (2018) 88% after 11 months, 85% after 13 months; in Fafchamps & Quinn (2017) 84% after 6 months.

In order to reduce bias of self-reported data on my main outcome variable, which is setup a business, I triangulated the self-reported self-employment status with valid trade license for those claimed to have registered formally in the survey response. Two methods were used to collect their trade license. First, depending on the business status of respondents, enumerators were tasked to ask the respondent at the end of the interview to send them the photo or scanned copy of the trade license as email attachment or through social media platforms (WhatsApp, telegram, viber) which can send photo or file. This method was applicable for educated respondents who can use internet and own social media. The good thing is most of the respondents are educated as more than 82% of them graduated from university or college or TVET. One concern we had was it costs respondents to send files (data usage cost). As a response to this, a 50 Birr worth mobile airtime top up was paid as incentive payment for all respondents participated in the survey. Through this we collected considerable number of licenses for our verification purpose.

Second, for those respondents either unable or unwilling to send copy of their license, an independent data verifying expert was hired after the completion of the phone survey to cross-check their existence with administrative records of federal and local regulatory agencies. Using the tax identification number (TIN), business name, and other identifiers collected in the phone survey, the expert was tasked to verify the claim of operating a business with up-to-date records of ministry of trade and regional integration, The federal urban job creation and food security agency, and Addis Ababa trade bureau. Using both methods, we managed to independently verify the operation of 61.5% formally registered businesses.

Other outcome variables like employment, sales, and profit are based on self-reported data since administrative data are not available. Even in a situation where administrative data are available on business performance indicators, its reliability as compared to self-reported ones is not guaranteed. I do not expect businesspersons to report performance data like profit more accurately for local authorities than for that of researchers. There are also empirical evidences showing misreporting of business data to authorities (Pomeranz, 2015; Kumler & Verhoogen, 2020; Carrillo et al., 2017). Similarly McKenzie (2017) reported that administrative data on employment is unreliable. As a results, self-reported profit, sales and employment data are used in this study.

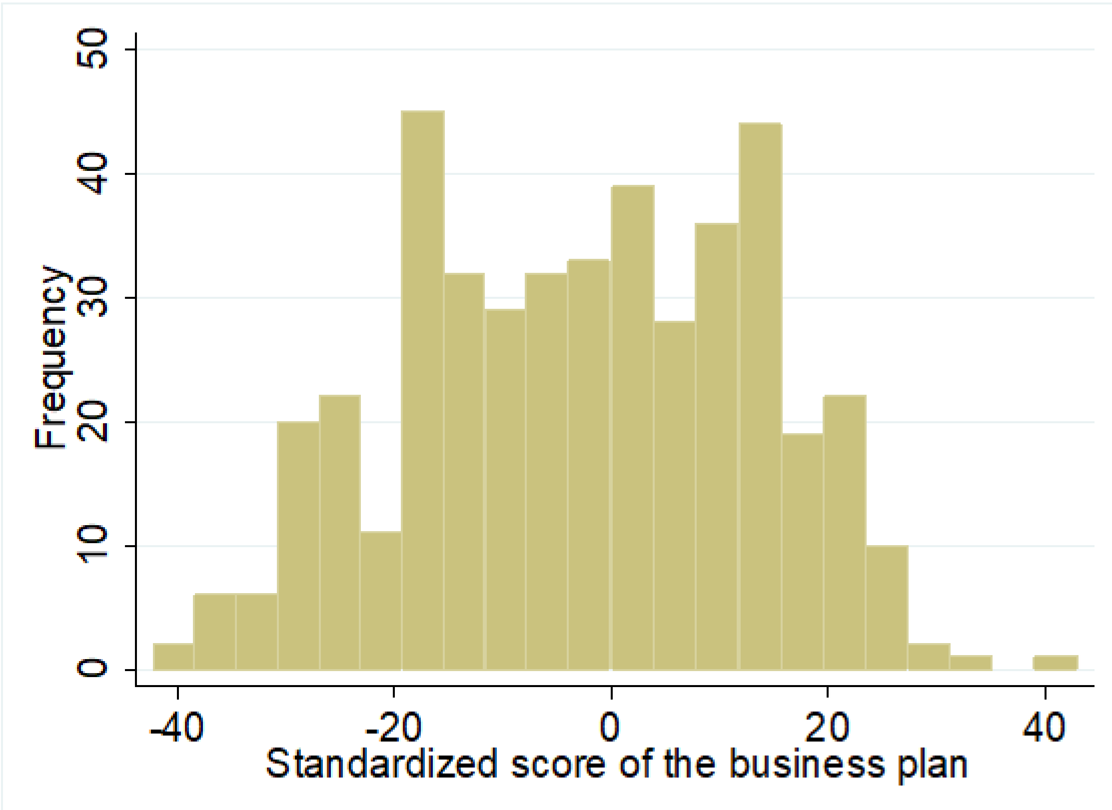
2.3.2. Description of the Data

i. Score of the business plan

Score is the one of the key variables I compiled from the administrative records of the competitions. It is used as a running variable in the chapter 3 and the regressor of interest in chapter 4 and it is worth describing it briefly. As stated in the previous sub-section, judges or experts scored the business plan of each applicants using pre-determined criteria. The score averaged over the criteria and judges were used to the training placement. In order to make the scores comparable across competitions considered in this study, I standardized the score by centering it at the cutoff.

Throughout this study, I excluded 38 applicants that were given zero score by the special decision of the committee in Bruh competition for missing information about their business model since this is an exceptional score which do not reflect their potential. This case is just the same as those excluded as illegible applicant. Thus, the whole analysis is based on the sample size of 456 observation. The distribution of standardized score for this matched sample is depicted in Figure 2.1 using a histogram drawn with the frequency distribution. In this figure and all other analysis where score is used, zero is the cutoff point for the standardized score. Applicants with a standardized score of zero and above are applicants who managed to pass the first-round screening of the competition and offered for the training slot. On the other hand, those below the score of zero (negative standardized score) are applicants who got rejected in the first screening of the respective competitions.

Figure 2.1 Distribution of standardized score of the business plan.



Notes: This score is the standardized value of the first-round screening result of each applicant of Bruh and EDC startups entrepreneurship competitions. Zero is the cutoff.

ii. Summary statistics of selected variables

Summary statistics of key variables from the baseline data and follow-up survey used in this study are presented in Table 2.1. Considering the full data matched with the follow-up survey, 51% my sample is from Bruh competition while the rest 49% is from that of EDC. About 27.8% of the applicants had operational young business at the time of application while this proportion jumps to 41.7% after a year. This shows that business ownership rate increased by about 14 percentage points.

Looking deep into the profile of applicants, most businesses are owned and managed by male and the participation of female in this respect is low. The applicants are much more educated than an average Ethiopian youth with more than 72% of have university degree. This is consistent with the business plan applicants reported in other developing countries as the competition requires paper works, more able ones self-select to the competition. Businesses which were being operated by the applicant at the time of the follow-up survey are small and young business with about, average, 7.8 number of workers and 2.5 years since operation, implying 1.5 years on average when applied for the competition.[6]

[6] One applicant reported a firm with the age of 28.6 years which could be acquired either through inheritance or purchasing of an existing business. Some applicants with a medium sized firm, in Ethiopian standard, have applied to EDC competition with the intention to get their business development supports.

Table 2. 1 Summary statistics of selected variables used in this study

Variable	Obs	Mean	Std. Dev.	Min	Max
Sample					
Case (Bruh=1; EDC=0)	494	0.510	0.500	0	1
Entrepreneur's and enterprise characteristics					
Existing business at baseline	494	0.287	0.453	0	1
Operates Business (currently)	494	0.417	0.494	0	1
Works for wage (currently)	494	0.449	0.498	0	1
Gender (male=1)	494	0.826	0.380	0	1
High school or below education	494	0.176	0.381	0	1
TVET or some College education	494	0.101	0.302	0	1
Has university Education	494	0.723	0.448	0	1
Firm age in years(currently)	206	2.576	2.995	0.083	28.6
Number of total workers (currently)	206	7.864	9.788	1	80
Sector dummies					
Agriculture	494	0.140	0.347	0	1
IT	494	0.275	0.447	0	1
Manufacturing g	494	0.310	0.463	0	1
Retail	494	0.219	0.414	0	1
Construction	494	0.057	0.231	0	1
Region					
Addis Ababa	494	0.623	0.485	0	1
Oromia	494	0.095	0.294	0	1
Amhara	494	0.194	0.396	0	1
Other regions	494	0.087	0.282	0	1

Notes: Currently refer the time of the follow-up, which is a year after the application to the competition. Means of firm age and numbers of workers are conditional on operating a business.

iii. About the training status: introducing the nominal versus effective treatment indicator

In this study, treatment is attending the training that the business plan competition offered as an integral part of the competition process right after the first screening. The total applicants offered the slot and attended the training have already been presented in the program description section. Here, let us summarize the proportion of trained entrepreneurs for the matched sample for which follow-up survey data are available.

In my matched sample, 142 of 494 (28.74%) of the respondent attended the training offered by the business plan competition and this group is my treatment group based on this treatment indicator. However, this indicator is ***nominal treatment indicator*** which considers training only within the program of interest and disregards the possibility of treatment by substitute programs as such programs are ubiquitous in the market as discussed in the context section. Therefore, I called the resulting first-stage equation estimated which is presented in chapter 3 as ***Nominal First-Stage.***

In the follow-up survey, however, in addition to measuring the business outcomes of participants respondents were asked if they had taken any entrepreneurship training in other similar programs. The very reason why I asked this is that I wanted to ensure that rejected applicants of the business plan competition (our control group) should not be treated elsewhere for this group to serve as a clean control group in this evaluation. If the rejected applicants are found to have taken similar training in other programs particularly after they got rejected from Bruh and EDC business plan competitions, they cannot be the ideal control group that I aspire to have for a precise causal estimate of the program's effect.

Surprising, in the follow-up survey, I found that about 78% of the applicants have had entrepreneurship training in any program (the business plan competition under evaluation and others). Note that this program has reached out only 28.74% of the applicants in its training intervention, implying that the remaining 71.3%. are rejected applicants that are supposed to be my control group. However, majority of this rejected applicants have ever had similar training in other programs. When I limit the timeline to be since these business plan competitions, again about 60.73% of all applicants and near to 53.69% rejected applicants of the business plan competition reported that they got at least one entrepreneurship training (Appendix 2.4, Table 2.A.2). In appendix 2.4, Table 2.A.2, I presented the cross-tabulation between treatment status in this program and treatment in any program disaggregated by Bruh and EDC; while in Table 3.A.3 the distribution of trainees who got trained by any program by the types of training provider is presented.

These results clearly show that training opportunity for startups in the market is ubiquitous and many of my control groups have been treated by substitute program. Therefore, the nominal training indicator is not reflective of one's real status since many applicants below the cutoff got trained in substitute programs in the same period even if they were rejected for the training by Bruh and EDC. What matters is getting the training regardless of who offer it. That is why I call the previous indicator as nominal indicator. Then, I constructed a new treatment indicator from the self-reported data which takes the value 1 if an entrepreneur had ever taken any entrepreneurship training and 0 otherwise. This is the true or ***effective treatment indicator*** that considers not just the treatment within the program of interest, but also other substitute treatments offered elsewhere. The first-stage estimate stemmed from this treatment indicator would be a real estimate and thus I call this first-stage the ***Effective First-Stage (EFS),*** which will be presented in the next chapter**.**

CHAPTER 3

3. Nurturing Startups: The role of trainings of the business plan competitions

3.1. Introduction

Poverty, income inequality, and unemployment are among the top challenges of developing countries. As a justification in their subsequent studies in developing countries, Sonobe & Otsuka (2006; 2011) underline the utmost importance of industrial development for tacking these problems through creating more jobs. In this respect, there seems a wide consensus that the role of startups, which are usually small in size, is pronounced and entrepreneurs are the main drivers of the economic dynamics (Gries & Naudé, 2010; Fritsch, 2008; Noseleit, 2013). Both cross-country and country specific studies confirm that small businesses have higher rates of job creation than larger firms (Ayyagari et al., 2011; Bigsten & Gebreeyesus, 2007), which could help reduce poverty and lay a foundation for industrial development in low income countries.

Again practically, it is obvious that in low-income countries like Sub-Saharan Africa where wage employment opportunities are limited, self-employment or microentrepreneurs are ubiquitous. Even when opportunities to work for large industries are available, entrepreneurship could be preferred to industry jobs, particularly when the latter one is either under remunerated or involve high health risks (Blattman & Dercon, 2018). These may justify the renewed commitments of governments, policy makers, NGOs, donors, and other development organizations operating in low- and middle-income countries to stimulate new business formation, improve the business environment, and support startups and small businesses to thrive.[7]

[7] As opposed to this widely accepted narrative, there is another strand of literature which argues the role of entrepreneurs is overestimated and supporting SMEs is just a waste of money (see Hessels & Naudé (2019) for review of this view).

Despite the increasing recognition to the roles of small business and startups in any economy, these fledgling enterprises face various constraints which primarily stem from market failures. The market failure for small firms is more pervasive associated with their size and this is referred as size-induced market failure (Vandenberg et al., 2016). The credit market failure and financial constraints as impediments for growth of small businesses are well documented in the literature (World Bank, 2008; Beck & Demirguc-Kunt, 2006; Nichter & Goldmark, 2009). It is not only in physical capital, there is also a considerable market failure to access entrepreneurial capital (Bruhn et al., 2010).

This condition requires an active policy intervention that aims at correcting these market failures and labeling the playground for small and young firms to join the market and play their role. However, as to what policy could unlock potentials of constrained entrepreneurs remains an open question in the policy and academic dialogues.

Recent empirical studies in developing countries demonstrate that business plan competition could be one of the potential policy options available to nurture entrepreneurship through directly its skill development and grant interventions or indirectly through facilitating the development and fundability of the business ideas (McKenzie, 2017; Brinckmann et al., 2010). This intervention is a recent phenomenon in low-income countries usually targeting both startups to help them successfully establish enterprises or existing small businesses to help expand their operation. Business plan competition usually involves direct financial rewards (money prize) for final winners which could be helpful to relax the financial constraint. In some cases, the design of the business plan competition may also include non-monetary supports such as group training, coaching, one-on-one counselling or advice, networking, publicity, and others for participants depending on their advancement in a competition. The latter one is more or less similar to the types of interventions that are made by business incubators and accelerators (González-Uribe & Reyes, 2021).

However, empirical studies that disentangle the causal impact of these new and innovative forms of interventions are scant (Mckenzie et al., 2020). Most of the existing literature in this area is about the traditional business trainings due to the fact that it is one of the most widely available support programs for small firms across the globe (see McKenzie & Woodruff (2014) and Mckenzie et al. (2020) for review).[8] Nonetheless, the nature of the business plan competition is different from the ordinary training interventions by contents and modality of the training, composition of participants, and other intangible benefits with a far-reaching business implication that contestants gain during and after the competition. As a result, it needs to be evaluated as a one independent area of intervention for policy learning.

Some studies on impacts of business plan competition that address the self-selection issue have been published over the last few years, and in general, these studies show that the grant component of the business plan competitions positively affect entrepreneurship activities (McKenzie, 2017; Fafchamps & Quinn, 2017). Nonetheless, evidence on the effects of skill development (or training) interventions inherent to business plan competitions are mixed (Klinger & Schündeln, 2011; Fafchamps & Woodruff, 2017). While the previous findings about the training components mostly reported to have no detectable impact on business operations of participants, adequate explanations are not usually provided why programs fail. The literature also commonly ignores the possibility of treatment by close substitute programs for the control groups in researching impacts of such entrepreneurship development programs which could potentially cause flawed conclusion about

[8] For evaluation of managerial and Kaizen training programs in developing countries see Higuchi et al. (2019) and Higuchi et al. (2015).

programs' effectiveness. In short, the existing studies are not only inconclusive but also diverse in their nature, underlining environment and population, which is a caveat to generalization. As any under researched area, they posed more questions than answers.

This study is, therefore, designed to examine the causal effect of an entrepreneurship program in Ethiopia consisting of two national business plan competitions on startups business entry and expansion using a quasi-experimental design. These entrepreneurship competitions, called *Bruh* and *EDC* startups' competition, were conducted in early 2021 by two organs of the federal government of Ethiopia (specifically, the Jobs Creation Commission (JCC) and Entrepreneurship Development Center (EDC)) with the aim to nature high-growth potential firms. The program involves two main interventions: non-monetary support (hereafter training) for applicants who manage to pass the first screening and cash grant for final winner. The training, particularly for Bruh, was provided intensively for about one-month within a bootcamp, where qualified contestants, trainers, and facilitators camp together in a dedicated facility. The entire process of the program is a combination of incubation and competition (thus, *incupetion*).

The panel of judges scored business plans of more than 500 eligible applicants of the first editions of the competitions and placement to the training was determined based on average score. By exploiting the business plan scores and exogenous cut off points of the competitions, I used the fuzzy regression discontinuity (RD) design to disentangle the causal effect of training component of the program by comparing applicants just above and below the cutoff.[9]

---

[9] In this study, I am interested in the training components of the program as the other arm (grant) had too small beneficiaries to consider it for quantitative evaluation.

A year after the application to the business plan competitions (or about 8 months after the completion of the program), I traced the universe of applicants and collected data, mainly on their business outcomes, to evaluate the short-term impacts of the program. The first-stage estimates of the model shows that marginally scoring above the cutoff (offered for the training slot) increases the training probability by about 60 percentage points more as compared to the rejected applicants who scored below the cutoff. This is a ***Nominal First-Stage (NFS)*** as it only accounts for treatment within the program of interest.

However, tracking the related training exposure of all applicants in the follow-up survey revealed that rejected applicants of the business plan competitions (the intended control group of the study) had similar training by other substitute programs in the same period. Then I ran the first-stage regression by considering treatment (training in this case) by any programs and I called it ***Effective First-Stage(EFS)***. Surprisingly, the strong first-stage we have seen in the nominal case disappeared in the estimation of the effective first-stage.

As a result, the reduced-form estimate become negligible, implying that I did not find any improvement of business outcomes of offered applicants of the business plan competition as compared to their rejected counterparts. However, it does not mean that the program is ineffective. Given the substitute treatment of the control groups, the result is not informative about the program effectiveness. This study provides a caveat for any impact study to consider the possibility of substitute treatments of the control group before claiming any causal result and concluding about effectiveness of policies and programs. It is essential to estimate the effective first-stage from the self-reported treatment status and see its significance before attributing any change of outcomes to a program’s intervention.

The chapter contributes to the existing literature at least in three arenas. First, this study provides additional evidence on the current scant literature about the experimentation of such new entrepreneurship policy options to foster entrepreneurship in developing countries. The unique features of this program, particularly the bootcamp and *incupetion* aspect, could give a good lesson regarding the alternative and innovative delivery mechanisms of entrepreneurship training. Second, as to my knowledge at least in this area, this study is the first to deliberately document evidence about the substantial take-up of substitute programs and in fact the treatment of control group elsewhere. As a result, this situation could be taken as a plausible reason for the negligible impact estimates of the entrepreneurship training programs reported everywhere even under experimental and quasi-experimental settings.

Third, the study also contributes to the improvement of future research design of impact studies by introducing the issue of effective first-stage and indicating the need for tracking subjects for similar treatments after the baseline to ensure validity of the counterfactuals. In addition, to improve the credibility of the evaluation through minimizing self-reporting bias in the follow-up data, I made the data collection process a completely program-blind whereby it was not mentioned for enumerators and respondents that the survey was meant for evaluating a specific program. Further, I cross-checked the self-reported data on the main outcome variable (business entry) with the administrative data from local authorities. These additions make the study to have valuable contributions in the area of entrepreneurship development policy.

The remining section of the chapter are organized as follows. Section 2 presents the review of related literature. The third part of the chapter outlines the empirical strategy and the test results of the basic identification assumption of the fuzzy RD model. In the fourth part of the chapter, I presented results and discussion. The last section makes concluding remarks.

3.2. Literature Review

This study is broadly related to the entrepreneurship and SME policy debate including literature in entrepreneurship training (Mckenzie et al., 2020; McKenzie & Woodruff (2014), for review), roles of entrepreneurial ability (Hessels & Naudé, 2019, for review), and grants to microentrepreneurs (Buera et al., 2020 , for recent review); and more specifically, it lays in the business plan competition and startups' acceleration and incubation programs literature (McKenzie (2017), Fafchamps & Quinn (2017), Klinger & Schündeln (2011), González-Uribe & Reyes (2021), Lall et al. (2020), and Bone et al. (2019)).

This section summarizes the existing literature that apply reliable identification strategies and provide evidence on causal effects of business plan competition and acceleration programs around the world. While empirical evidences on traditional business training or access to various forms of finance are ubiquitous, there are a dearth of studies on the emerging and innovative programs of entrepreneurship development including business plan competitions, acceleration, and incubation programs. Mckenzie et al. (2020) in their review also underscore the lack of empirical studies in this area. Table 3.1. summarizes the existing studies related to evaluation of business plan competitions (or sometimes called entrepreneurship competition) and acceleration programs with a special emphasis on their training intervention while some of the major ones have been highlighted as follows.

By conducting a large scale experiment, McKenzie (2017) evaluated the first round of a generous business plan competition in Nigerian called *YouWin* where winners were awarded about 50,000 USD, on average, upon achieving certain milestones and a 4-days entrepreneurship training for applicants qualified the first screening. The study documented that, three years after the competition, winning the grant had a larger effect on business entry and survival, employment and other outcomes while the training component of the business plan intervention was ineffective.

Likewise, Fafchamps & Woodruff (2017) did not find a significant effect (even negative impact on aggregate growth of firms though insignificant at conventional level) on a 5-days standardized training and a 1-2 days consultancy for small firms participated in a business plan competition in Ghana. Fafchamps & Quinn (2017) also conducted their own business plan competition called *Aspire* in three African countries (Ethiopia, Tanzania, and Zambia) in which winners were given 1000 USD without any condition and investigated whether winning the prize improves business performance. By pooling data from the three countries, they compare the performance of 39 winners with 82 runners-up using the regression discontinuity design. Their study reveals that, six months after the program, winners employ 2 additional permanent workers and 33 percentage points more likely to be self-employed as compared to the runners-ups. Nonetheless, this competition did not have training intervention at all, and this cannot provide any lesson for skill constraints and the role played by business plan competitions. Despite their effort to pool data from the three countries and use appropriate statistical techniques to address the small sample problem, the sample size seems still a concern.

One more study which is worth mentioning is business plan competitions held by TechnoServe in three different countries in central America (El Salvador, Guatemala, and Nicaragua) during 2002–05. The 3-7 days training programs of these competitions were evaluated by Klinger & Schündeln (2011) using quasi experimental approach. Their finding shows that the first-round group training was more important for expansion of existing businesses while advanced training (second round) which involves a one-on-one assistance and winning the prize money significantly improved the establishment of new businesses. This program seems more related to the program being investigated in a way that it has staggered training interventions. Nonetheless, it is different from the program of interest not only geographically but also in terms of intensity, delivery mechanism, and other components of the interventions.

The major limitation of this paper is they designed the evaluation study four years after the programs were completed. The timing of evaluation seems inappropriate particularly for startups since short-term impacts are more appropriate for small grants and microentrepreneurs (Fafchamps & Quinn, 2017). In addition, the fact that they reconstruct the data for control groups by recall after many years is more likely to escalate the recall bias which is likely to be systematic for the control groups as the data for treated firms are real time data collected beforehand by program implementers. As a result, the choice of their outcome variable was dictated by this situation and could not evaluate impacts on key performance measures like sales and profits.

Further, recent studies also show that acceleration programs have a promising effect on the operation of early-stage venture through their intensive and more customized supports and the notable study in this regard are Lall et al. (2020), González-Uribe & Reyes (2021) and Bone et al. (2019). Other group of study outside of the business plan competition context which provided causal evidence on entrepreneurship training and other supports programs in developing countries including Higuchi et al. (2019), Blattman et al. (2022), and Blattman & Dercon (2018) have also mixed results on the effectiveness of the supports (see Table 3.1 for details).

In sum, majority of the existing evidence on entrepreneurship training is an ordinary business training usually provided to average MSEs while the stepwise training offered as part of the business plan competition or in acceleration programs differ in its intensity, target group, delivery mechanism, existence of complementary supports, and other features. This needs to be evaluated as a different area of research. The review the existing literature also asserts that so far there is not any study on impacts of *incupetion* programs in a single experiment. The few dependable studies on business plan competition are few and diverse in their nature which make them difficult to generalize. More importantly, most of the previous studies directly dealt with the program of intervention in their survey which automatically becomes clear for the respondents that the survey is all about the program of interest. This is likely to make the respondents to respond differently and thus the true impact of the programs might have not been found.

On the top of all, previous studies disregard the possibility of treatment by substitute programs. Most researcher who failed to find significant impact on training did not provide a plausible reason for this. It could be because their control group (rejected applicants) got similar training elsewhere, as it is uncovered in this study. It is inconceivable for the control group at the start of the program remained untrained by any program in the same setting for many years that the evaluation claim to cover after the intervention. That mean if we seriously track training access of control group by similar programs, we are likely to get contaminated control group which could be one of the reasons for insignificant effect of training programs we commonly see in impact studies.

Finally, the empirical studies in this area commonly suffers from small sample problem, with the exception of McKenzie (2017), despite their attempts to pool even heterogenous cases and multiple periods with the intention to increase the sample size. This study contributes its own share to fill some of these gaps in the current literature by studying an incupetion program in Ethiopia as an entrepreneurship development policy option, carrying out own surveys completely independent of the program or program owners, and considering the possibility of substitute treatment by similar programs while trying to disentangle the causal effect of training components of Bruh and EDC business plan competitions.

Table 3. 1 Summary of literature on impact of business plan competitions, accelerators, and related interventions

S.No	Author(s)	Case/country	Type Intervention(s) studied	Outcome of interest	Methods	Major findings
1.	McKenzie (2017)	First round of YouWin business plan competition (BPC), Nigeria, which attracted 23844 (3614 are existing) businesses	-Grants, approximately US$50,000 for each winner paid out in four tranche payments conditional on achieving basic milestones -a 4-day business plan training course for top 6000	-whether set up a business and have it subsequently survived (for new) and survival (for existing) -own employment -Total employment (owner, wage and salaried, daily and casual workers but unpaid excluded) -Dummy for surpassing 10 employment -Innovation index (12 measures) -Monthly profit and sales with different measures	-track applicants for 5 years -By 475 regional and overall winners from the experiment, 729 additional winners were randomly selected from a group of 1,841 semifinalists -He used 3 rounds of data, with the round collected 12 to 18 months, and 27 months after receiving the first and final payments , respectively. -Then, he implemented RCT for experimental sample; 4 days training evaluated using RD design, and the impact on non-experimental winners evaluated using PSM.	Three years after applying winning the grant had large effect on: -New firms: 37 and 23 percentage points more likely to operate a business and cross the 10- employment threshold, respectively. Existing firms: 20 and 21 percentage points more likely to survive and cross the ten-employment threshold, respectively, than the control groups. -Winning also leads firms to be more innovative, more profitable, large sales **Mechanisms:** grants cause to purchase more capital and hire more labor; But it did not affect business networks, mentors, self-efficacy, or uses of other sources of finance -Training did not have any effect
2.	Fafchamps & Quinn (2017)	Three African countries (Ethiopia, Tanzania, and Zambia) with a total of 750 applicants of	- Prize of US$1000 to spend at his or her discretion (unconditional)	-Dummy for self-employment - Firm performance: number of permanent employees, average sales over the last month, average costs, self-reported profits,	-12 candidates evaluated by a committee of judges and each committee selected one winner. They use RDD and compare winners with the two runners-up in each committee. The restricted sample pooled from the 3	Winners are 33 percentage points more likely to be self-employed after 6 months; employ 2 additional permanent workers, have better performance, and larger firm size.

S.No	Author(s)	Case/country	Type Intervention(s) studied	Outcome of interest	Methods	Major findings
		which 481 have actually participated		and profits calculated as sales minus costs	countries consists of 39 winners Vs 82 runners-up (16 winners and 30 runners-up in Ethiopia, 16 winners and 31 runners-up in Tanzania, and 7 winners and 21 runners-up in Zambia) Data: they collected the follow up just six months after the treatment and this timeline is argued to be appropriate for such small sizes of grants and enterprises.	
3.	Klinger & Schündeln (2011)	Business training programs that the NGO TechnoServe held in Central America (El Salvador, Guatemala, and Nicaragua) during 2002–05	3 treatments - being in the first training program (including other training and finance recipients) - additional training, conditional on having been in the first training - monetary prize for winning the competition, which sums US$6,000 to US$15,000 (depending on country and year) (conditional)	- new business launched or existing business expanded (combined) -starting a business, -significantly expanding a business (measured as binary variable as perceived by the owners)	Fuzzy RDD for the trainings and sharp RDD for the monetary prize. - They got robust result for using a window of two standard deviations, and a window of 0.5 standard deviations	-total effect of the trainings (both rounds) led to 17 to 22 percentage points increase in the probability of opening or expanding a business; 25 to 56 percentage points rise in probability of expansion (separate), but insignificant effect (4-9 percentage point) effect on launching new business. -First round training (group training) is more important to expand existing business than to establish new ones. - second round training (more of one-on-one and involves development of full business plan) has a larger and significant impact to start new business. -Getting the prize money had a significant effect on launching new business, but negligible impact to expand the incumbent ones.

S.No	Author(s)	Case/country	Type Intervention(s) studied	Outcome of interest	Methods	Major findings
4.	Fafchamps & Woodruff (2017)	Business plan competition in Ghana participating 140 existing firms with 2-20 employees.	a five-day standardized training (for 70 firms) and customized consultancy service for 27 firms	Aggregate growth measured 1 and 2 years after the meeting of the panel	RCT (randomizing the probability of winning the training based on the quartile of the ranking by the panel)	Training has no effect on growth; even the coefficient is negative and significant at 0.15 level. Assigning to training increased firm exit by 7 percentage points.
5.	González-Uribe & Reyes (2021)	Business accelerator (ValleE) in Colombia	Group training, customized advice and visibility	Revenue	Used 675 firm-year observations from 135 applicants where the last observation is 3 years after the application. They employed IV constructed from exogenous differences in judges' scoring generosity	After 3 years, participation in the accelerator increases annual revenue by $20 K USD. They reported this effect is equivalent to a 166% (130%) increase from the rejected applicant (average applicant) revenue. Customized advice and visibility were more impact full than group training.
6.	Lall et al.(2020)	Accelerators around the world (members of Aspen Network of Development Entrepreneurs)	Training workshop, individualized training, technical assistant, mentorship, networking	Level of equity investment reported on follow-up surveys	Used data from 1647 entrepreneurs who applied to 77 impact-oriented accelerators between 2013 and 2016. They exploit the exogenous variation in the starting period of acceleration program within a year which exogenous variation in number of treated and untreated months after selection to the accelerators.	In the first follow-up year, beneficiaries of accelerator programs attracted more outside equity investment than rejected applicants. This promising positive effect is not observed for women owned ventures and venture located in emerging markets.
7.	Bone et al. (2019)	Business accelerators and incubators in UK (London)	-attending accelerator	-survival (measured by continued online presence), -employee growth, and -funds raised.	Used data from 5 cohorts of 638 startups applied for business accelerator during 2013-2016 period. The estimated the LATE using fuzzy regression discontinuity approach (RD) that exploits the top-20 interview threshold rule	-Accelerators have positive impact on startups performance - it has also positive spillover effect on the wider business ecosystem

S.No	Author(s)	Case/country	Type Intervention(s) studied	Outcome of interest	Methods	Major findings
8.	Higuchi et al.(2019)	Management training including Kaizen or lean production in Tanzania	-class room training for 40 hrs -on-site training -Class room +on-site training	-management practice -sales revenue -value added	RCT on 113 existing small manufacturers with at least 5 hired workers in Tanzania Interventions were made in 2010; follow-up surveys were conducted in 2011, 2012, and 2014 while baseline was done early 2010.	-Intervention improved management practice and overtime trainees keep practicing only selected once and drop the rest -business performance is improved by the combined program in the medium run, but not in short run. Sufficient assimilation required for training to bring impact on performance
9.	Mckenzie et al. (2020)	Meta analysis of various entrepreneurship training impacts around the world to summarize what we know till date	-Traditional entrepreneurship training (in-class training) - Personal initiative and heuristic training -Kaizen -Consulting incubators and accelerators -Mentoring -Matching firms with well-performing peers -alternative delivery methods (online training, television edutainment, and SMS messages)	-sales -profit	Meta analysis on various published works that applied experimental or quasi-experimental methods	*Traditional training*: modest positive effect on business practice and outcomes for microentrepreneurs -*Personal initiative and heuristic training*: effective for micro entrepreneurs. *Kaizen*: promising result for manufacturing firms above subsistence level *Consulting:* works for medium and large size firms, even for smaller firms with 14 average workers -*Incubators and accelerators:* evidence in developing countries are scarce; it is not clear as to which component is relevant -*Mentoring*: good for advanced firms working to innovate as substitute for training -*Matching firms with well-performing peers*: seems effective but the quality of the peer matters

S.No	Author(s)	Case/country	Type Intervention(s) studied	Outcome of interest	Methods	Major findings
						Alternative delivery methods: no enough evidence
10.	Blattman & Dercon (2018)	Ethiopia	Five days of business training and planning, followed by an unconditional cash grant of nearly 5,000 birr, or roughly $300	Income *Employment/Occupational Choice* *Physical Health* *Mental Health and Happiness* *Other indirect outcomes measured after a year.*	A randomized control experiment on near to 1000 job applicants in 5 firms where 304 were assigned to receive a job offer, 285 to receive the entrepreneurship program, and 358 to a control group. They estimated intention – to-treat (ITT) and complier average treatment effect (CATE) various outcomes using two rounds of data collected 11 and 13 months after the experiment	A year after the program, being in the entrepreneurship arm raised income by a third. The provision of training and initial capital caused young entrepreneurs to shift from casual labor and industrial work to their own farms and petty business. This also decreases their probability to work in formal industries. Unlike the industry job, the health cost of self-employment is negligible. The result also demonstrated that industry jobs in Ethiopia is not attractive as it is less rewarding and more hazardous. They ruled out the hypothesis that entrepreneurship is either undesirable or likely to involve high risks of income.
11.	Blattman et al. (2022)			The same outcome measured after five years (long-term)		Outcomes of the three arms converge in the long run. Most recipients of the grant exit their businesses. The short-run increases in self-employment, productivity and earnings and the negative health effect observed after a year have dissipated over time (after 5 year). The interventions had no long-term effect.

3.3. Empirical strategy

3.3.1. Specification of the model

This chapter intends to examine the causal impact of the program's intervention on self-employment (or business establishment) and business expansion (measured by employment, sales, and profit). Throughout this study, I will be limited to the training intervention of the program since the recipients of the other intervention of the program, that is grant, are too small (just 26) to do meaningful quantitative impact evaluation for this arm. As discussed in chapter two, in both competitions, judges scored business plans of each applicant based on pre-determined criteria and average score of the first-round screening was used to determine placement for the training program.

The competition organizers had determined their admission capacity to the training beforehand while the judges were tasked to score all the eligible applicants. Then, the organizers invited top applicants for the training based on their score and availability for the entire duration of the training period starting from the highest scorer until their capacity is filled. For instance, the training program in *Bruh* was designed to be offered in the bootcamp for one month and JCC had already determined to admit top 70 applicants based on their first-round screening result. The judges scored 275 applicants and delivered to JCC. JCC started inviting applicants starting from the applicants ranked first and during invitation applicants were required to commit one-month full time for the bootcamp. At this stage, some applicants who were among the top 70 declined the offer for various reasons and JCC replaced them with the next best applicants again based on their score. Through this process, JCC had to invite top103 applicants to fill the 70 quota and 63.05% was the cutoff. Similarly in EDC, each center determined the number of applicants admitted for the training and the scoring was done by the respective centers, each center had its own cutoff.

Therefore, as a rule of the competitions, applicants above the cut-off were offered slots for the training intervention and pass to next step of the competition while those below the cut-off did not have that chance to access the training prepared by these business plan competitions. This allows us exploit the scores given to each contestant and the exogenous cut-off points used to select best applicants to estimate the causal effect of the training interventions using regression discontinuity (RD) technique.

Closely looking at the implementation of the training program, there was no crossovers, meaning that all applicants below the cut-off did not receive for the training intervention (thus, they perfectly complied). However, some of those above the cut-off did not show-up because of various reasons as discussed in chapter 2. This implies there was imperfect compliance above the cutoff and thus the attainment of the training program is not a deterministic function of score. This lends itself to a fuzzy Regression Discontinuity Design (RDD). Thus, in this chapter, I employed a fuzzy RDD to identify the causal impacts of the training intervention through comparing offered applicants who are just above the cutoff (treatment group) and rejected applicants just below the cutoff (control group). Considering the treatment in substitute program as a crossover case and with the fact that there are cases of no-shows, the estimated causal parameter would be the Local Average Treatment Effect (LATE).

Our generic fuzzy RD model is specified as three sets of equations as follows.

The Outcome equation:

$$Y_{ij} = \beta T_{ij} + f\left(S_{ij} - \bar{S}_j\right) + \delta X_{ij} + \upsilon_{ij} \tag{3.1}$$

The First stage equation:

$$T_{ij} = \pi \mathrm{I}\left\{S_{ij} - \bar{S}_j \geq 0\right\} + h\left(S_{ij} - \bar{S}_j\right) + \theta X_{ij} + \nu_{ij} \tag{3.2}$$

Reduced-form equation:

$$Y_{ij} = \beta\pi\mathrm{I}\{S_{ij} - \bar{S}_j \geq 0\} + g(S_{ij} - \bar{S}_j) + \gamma X_{ij} + \varepsilon_{ij} \qquad (3.3)$$

Where the subscript *i* represents applicant and *j* indicates the competition centers based on which the jury and the cutoff vary (Bruh and 4 centers of EDC separately); Y_{ij} denotes the outcome variable of applicant i competed in competition j or evaluated by a jury *j*. Outcome variables includes measures of business entry and survival (operating a firm) as well as expansion (total numbers of workers, monthly sales in Ethiopian Birr, monthly profit) which were observed 8 months after the intervention. T_{ij} represents the treatment indicator which takes the value 1 if an applicant attended the training and 0 otherwise; S_{ij} is the average score in the first screening and $\bar{S}_j$ the cutoff point to be admitted for the training program; $S_{ij} - \bar{S}_j$ is the average score centered at the cutoff which is my running variable; $\mathrm{I}\{S_{ij} - \bar{S}_j \geq 0\}$ is an indicator function for applicant *i* to be above the center j's cutoff; $f(.)$, $h(.)$ and $g(.)$ are the polynomial functions of the standardized score; X_{ij} is a vector of exogenous controls; β is the parameters of interest for the outcome equation (causal parameter) that measures effect of the training on business outcomes; and π represents the first-stage parameter which captures the effect of qualifying for the training (i.e. scoring above the cutoff) on training participation. The product of the two parameters (βπ) gives the reduced form estimate, another parameter of interest in eq(3.3). If the first-stage is negligible, the reduced form is expected to be small for any value of $\beta.\ \theta, \delta, and\ \gamma$ are coefficients of the control variables; and $\upsilon_{ij}, \nu_{ij}\ , and\ \varepsilon_{ij}$ are the error terms.

In estimation of these equations, an optimal bandwidth selection method recently developed by Cattaneo et al. (2020) and Cattaneo et al. (2021) is utilized to avoid the bias that could come from the subjective selection of bandwidth.

3.3.2. Test of the basic identification assumption: Manipulation test of score

Before presenting results from estimation of the model, the validity of the basic identification assumption of the RDD should be tested in my data. The basic identification assumption of the model is that entrepreneurs who are just above or below the cut-off are similar in their observable and unobservable characteristics, implying that participants are unable to manipulate the running variable (the score). This is also referred in the literature as continuity assumption (Lee & Lemieux, 2010). If the scores are not manipulated, the training probability function is expected to be smooth at the cutoff. Since I have transformed the score as indicated in the specification of the model, the cutoff in this study is zero.

Using the first-round screening scores data obtained from the administrative records of *Bruh* and *EDC*, I tested the plausibility of this assumption using two methods: Falsification test on pre-determined covariates (Lee, 2008; Lee & Lemieux, 2010), and the density test (McCrary, 2008 and Cattaneo *et al*., 2020). The test results of each method are presented as follows.

3.3.2.1. **Falsification test on pre-determined covariates**

The first simple method available to test the continuity assumption is that the falsification test of pre-determined exogenous covariates (Lee, 2008; Lee & Lemieux, 2010). Accordingly, the distribution of the pre-determined covariates must be continuous at the cutoff if scores are not manipulated. This is analogous to the balance test of treatment and control groups using their baseline characteristics in case of random experiments.

I conducted the test using the entrepreneur and enterprise characteristics data which were collected at the baseline (application period) as pre-determined exogenous covariates. These includes the entrepreneur's gender (dummy for female owner), education (measured as three categories: high school or below, Technical and Vocational Education & Training (TVET) or some college level education, and undergraduate or graduate degree), dummy for having existing business at the time of application, sector dummies of the (proposed) businesses applied for the competitions, and regional dummies. Using these exogenous covariates as outcome variables and the running variable (score) as a regressor, I estimated the RD estimators of the coefficients with data-driven automatic bandwidth selection proposed by Cattaneo et al. (2020) by allowing variations in bandwidth to the left and right of the cutoff.

The results of this test for each outcome variable are summarized in Table 3.2 for the full sample as well as for Bruh and EDC sub-samples. As shown in Table 3.2, the estimated coefficients in almost all the models are not statistically different from zero, suggesting that the score is continuous at the cutoff and continuity assumption is satisfied.

Table 3. 2 Falsification test results of the running variable based on pre-determined covariates

Dependent Variable	Full sample	Bruh	EDC
Female owner	-0.0279	0.0857	-0.187
	(0.110)	(0.126)	(0.193)
High school lor below education	0.185*	0.142	0.502
	(0.102)	(0.168)	(0.324)
TVET or some college level	-0.0136	0.186*	-0.303
	(0.0806)	(0.112)	(0.186)
Undergrad or grad degree	-0.170	-0.366**	0.0950
	(0.118)	(0.181)	(0.209)
Existing business at the application	-0.0921	0.128	-0.435*
	(0.107)	(0.119)	(0.226)
Manufacturing sector	0.0156	0.352*	-0.230
	(0.119)	(0.212)	(0.243)
Construction sector	-0.0323	-0.104	0.165
	(0.0759)	(0.102)	(0.100)
Agriculture sector	0.155*	-0.0914	0.306*
	(0.0925)	(0.0705)	(0.171)
IT sector	-0.0204	-0.0623	-0.0410
	(0.111)	(0.148)	(0.175)
Retail sector	-0.140	-0.174	-0.118
	(0.126)	(0.195)	(0.203)
Addis Ababa	0.142	0.196	0.0962
	(0.151)	(0.210)	(0.205)
Oromia region	-0.00281	-0.0830	0.103
	(0.0723)	(0.114)	(0.0708)
Amhara region	-0.0723	-0.0005	-0.232
	(0.124)	(0.150)	(0.210)
Other regions	-0.0170	-0.0583	0.0188
	(0.0789)	(0.119)	(0.139)

Notes: The reported coefficients are the RD estimates for coefficients of standardized score on exogenous covariates. For each regression, data-driven and varying optimal bandwidth to the left and right side of the cut-off (MSE-optimal bandwidth) are used. Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively.

3.3.2.2. The Density Test

The second and formal method to test the continuity of the running variable at the cutoff is the density test. This test was introduced by McCrary (2008) and recently improved by Cattaneo et al., (2020) for the generating local polynomial density estimators and by Cattaneo et al. (2021b) and

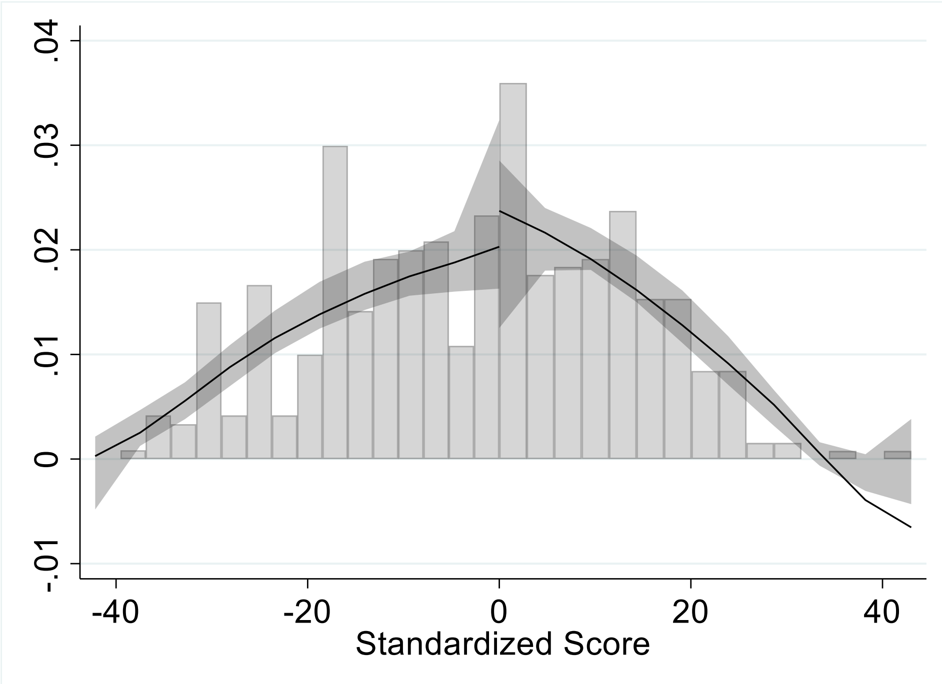
Cattaneo et al. (2022) for graphical procedures with valid confidence bands. The idea of this test is by obtaining a histogram of the running variable and see if the estimated densities obtained from the local polynomial regression separately run in the left and right sides of the cutoff are similar in both sides. The null hypothesis of this test is that there is no manipulation of the running variable or densities of the running variable is the same to the right and left of the cutoff.

The test results of this hypothesis are presented in Figure 3.1 for the full sample (panel A) and for the disaggregated one by types of competition (Panel B and Panel C). In this test, the smoothness of the distribution at the cutoff is objectively gauged from the resulting test-statistics. For the full sample, I fail to reject the null hypothesis of no manipulation with a p-value of 0.7097 implying that the density function is smooth at the cutoff. This is consistent with what we visualize in Figure 3.1, panel A. Similarly, by disaggregating the data into Bruh and EDC cases, I fail to reject the null hypothesis with a p-values of 0.4936 and 0.9656 for Bruh and EDC, respectively. This again bolsters the finding that score was not manipulated by subjects of this study.

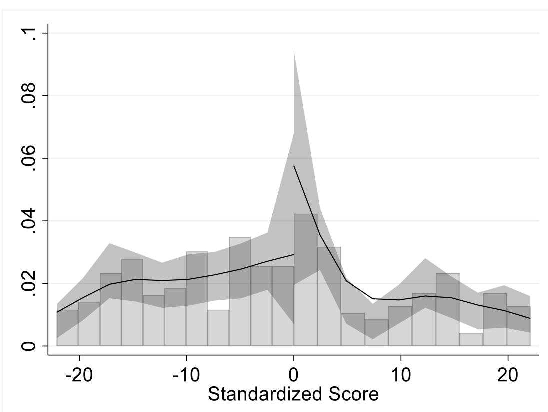
As suggested by both the statistical tests performed so far, the model passes the main identification assumption, that is, continuity of the running variable at the cutoff. This is quite consistent with the intuition of the program. The fact that scores are given by a panel of judges where the average score from all criteria and all judges determine final placement. In this situation, there is no way for a contestant to know the scores of any competitor before the result is revealed. Similarly, a member of the jury cannot know what other score is given by other members before the score is submitted for the computation of the average score. This situation rules out the possibility of manipulation of scores by the contestant as well as a member of judges.

Figure 3.1 Density Test for manipulation of scores

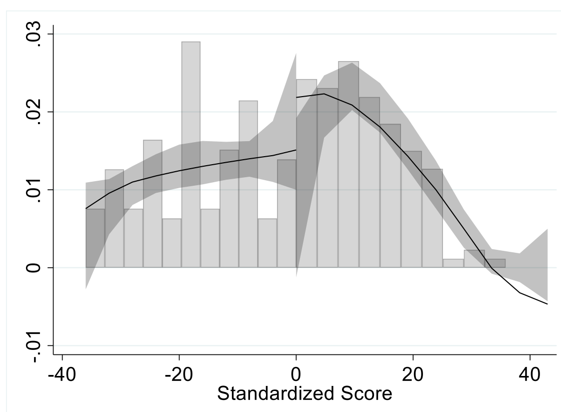
Panel A: For the full sample



Panel B: For Bruh sub-sample



Panel C: For EDC sub-sample



Note: Thes graph summarize the density test for manipulation of the running variable (score) for the full sample, Bruh sub-sample, and EDC sub-samples. The resulting P-values of the test is 0.7097, 0.4936, and 0.9656 for tests displayed in Panel A, Panel B, and Panel C, respectively. The null hypothesis is that there is no manipulation of the running variable (score).
In addition, the number of contestants who pass to the next stage is already determined beforehand.

Once the scores are averages and top scorers are selected in order of their scores until the pre-planned capacity (quota) for a given intervention is filled. That also makes the cutoff purely exogenous.

Other assumptions to identify the model are exclusion restriction and monotonicity assumptions. These assumptions are easily satisfied in this application as scores affects the outcomes through only allowing access to training of the program, for the former, and as contestants below the cutoff are unlikely to refuse to get the training, for the latter assumption. All these tests clearly imply that RD is a valid design for this study.

3.4. Estimation results of the models

3.4.1. The First-Stage Estimates

This section presents the results of the First-Stage (FS) equation specified in equation (3.2) mainly using a binned scattered graphs to make the results easily understandable while the estimated coefficients and their standard errors are also tabulated at the end of the sub-section for further reference and to summarize the results. Any causal inference analysis that attributes a change of the outcome of interest to the intervention of a program should be preceded by a strong first-stage result. Here, I will demonstrate the change in training probability at the cutoff with varieties of specifications.

3.4.1.1. The Nominal First-Stage (NFS)

The dependent variable for the first-stage of my fuzzy RD model is attendance of the entrepreneurship training (treatment) which takes the value 1 if an entrepreneur attended the training of the business plan competitions and 0 otherwise. As discussed in chapter 2, this is a nominal indicator and the resulting first-stage is the Nominal Fist-Stage (NFS). Data about the status of training attendance was taken from the administrative records of the competition.

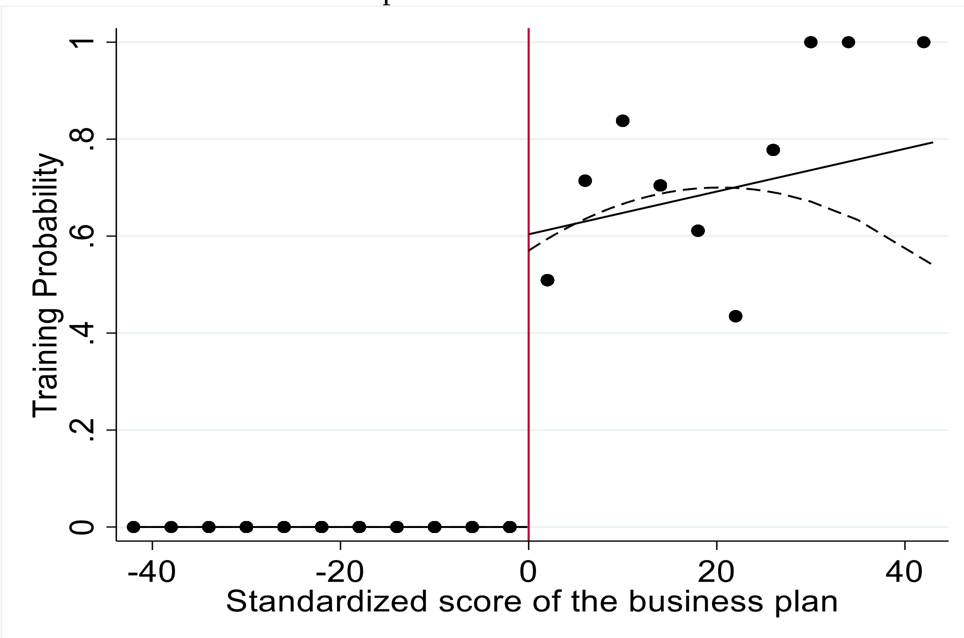
Figure 3.2 depicts the nominal first-stage result which summarizes the probability of attending the entrepreneurship training offered by the program under study (Bruh and EDC business plan competitions) as a function of score. Panel A, panel B, and panel C presents the results for the full sample, Bruh sub-sample, and EDC sub-sample, respectively. In these figures and all others presented next, the scattered dots are the bin means which are computed with a bin width of 4. The solid line and the dashed curve are the linear and quadratic regression fits separately estimated to the left and right of the cutoff (zero). Rejected applicants in the first-round screening of the competition are those below zero and their offered counterparts are those of zero and above.

This nominal first-stage estimates show that there is a clear jump in training probability at the cutoff in all the three panels. The difference in the training probability between applicants above and below the cutoff is the first-stage parameter, which is estimated to be about 60% and 57% in the linear and quadratic specification, respectively, for the full sample for instance. The point estimates of the first stage parameter for each sub-sample (full, Bruh, or EDC) are also presented in Table 3.3 which reaffirms that I have a strong first-stage as the reported parameter is statistically significant at 1% level in all cases (Figure 3.2 and Table 3.3).

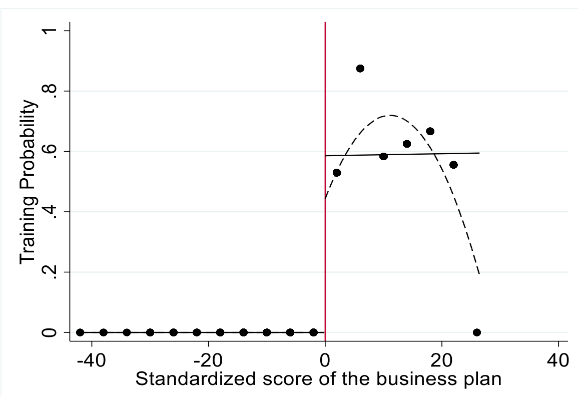
To the left of the cutoff, the regression fit lines coincide to the horizontal axis at $Y = 0$, implying that there are no crossovers in the training of the program while a less than one training probability to the right of the cutoff indicates the existence of no-shows, consistent with my discussion in the program description. The fact that score is a strong predictor of treatment (i.e. training participation) with a clear jump at the cutoff coupled with the satisfaction of the main identification assumption tested in the previous section make the fuzzy RD a perfect design for this study.

Figure 3. 1 Nominal First-stage (NFS) result of the probability of attending training of the program.

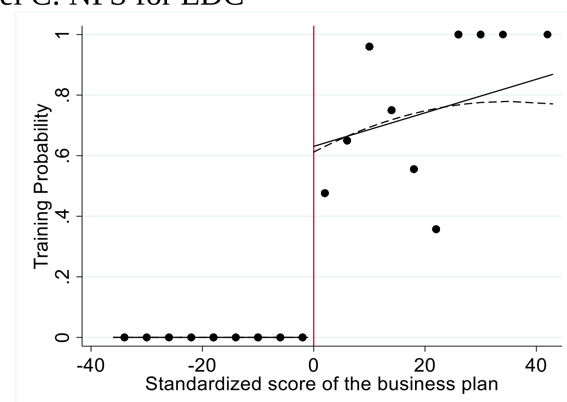
Panel A: NFS for the full sample



Panel B: NFS for Bruh



Panel C: NFS for EDC



Note: These graphs show the nominal first-stage results for the full sample (Panel A), for *Bruh* (panel B), and *EDC* (panel C) sub-samples. The dependent variable in all cases is the dummy for attending the entrepreneurship training prepared by the competition organizers. Zero is the cutoff for the running variable(score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

3.4.1.2. The Effective First-Stage (EFS)

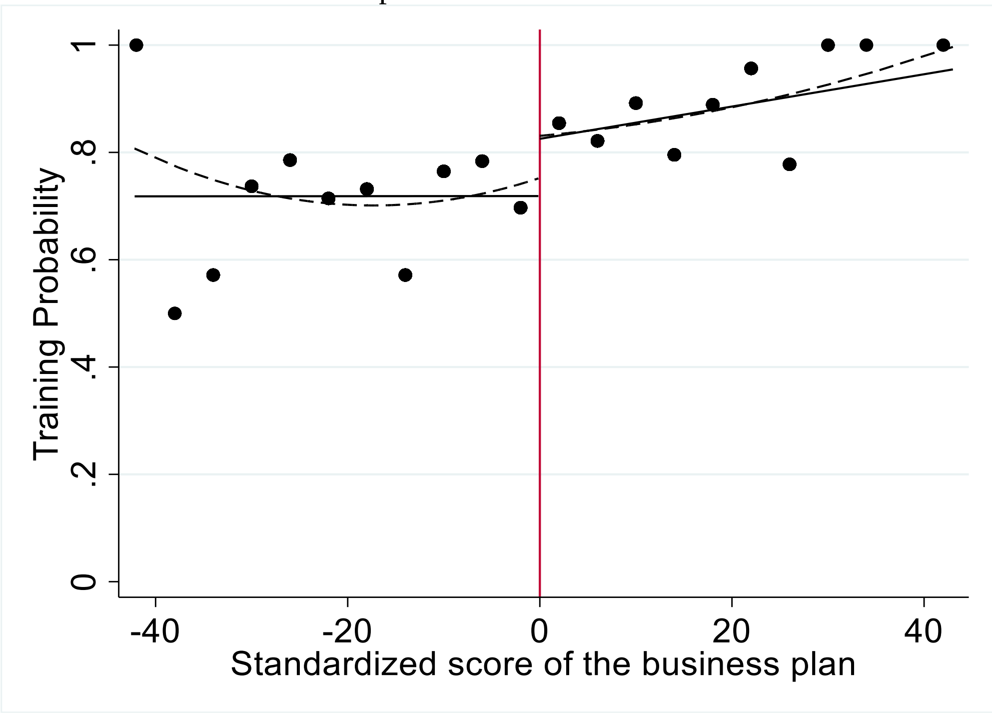
A) The main EFS result

As described above, I had the nominal firs-stage result which shows a perfect set up for RDD before going for the follow-up survey. In the follow-up survey, however, I documented the training status of all applicants in substitute programs, and I found that many of the control groups were actually treated elsewhere by substitute program as described in chapter 2. Using the effective treatment indicator which takes the value 1 if an entrepreneur had ever taken any entrepreneurship training and 0 otherwise, I have re-estimated the first-stage equation, which is the Effective First-Stage (EFS).

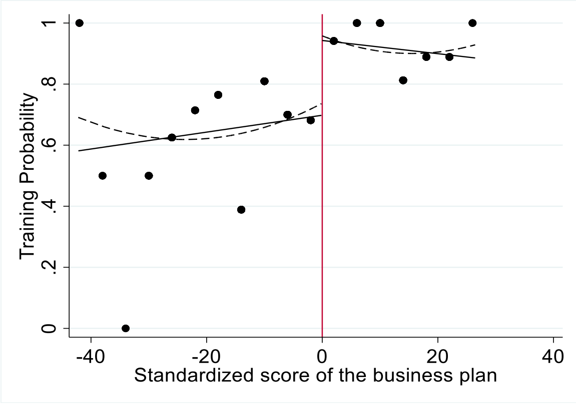
The results of the effective first-stage presented in Figure 3.3 and Table 3.4 reveal that the first-stage we observed for the nominal case disappeared when we take the substitute treatments into account. For instance, if we look at the full sample case in panel A of Figure 3.4, we find that both applicants to the right and left of the cutoff have high and comparable level of training probability. At the cutoff in the linear specification, applicants that were rejected in the business plan competition (left of the cutoff) has a training probability of more than 0.7 while it is a little bit higher than 0.8 for their offered counterparts (right of the cutoff). As a result, the estimated effective first-stage parameter dwindled to about 10 percentage points. Even this estimate reduces to 7 percentage points if we consider the quadratic specification and in both specification the coefficients are not statistically significant as shown in Table 3.4, column 1 and column 2. This implies that both control and treatment groups have the same status in terms of attending training when we consider substitute programs.

Figure 3. 2 Effective First-stage (EFS) result of the probability of attending entrepreneurship training from any program

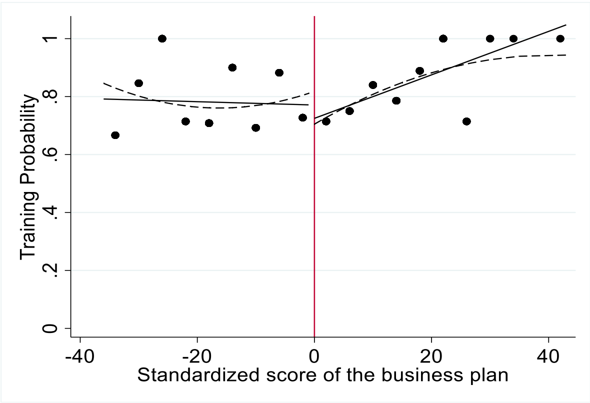
Panel A: EFS for the full sample



Panel B: EFS for the Bruh



Panel C: EFS for the EDC



Note: These graphs show the effective first-stage results for the full sample (Panel A), for *Bruh* (panel B), and *EDC* (panel C) sub-samples. The dependent variable in all cases is the dummy for attending the entrepreneurship training ever offered by any program or training provider including the program of interest. Zero is the cutoff for the running variable(score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Table 3. 3 Nominal First-Stage Estimates of the Effect of Scoring above the Cutoff on Training Attendance

	Full sample		Bruh		EDC	
	Linear (1)	Quadratic (2)	Linear (3)	Quadratic (4)	Linear (5)	Quadratic (6)
π	0.6038***	0.5700***	0.5857***	0.4434***	0.6308***	0.6120***
	(0.0548)	(0.0710)	(0.0869)	(0.1087)	(0.0707)	(0.0984)
Observations	456	456	214	214	242	242
R-squared	0.4965	0.4982	0.4536	0.4803	0.5228	0.5232

Notes: Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is Dummy for attending the training offered by the business plan competitions under study. The reported coefficients are estimates of the first-stage parameter(π); the standardized scores and its interaction with the indicator of being above the cutoff have been controlled; Estimations are based on the full support. Linear and Quadratic are linear are types of functional forms of the model.

Table 3. 4 Effective First-Stage Estimates of the Effect of Scoring above the Cutoff on Training Attendance

	Full sample		Bruh		EDC	
	Linear (1)	Quadratic (2)	Linear (3)	Quadratic (4)	Linear (5)	Quadratic (6)
π	0.1066	0.0784	0.2449***	0.2206*	-0.0460	-0.1130
	(0.0667)	(0.0905)	(0.0825)	(0.1160)	(0.1007)	(0.1444)
Observations	456	456	214	214	242	242
R-squared	0.0303	0.0314	0.0969	0.0992	0.0141	0.0161

Notes: Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is Dummy for attending entrepreneurship training offered by any program at any time. The reported coefficients are estimates of the first-stage parameter (π); the standardized scores and its interaction with the indicator of being above the cutoff have been controlled; Estimations are based on the full support. Linear and Quadratic are linear are types of functional forms of the model.

Disaggregating the analysis into the two types of competitions yields similar pattern. The strong first-stage parameter I have estimated in nominal case for Bruh sub-sample (about 59% in linear and 44% in quadratic specification from Table 3.3) has sharply declined to about 24% and 22% in linear and quadratic specification, respectively, as shown in panel B of Figure 3.3 and column 3 and column 4 of Table 3.4. Though the magnitudes of the estimated coefficients have declined, it is still statistically significant implying that the first-stage has somehow sustained for Bruh even after accounting for training by substitute programs. On the other hand, alike the full sample, the first-stage for EDC has completely disappeared, even it turned to negative though insignificant, when I estimate the effective first-stage (Colum 5 and column 6 of Table 3.4 and Panel C of Figure 3.3).

To sum up, when treatment only within the program under evaluation (the business plan competitions) is considered, as many researchers do, there is a strong first-stage for all sub-samples. Nonetheless, I argue that balance of the treatment and control group at the baseline line (time of application in my case) is not enough to ensure the validity of the counterfactual. It is also essential to check the exposure of the control groups for substitute treatments after the placement of the program. My follow-up data in this study revealed that the control group of this study when designed (rejected applicants) got trained elsewhere by substitute programs. Given this situation, we do not expect the business outcomes of training beneficiaries to be better than that of their rejected counterparts as the latter group (controls) had similar training from other trainers.

B) Effective First-Stage by contents of the training

In the effective first-stage result presented so far, I have demonstrated that both group of applicants had the same level of treatment (training) no matter where they got treated. One possible concern which could weaken this conclusion is that the two groups may not necessarily had the same contents of the entrepreneurship training. For instance, the types of training modules developed for offered applicants of the business plan competitions were more tailored for startups. On the other hand, my control groups who got entrepreneurship training elsewhere may not have had the same contents of training and its relevance to their business development may be questionable. If there is variation in the contents of the training each group covered, the effective first-stage may not necessarily be effective, and the nominal first-stage could rather be more relevant.

In order to address this potential concern, I collected detailed information in the follow-up survey regarding the types of training modules they have ever covered in the training programs they attended. By doing so, I uncovered that the respondents have had various type of entrepreneurship or business trainings which are categorized into 11 themes or modules. Comparing these 11 modules with the types of training modules covered by the business plan competitions, which are reported in chapter 2, gives us two major categories of modules. These are:

i. Modules (training contents) which are ***the same*** as the modules covered in the program of interest (the competitions). This group consists of six modules, namely, entrepreneurship competency, business idea development and business plan preparation, pitching skills, marketing, visual prototype and product development, and legal business setup
ii. Modules which have ***different content*** as compared to the trainings covered by the business plan competitions. Kaizen, technical training, bookkeeping, management training, and other trainings like time management are the types of modules included in this category.

If applicants to the right of the cutoff are found to have higher probability of accessing training of category one as compared to their counterparts to the left of the cutoff, the training that the latter group had in other programs cannot be considered as a close substitute for the training offered by the program of interest. Statistically speaking, if the groups have similar access to the same contents of this program's training, the concern we had about substitute treatment as well as my effective first-stage remain valid.

To test this claim, I defined two treatment indicators associated to the two categories of the training modules and estimated the first-stage for each types of training. For instance, related to the first category, the treatment indicator is defined to take a value 1 if a respondent has ever covered one of the six modules of category one (which are the same as offered in the business plan competitions) and 0 otherwise. Again, for the second category, an indicator for having different contents of training as compared to that of Bruh/EDC is constructed to take the value 1 if the entrepreneur has ever covered at least one of the 5 modules in category two and 0 otherwise. The first-stage results which are generated using these treatment indicators are presented in Figure 3.4.

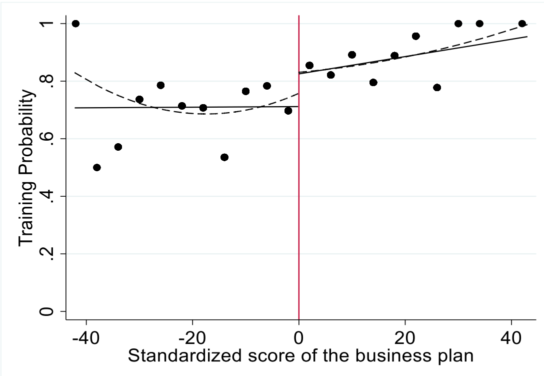
For each sub-sample, I reported a pair of graphs that compare the effective first-stages for the same content (category 1) and different content (category 2) trainings. The results shown from panel A to panel F in Figure 3.4 again confirmed the previous finding that the effective first stage is generally negligible as we do not see clear and big jumps in the training probabilities at the cutoff. To be precise about the estimated first-stage parameter, the results depicted in panel A, C, and E, for instance, entails us that marginally qualifying for the training increased access to alike-program's type training by 11.4% for the whole sample, 25.5% for Bruh, and almost nil for EDC. This result is the same as the result I reported in Figure 3.3 for the aggregates treatment indicator which disregards the contents.

This finding implies that the trainings that my rejected applicants accessed elsewhere are of the same content as the trainings offered by Bruh and EDC business plan competitions for the accepted applicants. The fact that the close substitutability of treatment by other programs is confirmed and thus results of my effective first-stage is more realistic than that of the nominal one.

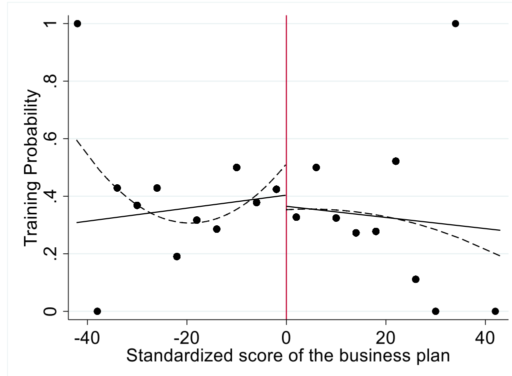
Another observation we can have from results in Figure 3.4 is that the level of training probabilities in the two categories of trainings. For instance, if we compare training probabilities in panel A and panel B of Figure 3.4 for the full sample, the training probability for the first category (panel A) is on average more than 70% while it is not more than 40% for the second group (B). This shows that the types of trainings designed and offered to the successful business plan applicants by the competition organizers is among the types of training which are widely available in the market. Had it been of a unique content, applicants to the left of the cutoff would not have reported the same level of access and content of training.

Figure 3. 3 Effective First-stage result disaggregated by contents of the training covered

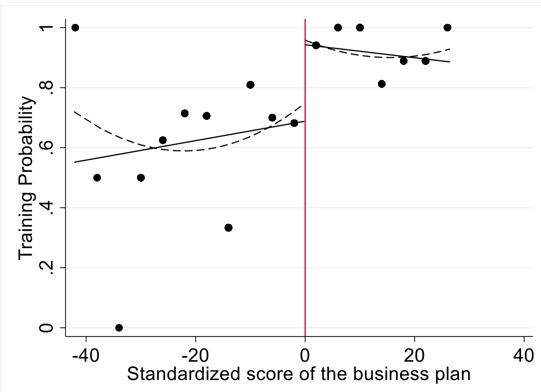
Panel A: Full sample's EFS for similar contents



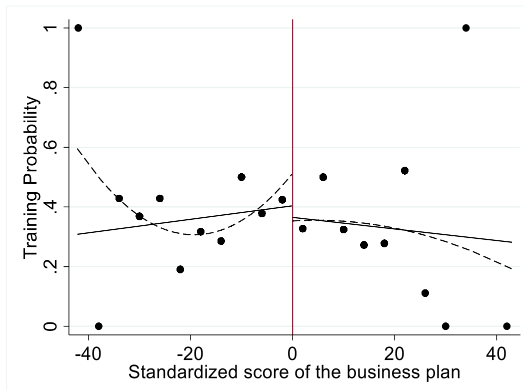
Panel B: Full sample's EFS for different contents



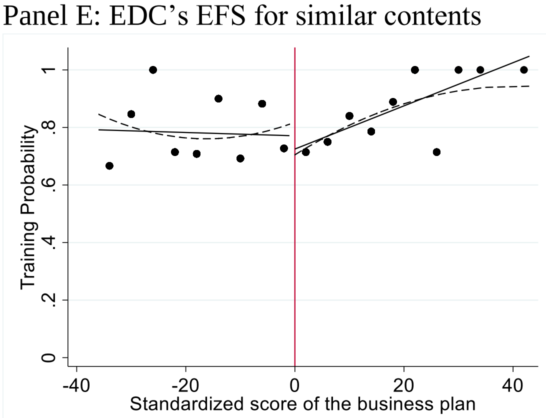
Panel C: Bruh's EFS for similar contents



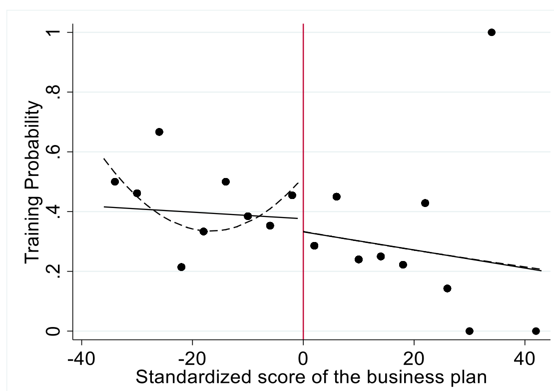
Panel D: Bruh's EFS for different contents



Panel E: EDC's EFS for similar contents



Panel F: EDC's EFS for different contents



Notes: Similar content means training modules trainees covered is the same as that of used by the program under evaluation, and different content otherwise. Dependent variable used in panel A, C, and E is a dummy for covering at least one of the 6 training modules offered by the competitions. Dependent variables used in panel B, D, and F are dummy for covering at least one module which are categorized as different from that of the program of interest. Zero is the cutoff for the running variable(score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

C) First-stage for more disaggregated measures of training content

Even though the training experience of respondent was further qualified by disaggregating it into similar and dissimilar to the contents covered by the program of interest in the previous sub-section, there has been still much aggregation in each category which could potentially mask the intra-group variations. For instance, in the previous analysis, an entrepreneur was categorized to have accessed a training of similar content like that of Bruh or EDC if she/he ever covered ***at least one*** module among the six modules in this category. This definition of the treatment indicator does not differentiate between respondents who had covered just one module from those who covered more or all the modules. This situation sparks a potential concern that the training intensity is not well captured by treatment indicator and thus the rejected applicants who got treated elsewhere could have covered only few modules while the offered applicants by default covered almost all the modules in this category.

In order to address this potential concern, I further disaggregated the analysis of the first group (similar content) into four. That is, I run separate first-stage for access of each of three major modules (business idea development and business plan preparation, entrepreneurship competency, and marketing) without any aggregation and the remaining three minor modules (pitching skills, visual prototype and product development, and legal business setup) together as '*other training of the program*'. The effective first-stage result associated with these four treatment indicators are presented in Figure 3.5 just for the full sample.[10]

[10] Results for the Bruh and EDC sub-samples are also generally the same, with a relatively higher positive first-stage for Bruh. These results are not reported here to keep the document readable.

As shown Figure 3.5, the training probabilities for each type of module (represented by panel A to panel D) are the same for both groups of applicants at the cutoff. That means rejected applicants have equally accessed trainings that offered applicants enjoyed within the program of interest. The only difference between the two groups may be the type of the training provider: the business plan competitions and other programs for the offered applicants while it is exclusive by other programs for applicants rejected in this program.

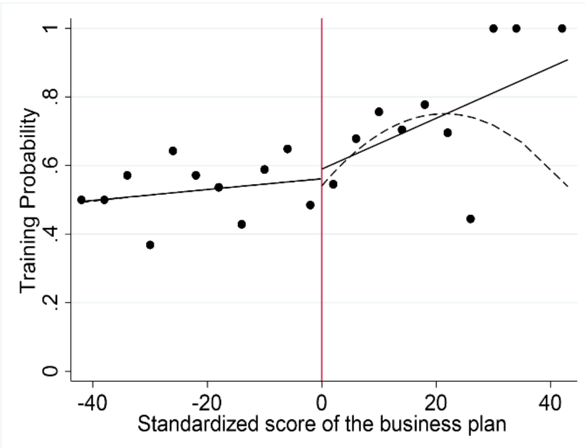
The additional results I reported here further bolster the validity of the effective first-stage results and weakens the potential concerns about it. Despite the attempts I made in this and previous sub-section to account for the intensity of the trainings while defining treatment group, I believe that treatment indicators used for the effective first-stage still have a couple of common limitations.

First, unlike the indicator I used for the nominal first-stage, the treatment indicators for all effective first-stage results are based on self-reported data. This suffers from some self-reported bias as I learnt from the cross checking of the self-reported status with the actual status in accepted applicants of this program. Second, I do not have data to measure the quality and duration of the self-reported trainings which could be important to explore.

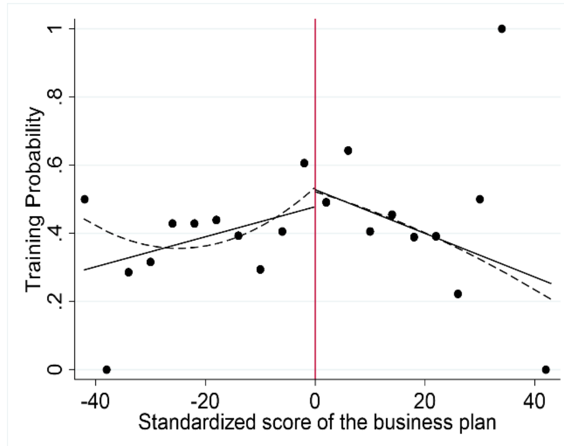
If our treatment indicator for the effective first stage is significantly imprecise due to these two limitations, we can expect a significant parameter estimate for the reduced-form equation. On the other hand, if we do not see any change in the reduced form equation at the cutoff, these limitations will have negligible effect on my effective first-stage estimate. I will prove shortly in the upcoming sections that the latter one is true. Before that let me address other concerns on the effective first-stage.

Figure 3. 4 Effective first-stage for covering major training modules of the program separately

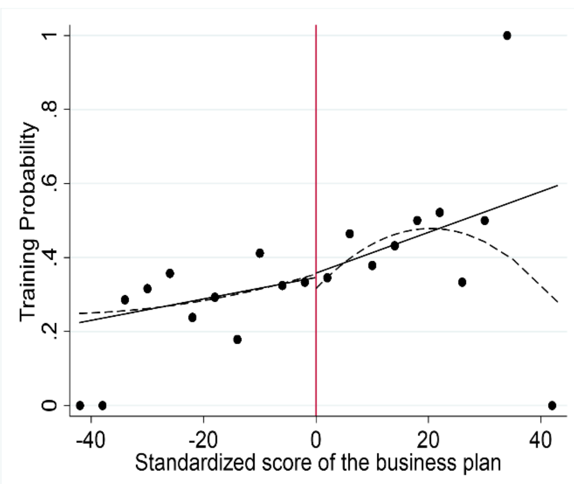
Panel A: EFS for Business idea development and business plan preparation



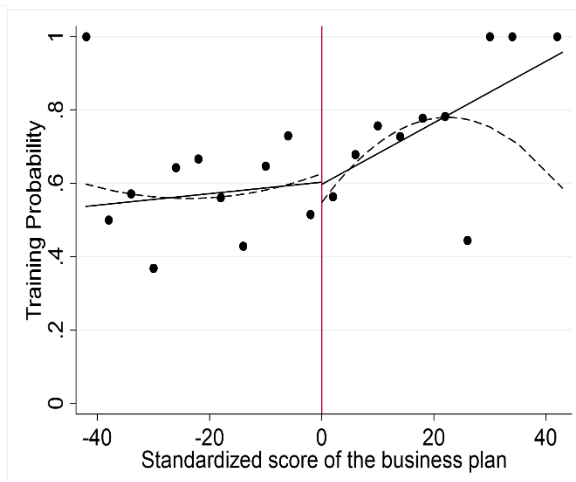
Panel B: EFS for Entrepreneurship Competency training



Panel C: EFS for Marketing training



Panel D: EFS for other trainings of the program



Notes: Dependent variable for each panel is dummy for taking the type of training specified in each panel. The types of training themes (modules) presented from panel A to panel D are those offered by the business plan competition. Data for access of each training are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

3.4.1.3. Additional Robustness Checks on the effective First-Stage

i. Effective first-stage by Synchronizing the timeline

In all the effective first-stage results presented so far, the treatment indicator (training dummy) was constructed by considering the training experience of the respondents they have *"ever"* had. This time-unbounded experience included the training they had before the business plan competitions were launched. One potential concern here is the substitute treatment for the rejected applicants may have been done before the program. If so, what we perceived as a substitute treatment during the program period may have been driven by the baseline balance, which is desirable, between the rejected and offered applicants regarding their prior training exposure.

As a response to this concern, the respondents were asked if they have attended any entrepreneurship training over the last one year, which covers from the start of the business plan competition to the follow-up survey. From this information, I constructed another treatment indicator that takes the value 1 if an entrepreneur attended any entrepreneurship training since her/his participation in Bruh/EDC and 0 otherwise.

The effective first-stage results estimated using this indicator are reported in Appendix 3.1, from Figure 3.A.1 to Figure 3.A.3. Still, the results remain the same as what are presented before with a first-stage parameter of 11% to 15% which is insignificant parameter for the full-sample, about 22% to 27%, depending on the specification, and significant coefficient for Bruh and almost zero for EDC. Therefore, the conclusion about the effective first-stage finding remains the same even after synchronizing the timeline to be consistent with the program period. Further, this result also shows that a large fraction of entrepreneurs got trained in a single year which again ensures the proliferation of training opportunities. As shown in Table 2.A.2 in the appendix most of the training program are offered by different types of government organizations followed by NGOs and international organizations.

ii. Effective First-stage by types of training providers or programs

Though applicants above and below the cutoff do not significantly vary by the contents as well as timing of the training each group had in any program (particularly in the full sample and EDC sub-sample), it would be also important to account for the variations in training providers. Because some organization or programs could be more effective than others due to either the variation the variation in the delivery mechanism or any other measures of quality. In this regard I did two additional analyses: by organizations of interest (EDC or JCC) or type of program.

First, I tested for the difference of rejected and offered applicants' access to trainings provided by EDC or JCC at anytime as these organizations have been providing entrepreneurship training in the programs in addition to the business plan competitions. No matter what the type of program is training provided by a given organization, it is likely to be of the same content at any time, particularly in the organization which provide standardized training like EDC. If the two groups of applicants are found to have the same level of access to trainings by these organization, one can simply rule out the variations I found in the nominal first-stage and further consolidate the claim for effective first-stage.

To this effect, I restricted the training experiences of respondents to only JCC and EDC at any time in constructing the treatment indicator and generated the first-stage results which are reported in Appendix 3.1, Figure 3.A.4 to Figure 3.A.6. This result varies by the type of competition. For Bruh applicants, marginally qualifying for the training in this business plan competition helped its participants to access training by these organizations by 36.6 percentage points more than their rejected counterparts and this estimate is statistically significant at 1% (linear specification of Figure 3.A.5). For EDC sub-sample, on the other hand, this estimate is not more than 10 percentage points and statistically not different from zero (Figure 3.A.6). This could be because Bruh applicants are younger, more with new business ideas, or with a younger startup as compared that of EDC and thus the former group had generally lower exposure for prior training opportunities than the letter one. On average, for the full sample (Figure 3.A.4), the program allowed offered participants to enjoy JCC or EDC's training by about 25 percentage points higher than reject applicants. In sum, in these business plan competitions, these organization, specially JCC, managed to benefit considerable number of entrepreneurs that would not have been trained otherwise.

Second, when the variation of training intensity by the nature of the program is considered, we witness from Bruh bootcamp as well as other incubators and accelerators that the training provided by incupetion programs is different from other ordinary business or entrepreneurship trainings. As a result, I run first-stage regression for access to training programs exclusively offered by incupetion related programs which includes training by any business plan competitions, Bruh or EDC programs, business incubation, and acceleration programs.

The results presented in the appendix Figure 3.A.7 to Figure 3.A.9 show that marginally scoring above the cutoff significantly increased the probability of attending trainings organized by incupetion programs, particularly for the full sample (by about 22 percentage point in the linear specification) and Bruh (40 percentage points) whereby the former is driven by the result of the latter. In Figure 3.A.10 to Figure 3.A.12, I estimated participation in incubations programs other than that of JCC and EDC and I did not find a significant first-stage. This implies that overall experience of incupetion program for the applicants is largely driven by their experience in the program of interest (Bruh and EDC). Likewise, in case of EDC, the estimated parameter in Figure 3.A.9 is negligible whereas in Figures 3.A.12 rejected applicants have more (about 19 percentage point) access to other incupetion programs. This shows that this business plan competition helped offered applicants of EDC competition to close the gap that would have been created otherwise by opening-up the opportunity for training within the program of interest.

Then, whether this relatively better access to trainings of incupetion programs for applicants above the cutoff will create a difference in their business outcomes will be tested in the reduced-form analysis in the next section. Before embarking on that, however, let us conclude this section by summarizing the results of the first-stage equation.

3.4.2. Summarizing the findings of the first-stage

The key takeaways of the first-stage findings are summarized as follows.

- There is a strong first-stage when program-specific treatment (training prepared as an integral part of Bruh and EDC business plan competitions) is used, which is true for the full sample, Bruh, and EDC sub-samples. I called this first-stage nominal first stage.
- I ran another first-stage where treatment (training) by other substitute programs is considered in addition to training by the program of interest. I called this as effective first-stage. In this case, the first-stage estimate becomes negligible for the full sample and EDC sub-sample, but it remains significant for Bruh even if the magnitude of the coefficient largely diminished as compared to the nominal first-stage estimate.
- The training that rejected applicants of the program enjoyed elsewhere is comparable with the training that JCC and EDC offered for their accepted applicants both in terms of contents covered and timeline, as confirmed in the robustness checks. This implies that

substitute treatments were indeed too close substitutes though the provider and the format of the program are different.

- Marginally qualifying the first screening was more important for Bruh applicants than their EDC counterparts to get training opportunity. It seems that the latter group gets training anyway.
- There have been many entrepreneurship training programs run by various government organization and NGOs, as I learned from the interaction that the applicants of our business plan competitions reported to have. The shares of JCC and EDC were found to be considerable in provision of entrepreneurship training in their various programs.

3.4.3.Estimation of the Reduced form equation

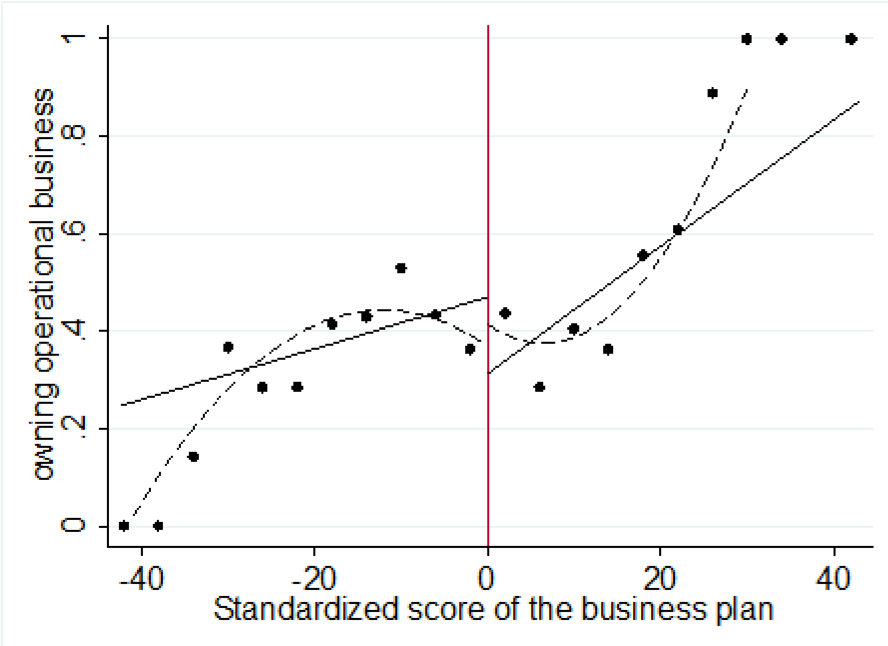
In this section, results and discussions of the reduced form equation specified in equation (3.3) are presented. I will start with results associated with the outcomes that measure business entry followed by outcomes which intends to measure business expansion. In order to ease the understanding of the findings and for more transparent communication of results, I presented the findings mainly using binned scattered graphs like that of the first-stage. To keep a good balance between readability of the document and provision of detailed information, attempt is made to focus on results of the full sample in the main body while additional results for Bruh and EDC sub-samples are also provided in the appendix.

3.4.3.1. Reduced form results for business entry and survival

This program mainly intends to support startups to facilitate to the establishment and survival of businesses that could be a source of employment for the owners as well as other people. Cognizant to this objective of the program, launching a business and have subsequently survived is the main outcome of interest for this study. To this effect, I defined the outcome variable, owing or operating a business, which takes the value 1 if the entrepreneur owns an operational business enterprise either in group or alone a year after the application to the business plan competition, and 0 otherwise.

The reduced form equation estimated for this outcome using the full sample is depicted in Figure 3.6. As shown in this graph, the probability of operating a business at the cutoff is statistically the same for applicants below and above the cutoff. This implies that marginally qualifying for the training program of the business plan competition did not make any difference on business ownership or self-employment. This is consistent with my expectation particularly after I learned from the effective first-stage estimates that the two groups are virtually the same in terms of their training probability (the treatment).

Figure 3. 5 Reduced form results on probability of owning a business (self-employment) for the full sample



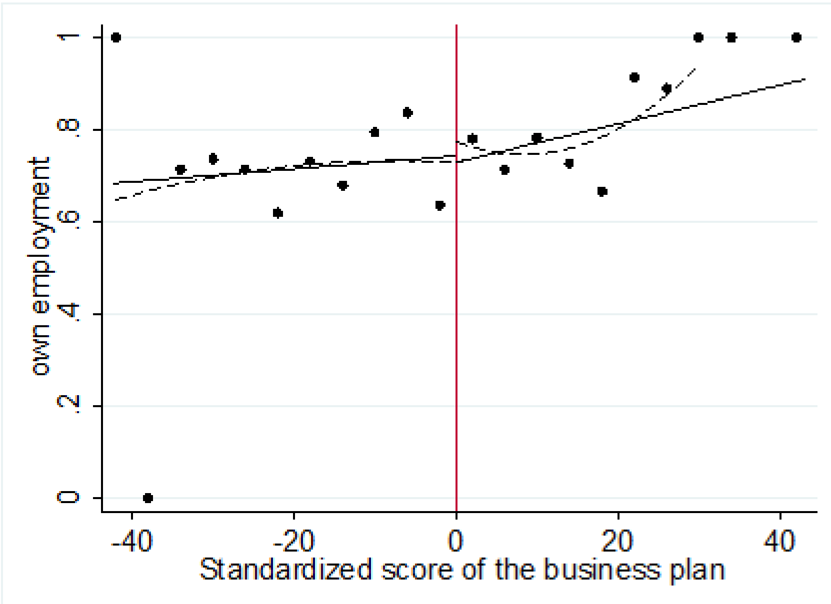
Notes: Dependent variable is dummy for owing (or operating) a firm one year after the application to the business plan competitions. Data about firm ownership is self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

It is also to consider the wage employment outcome of applicants while assessing their self-employability as one affects the other. I defined an outcome called 'own employment' which takes a value 1 if the respondent is either self-employed or wage-employed, or both at the time of the follow-up survey; and 0 otherwise.[11] The reduced form estimates for the full sample depicted in Figure 3.7 shows no jump in probability of own employment at the cutoff which means both rejected and offered applicants are not different from each other in this respect. The result in Figure

[11] The definition of this outcome was taken from McKenzie (2017).

3.7 further informs that about three-fourth (or 75%) of the applicants are employed and more than half of this is attributed to self-employment (Figure 3.6).

Figure 3. 6 Reduced form results on probability of own employment for the full sample



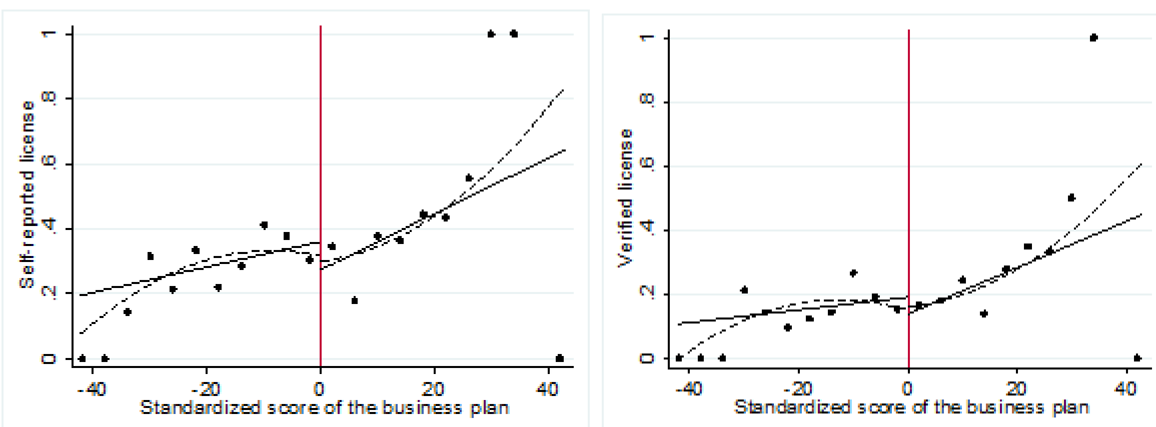
Notes: Dependent variable is dummy for own employment which takes the value 1 if the respondent is either self-employed or wage employed or both one year after the application to the business plan competitions. Data about employment status is self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Formalizing the operation of the small businesses through facilitating licensing of their business is one of objectives of government organizations working on entrepreneurship development like JCC, for instance. In this study, I have also evaluated if the program has any effect of owning licensed businesses. The reduced form results reported in Figure 3.8 are based on two different measures license outcomes. In panel A, the probability of operating licensed business based on self-reported license status of businesses is displayed while panel B presents for license which we independently verified its existence in the local and federal government agencies responsible for licensing business, as described in chapter 2. In both panels, the estimates are unconditional estimates which are not conditional on operating a business. To this end, the outcomes of respondents who did not own businesses one year after the start of the program are coded to zero. Doing so is quite important to address the issue of sample selection.

The result in both panels of Figure 3.8 revealed that applicants who scored above the cutoff are not better off in terms of operating formally registered businesses as compared to their counterparts rejected in the first round of the competitions.

Figure 3. 7 Reduced form results on owning licensed (or formal) business

Panel A: Self-reported license   Panel B: Verified license



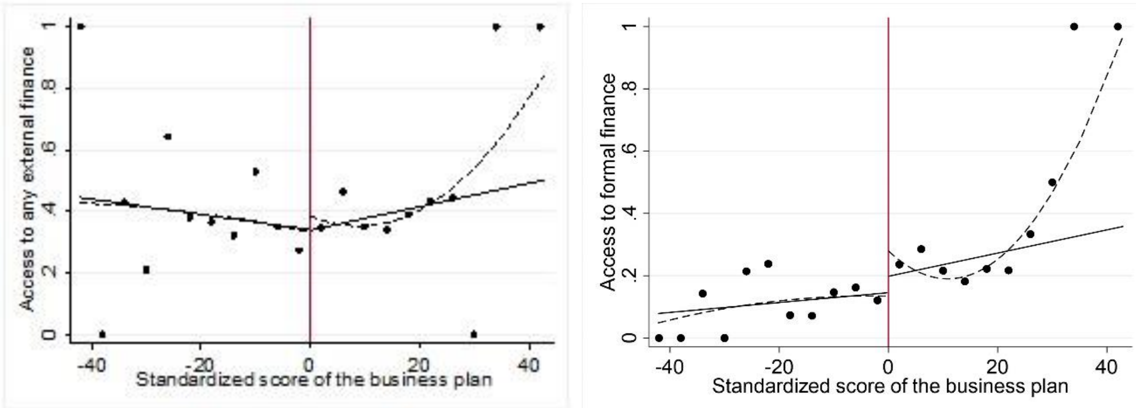
Notes: Dependent variable for each panel is dummy for owing (or operating) licensed business one year after the application to the business plan competitions. The results are not conditional on operating a business and outcomes of those who did not operate business were coded to zero. Data about license status of the firm in panel A is self-reported in the follow-up survey while in panel B it is independently verified from local regulatory agencies. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Finance is one of the biggest challenge for startups and small businesses in Ethiopian entrepreneurship landscape (Gebreeyesus et al., 2018; World Bank, 2015). Cognizant to this challenge, the business plan competitions under study were designed to relax this constraint by providing direct grants for final winners and facilitate external financing at least for accepted participants. Trainings on business plan preparation, pitching skills, networking with investors, and related activities of the competitions were primarily to improve the fundability of the business ideas. When these interventions are effective, we can see a higher probability of attracting external finance which includes bank loan, micro finance Institutions loan, equity investment, angel capital, grants from formal organizations, and other formal sources. When entrepreneurs' access to external finance is improved, their reliance on informal financial sources such as borrowing from families, friends, local money lender, and local traditional associations (like *Ikub* and *Idir* in Ethiopian case) is expected to diminish.

In figure 3.9, the reduced-form results for access to any external finance (Panel A) which included both formal and informal sources and for the formal financing (panel B) are reported. In both cases, the estimated coefficients are not statistically different from zero. This implies that even successful applicants in these business plan competitions are at low penetration rate of the formal financial sources. This group is no different from the rejected ones at least within one year since their participation in the program. This is also consistent with the direct responses of applicants in my follow-up survey where lack of finance is still the first business obstacles, as it was reported by a third of the respondents as their number one constraint of business operation.

Figure 3. 8 Reduced form results on access to finance

Panel A: Access to any external finance   Panel B: Access to finance from formal sources



Notes: Dependent variable is dummy for getting loan or grant for the business from both formal and informal sources (for panel A) and from formal sources only (from panel B) since application to the business plan competitions. Data about access to finance of the entrepreneur is self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

In a nutshell, the business entry and survival, formality, and access to finance outcomes of offered applicants are not statistically different from rejected applicants of the business plan competition a year after the application. These findings remain the same when we disaggregate the analysis into Bruh and EDC, as shown in the appendix 3.2 Figure 3.A.13 (for Bruh) and appendix 3.3 Figure 3.A.16 (for EDC). The reduced form equation estimates of all outcomes presented graphically and discussed so far disaggregated by the competition type are also summarized in Table 3.5.

Table 3. 5 Reduced form Estimates of the effect of scoring above the cutoff on business entry for the full sample

Outcome	Full sample		Bruh		EDC	
	Linear	Quadratic	Linear	Quadratic	Linear	Quadratic
	(1)	(2)	(3)	(4)	(5)	(6)
Operating firm	-0.1557*	0.0450	-0.1707	0.0703	-0.1823	0.0298
	(0.0801)	(0.107)	(0.1092)	(0.143)	(0.1215)	(0.172)
Own employment	-0.0134	0.0448	-0.0594	-0.0227	-0.0145	0.0912
	(0.0707)	(0.0982)	(0.1071)	(0.148)	(0.0947)	(0.139)
formal finance	0.0526	0.1460*	0.1186	0.1460	-0.0150	0.1120
	(0.0644)	(0.0811)	(0.0933)	(0.121)	(0.0927)	(0.125)
External finance	-0.0011	0.0510	-0.0238	-0.0378	0.0094	0.1570
	(0.0794)	(0.1040)	(0.1105)	(0.144)	(0.1184)	(0.165)
Licensed (self-reported)	-0.0871	-0.0162	-0.1341	-0.0220	-0.1102	-0.0474
	(0.0775)	(0.1040)	(0.0969)	(0.127)	(0.1217)	(0.174)
Licensed (Verified)	-0.0519	0.0108	-0.0925	0.0157	-0.0661	-0.0234
	(0.0627)	(0.0810	(0.0653)	(0.0793)	(0.1074)	(0.1540)
Observations	456	456	214	214	242	242

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Dependent variables are as defined before. The reported coefficients are estimates of the reduced-form parameter; the standardized scores and its interaction with the indicator of being above the cutoff have been controlled for; Estimations are based on the full support. The results are not conditional on operating a business and outcomes of those who did not operate business were coded to zero. Linear and Quadratic are linear are types of functional forms of the model.

3.4.3.2. Reduced form results for performance indicators

Entry to the market is not the ultimate objective of businesses as well as for entrepreneurship support policies and programs. It is from the growth and expansion of the new entrants or incumbent firms that an economy would benefit either in terms of job creation or productivity gains or any other metrics. That is why this program has also included existing young business as one of the target groups with the aim of helping them expand their operation. In line with this intention of the program, levels of sales, profit, and employment are selected as additional outcomes worth considering evaluating the effectiveness of the program.

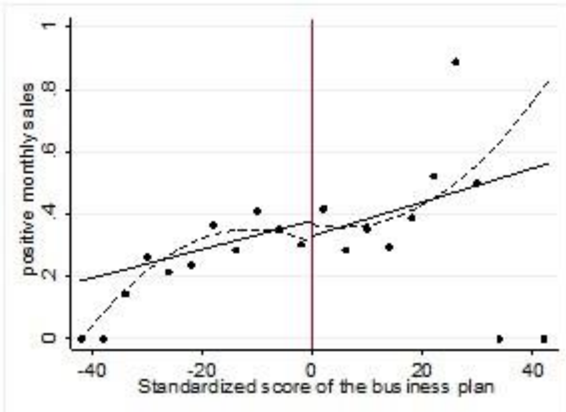
Considering more multiple outcomes in evaluation of small business interventions is much common in the literature as well. I argue that having multiple outcomes in this study is justified from different grounds. First, the credibility of self-reported business data is usually questionable and thus supporting findings of one measure of performance by other alternative measures enhances the robustness of the result.

Second, usually a program is designed and implemented to achieve multiple objectives of various stakeholders. Sometimes, in a social or economic setting, there may be contradicting objectives in which the failure in one objective of the program could be partly due to the success in the other. For instance, supporting the informal business to get registered and licensed by the local regulatory agencies could help apply for formal loans as having business license is the first criteria of banks or microfinance institutions in Ethiopia. However, doing so also automatically makes businesses subject to different taxes by local authorities which could negatively affect their survival. This condition demands the researcher to consider more outcome variables for the research to be informative about the policy makers. Therefore, the choices of the outcome variables should be synchronized with the program theory.[12]

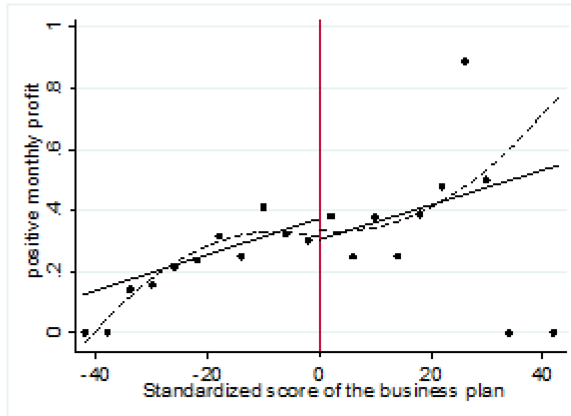
[12] These justifications are just relevant for this chapter. In the next chapter (section 4.3.2), I also provided additional justifications which are relevant for the research questions of that chapter.

Figure 3. 9 Reduced form results on dichotomous measures of firm performance (Full sample)

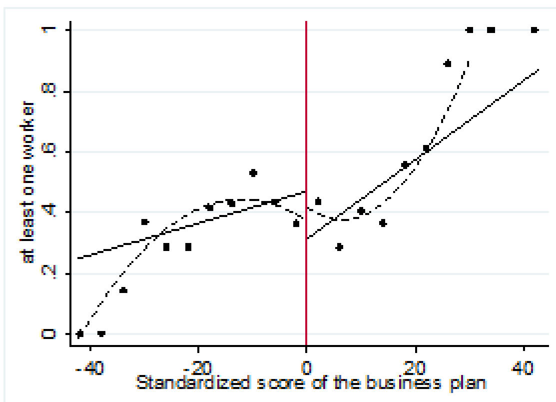
Panel A: Reporting Monthly sales



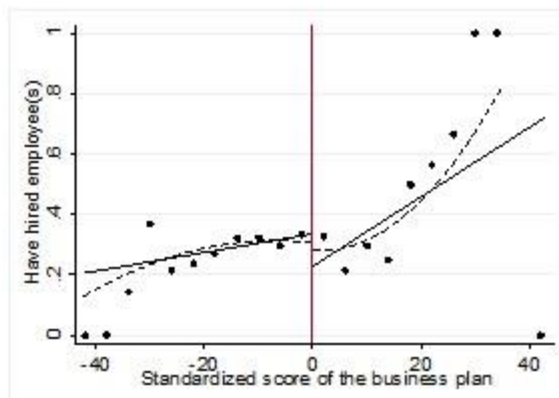
Panel B: Reporting Monthly profit



Panel C: Having at least one worker



Panel D: Having hired employee



Notes: Dependent variable for each panel is dummy for reporting any sales, profit, worker, and salaried worker (hired employee) for panel A, B, C, and D, respectively, one year after the application to the business plan competitions. The results are not conditional on operating a business and outcomes of those who did not operate business were coded to zero. Data about these outcomes are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Coming back to the reduced form results pertaining to the three performance indicators (sales, profit, and employment), I started by taking a simple indicator of each for reporting any sale, profit, or employment. For instance, for sales, the outcome variable takes 1 if the business owner reports any positive monthly sales a year after the start of the program or 8 months after the training of the program was completed. Redefining continuous business outcomes could be important to minimize measurement errors caused by misreporting of the actual figure.

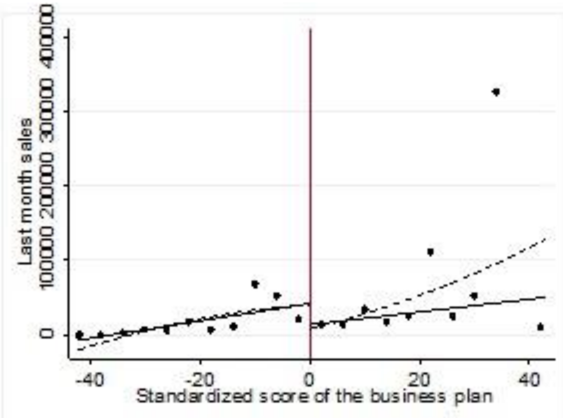
The reduced form results reported in Figure 3.10 for probability of reporting any sales (panel A), profit (Panel B), any worker (panel C), and any hired or salaried workers (panel D) show that scoring above the cutoff and qualified for the program's training did not improve any of these business outcomes. While Figure 3.10 is for the full sample, similar results are also reported in appendix 3.2 and appendix 3.3 separately for Bruh and EDC, respectively.

Last, but not least, I have also examined my reduced form equation on four continuous measures of business performance indicators as the dichotomous measures I presented above collapse variations within each group. Since exploiting the variations in these performance measures could give us a good insight, considering the continuous measures is also important. Here, sales and profit are last month sales (profit) in Ethiopian Birr measured a year after the application of the program. Employment is measured by the total numbers of workers and amount of external finance raised is the sum of all loans, grants, equity investment, and any other finance raised from formal and informal sources in Ethiopian Birr.

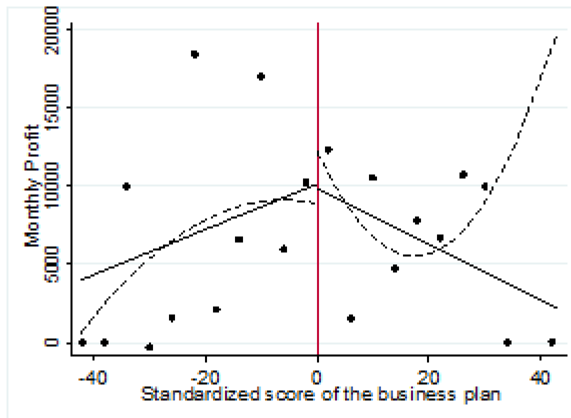
As it can be viewed in all panels of Figure 3.11 (for the full sample) and appendix 3.2 and 3.3 (for Bruh and EDC), the reduced form parameter is significant in none of the four outcomes in both linear and quadratic specifications. Even the coefficient on employment in case of Bruh is significantly negative (appendix 3.2 Figure 3.A.15 panel C). For the rest, applicants to the left and right of the cutoff are statistically the same in terms of various measures of business outcomes. In short, for both measures of business entry and expansion, there is no any evidence about the improvement of the outcomes at the cutoff showing that offered applicants are indistinguishable from the rejected ones.

Figure 3. 10 Reduced form results on continuous measures of firm performance (Full sample)

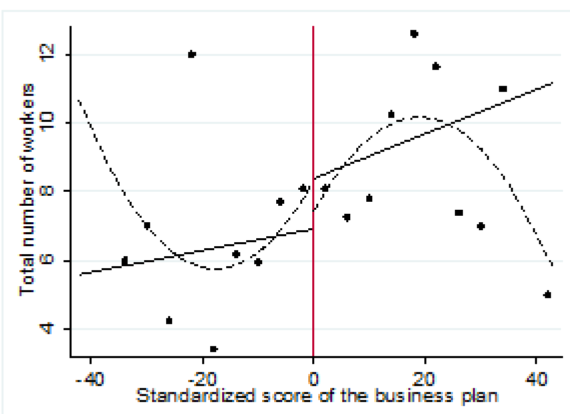
Panel A: Monthly sales



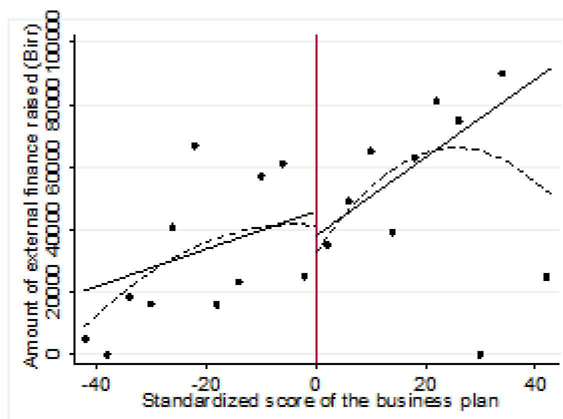
Panel B: Monthly profit



Panel C: Employment level



Panel D: Amount of external finance raised



Notes: Dependent variable for each panel the level of last month's sales (panel A) and profit (panel B) in Ethiopian Birr, total numbers of worker (panel C), and amount of external finance raised over a year in Ethiopian Birr (panel D), all as reported by the respondents a year after the application to program. The results are not conditional on operating a business and outcomes of those who did not operate business were coded to zero. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Final remarks on the findings

The finding of the reduced form parameter is consistent with our prior expectation based on the first-stage results. In the case of the EFS, the first-stage parameter (π) is near to zero. Therefore, the reduced-form parameter (πβ) will also be almost zero for any value of β. We have contaminated controls due to participation in substitute programs means the score would be a weak instrument and the estimation of the program effect using this framework (RD) would underestimate the true program effect. That is why I did not need to report the IV estimate, or the ITT parameter estimates with various specifications.

One potential solution for this seems estimating the causal parameter using the nominal first-stage using only non-contaminated part of the controls. However, I argue that this approach is flawed since the resulting parameter would also be biased dues to self-selection to the substitute programs. That means since treatment in the substitute programs were not assigned randomly a mere exclusion of the contaminated controls adds fuel to the flame. Similarly, comparing participants who had entrepreneurship training in any program against those who did not have any is the other option one could think of. However, leaving the weak instrument issue aside, this approach can only inform about the effect of entrepreneurship training in general, not about the causal effect of the program (Bruh ans EDC) which I intend to evaluate.

As a result, what we know in this study is not whether the program works rather the failure of the offered applicants to stood out of the crowed since they did not have a better access to treatment than their comparison group in the first place. Stopping the story here is by far better than reaching wrong conclusion about the program's effectiveness without ensuring the validity of the counterfactual, which is common in previous studies.

3.5. Discussion and Concluding remarks

This chapter aimed at disentangling the causal effect of the training (or bootcamp) program since it was the main intervention of Bruh and EDC business plan competitions. Though it was a perfect set up for the fuzzy RD design with a large effect of the first-stage estimate using treatment within the program, the follow-up survey uncovered that the rejected applicants of the competition (who were supposed to be the control group) got treated elsewhere. That means the significant first-stage I found in the nominal case disappeared when substitute treatments were taken into account (effective first-stage became negligible).

Since both groups of applicants had the same level of training and thus the first-stage is insignificant, we could not see any jump in the reduced form graphs when business performance measures are plotted against score. As implied by the reduced form parameter, I can conclude that training beneficiaries of the business plan competitions did not perform better than their rejected counterparts in both business entry and expansion. This does not mean that the program was not effective to impact business outcomes; nor the entrepreneurial skill constraint or training does not matter. Despite my initial intention, I cannot tell about the effectiveness of the program. The evidence provided in this study only tells that the program's effect, as implied by the reduced form estimates, is negligible because control groups got substitute treatment in similar programs.

Previous studies were not able to figure out as to why the training component of business plan competitions yielded negligible impact on entrepreneurship activities. For instance, McKenzie (2017) reported large estimates the first-stage (74 to 90 percentage points). However, his reduced form graphs did not show any jump at the cutoff and the estimated causal parameter became insignificant. Though he did not explore why the effect appeared to be negligible, it could be because there would not been strong first-stage in the first place had the effective first-stage been considered. Similarly, Fafchamps & Woodruff (2017) were not able to tell why their training were ineffective thought they had rightfully suspected the issue of substitute treatment as a possible reason.

This paper provided clear empirical evidence on substitute treatments of the control groups with a far-reaching implication for future research designs. This could be part of the explanations for the modest or small positive impacts of entrepreneurship training programs we witness around the globe (Mckenzie et al., 2020).

This study implies that it is important for impact researchers to go beyond the baseline balance check-up. Documenting the necessary information about the possible substitute treatments that subjects may have enjoyed during the program period or after that should be part of the follow-up data collection task. Then it is essential to estimate both the nominal and effective first-stage. If the estimates of the two types of first-stages converge, then we can be confident on the cleanness of the control group. Otherwise, the control group would be a contaminated control based on which one cannot say anything about the program effectiveness unless it is justified by any possible variations in treatment intensity or quality.

Many program evaluation studies try to provide explanations how the program works when it works through exploring various channels usually taken from the theoretical models. Nonetheless, it is not common, at least in this area, to see plausible and evidence-based explanations when the program fails. Figuring out why a program did not work would be much helpful for the policy makers. Most importantly, the result would be misleading if we conclude that the program is ineffective while our control group is contaminated by substitute treatment. Therefore, exploring the reasons behind any finding should be a prerequisite to accept any result. This is one of the key messages of this study.

CHAPTER 4

4. Predicting entrepreneurial success and identifying gazelles using business plan competitions

4.1. Introduction

Policy makers and governments in developing countries are primarily concerned with achieving robust economic growth and realizing the structural transformation in one hand and creating more jobs and lifting people out of poverty on the other. Startups, which are disproportionately small firms, plays a central role in the economic dynamics of countries. Startups are recognized as a source of productivity growth and net job creation since they are the main actors in the process of replacing old and inefficient incumbent firms with the new and efficient one. This process is usually referred as ‘creative destruction’ in the literature (Decker et al., 2014; Nanda, 2016).

However, these firms face several constraints including lack of entrepreneurial capital, limited access of credit, underdeveloped infrastructure, restrictive entrepreneurship landscape, and lack of self-confidence by the entrepreneurs (Bigsten & Söderbom, 2006; Blattman et al., 2014; Gebreeyesus et al., 2018; Kirkwood, 2009). These constrains held high-potential firms away from their optimal size and curb the contribution they would have played in an economy otherwise. This is one of the main justifications behind the policy interventions and numerous entrepreneurship development programs we witness around the world. Yet, it is not uncommon to see many entrepreneurship development programs to have limited impacts and entrepreneurship training is one example in this respect (McKenzie & Woodruff, 2014; Mckenzie et al., 2020).

One reason for the failure for policies and development programs to achieve the intended impact could be associated with firm heterogeneity. Firms differ in their intention, characteristics, constraints, potentials, and the like. Thus, generic policies or support programs which are not tailored to their needs are expected to have either heterogenous impacts or no impact at all. The main origin of firm heterogeneity is the type of entrepreneurship since small minority of the enterprises in developing countries are dynamic enterprises (or gazelles) which have high potential for growth while the rest majority are survivalist (or necessity driven) type of enterprises which apparently struggle to survive and with little or no ambition to grow (Mead & Liedholm, 1998; Cotter, 1996; Nichter & Goldmark, 2009; Berner et al., 2012). Therefore, the bifurcation of firms into gazelles and survivalist has been considered as an essential step towards successful entrepreneurship policy.

The proper bifurcation of firms enables policy makers to target the right group depending on their objective and design and implement more tailored policy approaches. For instance, a policy that aimed at tackling poverty and improving welfare could target survivalist types of enterprises with a distinct type of support while gazelles would be the right group for a policy that intends to boost industrial growth and wage employment. Basically, many development practitioners, NGOs, policy makers, investors, and other development actors intend to target gazelles and unlock their potential due to their unparallel roles in terms of job creation, productivity, and return to capital for any additional intervention they receive (Henrekson & Johansson, 2010; Shane, 2009; Sims & O'Regan, 2006; Nanda, 2016).

Then, how can we identify gazelles from the crowd of enterprises? Business plan competition is one of the policy options recently being popular in developing countries to separate high-growth potential businesses from the mass(McKenzie, 2017). Since differentiating promising startups or businesses ideas from the low-potential ones is the central tasks of any business plan competition, a panel of experts spend considerable time to rigorously screen applicants by scoring their business plans and/or organizing pitches. However, to what extent businesses plan competitions become successful in making accurate predictions of growth potentials of firms and select most promising applicants for the intended intervention remains an empirical question.

Thus, this chapter aims to examine if the business plan competition judges could predict entrepreneurial success and differentiate (potential) gazelles among startups participated in two national business plan competitions (Bruh and EDC) in Ethiopia. As described in chapter 2, the panel of judges scored the business idea of all eligible applicants right after the application period, that is, a first-round screening based on which placement to the training was determined. The average score reflects the growth potential of each firm (or applicant) as predicted by judges *ex ante*. A year after this prediction was made, I traced almost all the applicants (about 500 in number) and measured their actual business outcomes in my own follow-up survey. Using data on the predicted growth potential (i.e. score from the judges) and actual performance measured by firm entry and survival, employment, sales, profit, and aggregate growth, the study provides a fresh empirical evidence whether the business plan competition helps predict entrepreneurial success and identify potential gazelles.

The results of the study reveal that score from the business plan judges is a significant predictor of entrepreneurial success consistently in all measures of business outcomes. There is non-linearities in the relationship that score is more predictive in the bottom and top of the distribution, but not in the middle, implying that judges were more successful in identifying most promising (or high potential) projects and least promising (or non-serious) ones. In general, the business plan competitions under evaluation succeeded in identifying gazelles. However, only small variations in outcomes are explained by score and other baseline covariates, consistent with the previous findings. Further disaggregating the analysis by types of competition show that experts were able to predict success in the EDC competition but not in Bruh and this situation provides a good opportunity to explore the possible reasons behind the contradictory findings of the previous literature in this regard. We also noted evaluation of existing business as opposed to new ideas, using more detailed criteria, reducing the burden of each expert by allowing more time and assigning reasonable applicants to evaluate, and using experts familiar with local context and business environment of the target group help improve the accuracy of prediction.

This paper is related to the theoretical literature about the experts prediction ability and drivers of its accuracy (Shanteau, 1992; Kahneman & Klein, 2009). One of the interesting areas of empirical research with a far-reaching policy implication that such theoretical predictions can provide a good guide is the evaluation or scoring of business potentials made by panel of experts in cases of business plan competition, screening process of acceleration and incubation programs, and funding decisions procedures of venture capitalists or angel investors, among others. This paper is directly related to the empirical works that assess the ability of judges/experts in predicting future business outcomes. There are two contradicting findings in the literature. The first group confirms the idea that experts initial judgment about growth potential of businesses is a significant predictive of actual success (Fafchamps & Woodruff, 2017; Scott et al., 2020; Åstebro & Elhedhli, 2006) while others demonstrated that entrepreneurial success is hard to predict (McKenzie & Sansone, 2019; Kerr, et al., 2014b). The dearth of empirical evidence in the area makes the literature far from being adequate and conclusive.

This study makes at least two contributions in this literature. First, it builds up from the previous literature and provides new empirical evidence on the predictive ability of experts on entrepreneurial success of startups over a wide range of outcomes in context of a business plan competition in low-income countries. From this we can gauge if a business plan competition is a successful policy option in developing countries in differentiating gazelles or startups with the potential to be gazelles from survivalist types which is a critical step for effective interventions. Second, from the two natural experiments considered in this study, it is only in the case of EDC that the prediction was successful but not in Bruh even though both competitions were conduct in the same setting within the same country and period. This situation opened-up the opportunity to explore the possible conditions, stated earlier, under which prediction of entrepreneurial success could be improved. In this regard, some of the explanations I provided for the variation in results between the two cases could be used to explain the mixed results we witness in the previous studies for which adequate explanations are hardly available so far.

The remaining part of the of the chapter is organized as follows. The second part discusses the related literature in the area. The empirical strategy is outlined in the third section. Section 4 presents the main results from the econometric models and related analyses. The fifth part concludes.

4.2. Review of related Literature

4.2.1. The rationale for bifurcation of firms and targeting

The need for bifurcation of firms into high-growth and survivalist type has been acknowledge by researchers and policy makers at least since the last three decades (House, 1984; Cotter, 1996; Mead & Liedholm, 1998; Berner et al., 2012)[13]. Accordingly, survivalist or necessity driven enterprises are most of the small business population in developing countries which are created out of necessities mainly due to lack of wage employment opportunities in the labor market. Such enterprises tend to provide subsistence level of income and employ mostly the owners, family

[13] The concept is of course similar with the dual urban economy or recognition of the informal sector by ILO early 1970s.

members or few external workers. They are known for high patterns of exit at their early age and stagnant growth when they manage to survive(Grimm et al., 2012; Mead & Liedholm, 1998) . They operate in the environment that is characterized by overcrowded market niches (Berner et al., 2012)

The second group of firms, on the other hand, are fast-growing firms which are also know by various names such as dynamic, growth-oriented, and gazelles. These firms have a capacity and aspiration to growth. Even though they constitute small proportion of the firm population in many countries, gazelles play disproportionate role in an economy. They exhibit high level of productivity, good financial performances, high return to capital, produce high quality products, creates more jobs, and continue to grow (Berner et al., 2012; Sims & O'Regan, 2006; Abebe et al., 2017; Mottaleb & Sonobe, 2012). Particularly, their unparallel contribution in job creation is well pronounced. For instance, Decker et al. (2014) estimated that near to 50% of the jobs created in the US attributes to the fast-growth firms.[14]

As a result, gazelles have been a priority target for policy makers, NGOs, and other development actors. For instance, governments and policy makers who envisage to achieve growth of industries and structural transformation have an incentive to target high-growth potential firms since they are likely to engage in accumulation, specialization, rip opportunities, and keep improving their productivity. Similarly, NGOs and other development program implementers seek to work with gazelles as they are likely to yield a good return to interventions they receive which, in turn, ensures program effectiveness. It has also been argued that a policy targeting gazelles is in line

[14] Grimm et al. (2012) extended the bifurcation into three by creating another group of enterprises which they call it "constrained gazelles" that lays between the two common types.

with a policy supporting constrained enterprises since there is a remarkable overlap between facing constraints and potentials for growth (Fafchamps & Woodruff, 2017).

Targeting gazelles by some interested group to achieve certain objectives does not mean that the survivalist enterprises do not deserve support. The literature acknowledges the need for support policies for both types of firms (Mead & Liedholm, 1998; Berner et al., 2012). However, these distinct group of enterprises need different policy approaches. That is the very reason behind making such bifurcation.

For instance, Mead & Liedholm (1998) argues that while a more tailored business development policy, non-financial supports in the area of product and inputs marketing, and substantial credit which goes beyond working capital for the firms to be able to purchase fixed capital are the types of policies required for gazelles, a provision of working capital through developing credit and saving schemes could be helpful for survivalist type of enterprises to ensure their survival and earn somewhat better and more stable income. Thus, they argue that survivalist enterprises must be a primary target for poverty alleviation as they help "a large number of very poor people become a little less poor." Nonetheless, this oversimplified policy prescription particularly for the latter types of firms has been recently challenged by Berner et al. (2012).

4.2.2. Methods to identify gazelles

A numbers of statistical methods have been developed to identify gazelles the mass of existing enterprises (for instance, Sims & O'Regan, 2006; Grimm et al., 2012). These methods commonly rely on the computation of firms' growth footprints using a multiperiod actual performance data. While such methods may be useful to separate fast-growing firms according to their current performance, the potential of young startups like ours with no or little track records cannot be determined using such criteria. In addition, focusing on the potential rather than the current status is important. Nanda (2016) underscores this point as "*policymakers would do better to focus on policies that can identify high-growth-potential firms rather than on trying to pick high-growth firms in advance and then trying to keep them alive when they perform poorly*" (Nanda, 2016, pp 9).

When we zoom in the exercise to identify gazelles, experts' judgement, simple *ad hoc* linear model mainly utilizes data on the traits of entrepreneurs and characteristics of their enterprises, as well as the machine learning techniques could be the potential methods relevant for startups. For example, a modern machine learning technique can be a potential mechanism to identify gazelles using big data. Nonetheless, it is less likely to be effective in context of entrepreneurship competition where one cannot have big enough data like millions of observations to benefit from this potential method (McKenzie & Sansone, 2019). Therefore, experts' judgment which could be further augmented by econometric models can play a central role in predicting high-growth potential entrepreneurship. In the next sub-section, we discuss the available evidence in the literature regarding the role of experts in identifying gazelles or predicting entrepreneurial success.

4.2.3. Evidence on using experts' judgment to predict entrepreneurial success

The judgment made by the business plan competition judges and their ability to predict business performance can be associated with the theoretical foundation of a literature in organizational psychology, decision making, and cognitive science. Researchers pointed out the factors that led experts to perform well or poor in making the judgements. The theory of expert competence provides a good guidance in this regard (Shanteau, 1992). Accordingly, five factors, namely, domain knowledge, psychological traits, cognitive skills, decision strategies, and task characteristics affect the skills and abilities that emerge (or do not emerge) in experts. Shanteau (1992) underscores the importance of the last factor (nature of the task) and stipulates the characteristics of a task associated with good and poor performances in experts' judgment.

Similarly, Kahneman & Klein (2009) provide three conditions for the experts forecast to be more accurate: First, the outcomes being judged are reasonably predictable; second, the experts have extensive experience making those judgments; and finally, the experts receive rapid and continuous feedback on the accuracy of their initial judgments.

There is a dearth of literature that empirically tests the ability of experts in predicting entrepreneurial success by differentiating high-growth potential entrepreneurs from low-potential ones. The existing evidence provide two contradicting views in this regard. The first group of empirical literature supports success in experts prediction (Scott et al., 2020; Scott et al., 2015; Fafchamps & Woodruff, 2017; Åstebro & Elhedhli, 2006) while others underline the difficulty of succeeding in making such prediction (e.g. McKenzie & Sansone, 2019 and Kerr, et al., 2014b)

For instance, Scott et al. (2020) examined the evaluation of 537 early-stage ventures is the Massachusetts Institute of Technology (MIT) venture mentoring service (provided for MIT affiliated entrepreneurs) made by 251 experienced entrepreneurs, investors, and executives between 2005 and 2012. They found a positive and significant association between the mentors' initial interest and the actual commercialization of the proposed business. They also noted that the prediction exercise was successful for ventures in some sectors such as the hardware, energy, life sciences, and medical devices sectors but not in others like consumer products, consumer web and mobile, and enterprise software sectors.

In a business plan competition in Ghana with140 participating firms, Fafchamps & Woodruff (2017) reported business plan scores from the panel of judges significantly predicted firm growth. Nonetheless, the target group of this competition was existing firms with average age of 9 years in business and whether this result can be replicated for startups like my case is unknown. In another study in Canada, Åstebro & Elhedhli (2006) traced 561 projects among the applications evaluated by the Inventor's Assistance Program(IAP) from 1989 to 1994 to compare their business outcomes against the prediction made by experts. They reported that the IAP score is a significant predictor of probability of commercialization and internal rate of return (IRR).

As opposed to these evidences, McKenzie & Sansone (2019) compared judges' score, econometric model and machine learning techniques using data from more than 2000 business plan competition participants in Nigeria and underline the difficulty in prediction of future outcomes in any of the methods they considered. Again, in the *ex ante* assessment of venture capital investors in the US were not able to predict the ultimate success of the investment, conditional on financing, with a correlation of less than 0.1(Kerr et al. 2014b). However, since this finding is based on the applicants who were financed by the angel investors groups, it does not necessarily show the overall predictive ability of the investors over the entire pool of applicants.

In addition, in another paper published a bit earlier than the above one, Kerr et al.(2014a) apparently reported the significant association between the initial interest and future success though this paper is misquoted in some reviews for the opposite argument. They expressed this condition as:

> *"By itself, collective interest [the measure of growth potential] is very predictive of future outcomes; the coefficient on the angel funding dummy [the outcome] is 0.11 and significant at the 1% level. This positive association, moreover, holds when excluding companies that*

> *Tech Coast Angels ultimately funds. In unreported regressions, we find that the interest-level variable has a coefficient of 0.006 (0.002), indicative of the power of the screening mechanism" (Kerr, et al., 2014a, pp 44).*

They again reaffirm this finding in their concluding remarks by summarizing it as *"Overall, we find that the interest levels of angels at the stages of the initial presentation and due diligence are predictive of investment success" pp 52*. It may be because they reinterpreted this result in their next paper (that is, Kerr, et al., 2014b) as a weak association (probably due to the magnitude of the correlation rather than the statistical significance) or confusing between the two different articles that some studies misquote this study.[15]

There is also disagreement whether experts better distinguish high-quality ideas or low-quality or non-serious ones over the entire distribution of applicants. For instance, Scott et al. (2015) shows that experts better able to separate contenders at the top of the distribution while Fafchamps & Woodruff (2017) indicates judges are better at cleaving off the bottom of the distribution.

In sum, either in the business plan competition or other related interventions where entrepreneurs are screened based on a formal scoring procedure, the empirical evidence regarding the ability of experts or judges to differentiate high-growth potential entrepreneurs from the rest is not only scant but also inconclusive. The existing literature also fails to provide adequate explanations for such contradictory findings. Almost all studies in this area suffers from small sample problem since is hardly possible to come across large sample size in such natural experiments. Some studies also could not evaluate the prediction of experts beyond the probability of commercialization while it would be informative to assess the evaluation both at the intensive and extensive margins.

---

[15] For instance, McKenzie & Sansone (2019) and Scott et al. (2015) misquoted this paper.

This study aims to contribute to this scant literature by providing new empirical evidence whether business plan competition, via its panel of judges, can identify gazelles, at least potentially, from the mass of young entrepreneurs who applied for the program in Ethiopia. We extended from the previous related studies in several ways. For instance, this study is different from Fafchamps & Woodruff (2017) as we are dealing with startups while their study was for established firms with an average age of about 9 years. Even though startups were included in the sample of McKenzie & Sansone (2019), they failed to disaggregate the analysis from data of established businesses and thus the results does not necessarily inform about startups. Here, it should be also noted that near to 29% of my sample were existing businesses at the time of application. However, only early-stage businesses were eligible to apply for the competition and thus my 'existing' businesses are not established business; they are also different from what McKenzie & Sansone (2019) labeled as 'existing' where they did not have any age limit. This difference makes my entire sample to be startups while their sample is a mix of startups and established businesses.

Another closer study is Scott et al. (2020) with the fact that it exclusively deals with early-stage ventures but differs from this study not only it is a case in advanced economy (i.e. in the US) but also it was not in the business plan setting and only evaluated prediction only on commercialization. It does not shed light on the predictive capacity of experts about growth of ventures which is much more important particularly from the investors point of view.

Another worth mentioning point is that the correlation between score from judges and future business outcomes may not necessarily attribute to the judges' selection or predictive ability unless we control for variations in treatment. Because score could affect outcome through treatment which could be training or grant in case of business plan competition, for instance. McKenzie & Sansone (2019) acknowledge this and they disaggregate the sample into grant winners and non-winners and controlled for the variations of the grant even within the winners as the grant intervention of the program had a significant effect on business outcomes as reported in McKenzie (2017). Nonetheless, both McKenzie & Sansone (2019) and Fafchamps & Woodruff (2017) ignored the variation in training status in analyzing the score-outcomes nexus for the reason that the training attendance did not have a statistically significant effect on the outcomes of interest.

However, I argue that even if the coefficient of training is not statistically significant, it would be safe to control for it as lack of statistical significance does not mean that zero effect. In addition, even if it is insignificant by itself from the program point of view, it may not be negligible in affecting the relationship between other variables (score and outcome in this case). Therefore, regardless of its significance, controlling for the treatment effect makes the analysis of predictive ability (i.e. selection effect) neat.

In this study, hence, by disaggregating the sample based on the treatment status, I examined if business plan score from panel of judges is a predictive of success that is measured by a wide range of outcomes (measuring entry and survival as well as growth) for startups in the developing countries' context. In addition, I tried to provide preliminary explanation for the heterogenous results we observe in the existing literature.

4.3. The Empirical Strategy

4.3.1. The Model

Business plan competitions including this one intends to select high-growth potential applicants for various interventions which are usually designed as part of the program. In the business plan competition under evaluation, judges scored all eligible applicants of the competitions based on pre-determined criteria just right after the application was closed. This first-round screening result was used by the competition organizers to decide the best applicants who could advance to the next stage and entitled for the training intervention. The score from judges is a measure of growth potential for each applicant predicted at the time of application using the information provided in the business plan and without meeting the applicants in-person.

In order to gauge whether the business plan competition helped identify gazelles, we need to evaluate whether the jury was able to differentiate between promising businesses or business ideas with a high-growth potential from the low potential ones. This can be judged by analyzing the association between the actual business outcome and the potential of the businesses as measured by the average score of the business plan competition. To this effect, I measured actual business outcomes of all applicants in the follow-up survey a year after the prediction about their growth potential was made. If the business plan score is a predictor of future business outcomes, we can conclude that judges succeed to predict entrepreneurial success and thus the business plan competition is helpful to identify the nature of entrepreneurship (gazelles or survivalist).

Therefore, we can determine judges' predictive ability from the econometric model specified as:

$$Y_{ijsr} = \alpha + \beta Score_{ijsr} + \delta X_{ijsr} + \rho_j + \tau_s + \theta_r + u_{ijsr} \qquad (4.1)$$

Where the subscript i is for individual firm or applicant; subscript j is for panel of judges which varies by competition center which are 5 panel (the 4 centers of EDC and Bruh); subscript r and s are stands for region or location of the applicant and sector dummies of the business, respectively. Y_{ijsr} denotes the business outcome variable measured 1 year after the application. $Score_{ijsr}$ represents the standardized average score of the first-round screening from the judges. X_{ijsr} is a vector of exogenous controls mainly the entrepreneurial characteristics like gender and education measured at the baseline. ρ_j, τ_s, and θ_r judges, sector, and regional fixed effects, respectively. α, β, and δ are model parameters, and u_{ijsr} is the error term.

In this specification, β is the parameter of interest. The positive and statistically significant value of the parameter implies that score is predicts future entrepreneurial success.

4.3.2. Definition the outcome of variables

I want to test the predictive ability of judges over two groups of outcomes. The first one is whether the applicant was operating a firm at the time of follow-up visit. This outcome variable measures firm entry and subsequent survival. The second group of outcomes is about business expansion measured by employment, sales, profit, and aggregate growth. In this chapter, each of the outcomes are defined as follows.

- *Operating a firm:* this is a dummy variable that takes the value of 1 if the entrepreneur operates a firm at the time of the follow up survey and 0 otherwise.
- *Employment:* is the total numbers of workers measured during the follow-up survey or a year after the application to the business plan competition.

- *Sales:* Inverse Hyperbolic Sine (IHS) transformations of last month's sales in Ethiopian Birr measured a year after the application
- Profit: Inverse hyperbolic Sine transformations of last month's profit in Birr measured a year after the application.
- Aggregate growth index: This is an aggregate performance indicator aimed at summarizing the business outcome of an enterprises in a single variable while it captures various measures of entrepreneurial success. Following Fafchamps & Woodruff (2017), aggregate growth in this study is computed as the sum of standardized values of the IHS transformations of employment, sales, and profit following That means, first, I transformed each component (employment, sales, and profit) into IHS form. Then, I standardized each component was standardized before taking their summation to generate the aggregate index.

An outcome Y is transformed into its inverse hyperbolic sine value using the following formula.

$$Inverse\ Hyperbolic\ Sine\ of\ Y = log[Y + (Y^2 + 1)]^{\frac{1}{2}} \qquad (4.2)$$

This transformation helps me to reduce the skewness of the distribution and improve the fitness of the model, which analogous to the log transformation. Unlike logs, the IHS transformation enables us to transform not only positive figures but also negative and zero values.

I have four reasons for using five different outcomes in this study. First, as the program envisages to facilitate business entry and help business growth, the choice of the outcome variables should be consistent with the objective of the program being evaluated. In line with this view, these are basically the outcomes that the judges were implicitly tasked to predict.

Second, in a variable that can be measured in several ways, it is reasonable to consider more than one proxy measures so that the sensitivity of the result to the change of measures could be tested. In my case, business expansion or growth is measured by 4 different outcomes. This will not only allow me to show if the result remain robust across a range of growth measures and confidently conclude about the point being made but also provides alternative statistics for different users. For instance, the program owners like the Jobs Creation Commission (JCC) are more interested the potentials of the business in creating jobs and their outcome of interest would be employment. For investors who are contemplating to invest on startups, profit would be their outcome of interest as it informs them about the return to capital. Others may want to see the size of operation which could be inferred from the monthly sales figure. Some may be interested to gauge the overall prospect by a more aggregate measure like aggregate growth as it is emanated from multiple responses and likely to be more reliable.

Third, though my outcome variables are related of each other, their measurement do not suffer from carryover effect. Since these are business rather than psychological questions, the measure of the first outcome does not affect the latter in question ordering (i.e. no carryover effect) and in such scenario using each outcome separately will not create any problem (Chiang et al., 2015).

Finally, one intention of this study is to explain the reasons behind the contradictory findings of the previous studies on the same issue using results from two different business plan competitions in the same setting. In order to attribute various features of the program for observed results (or reasons), we need to keep the outcome variables the same as the previous studies. With different outcomes, neither comparison of the results with the previous findings nor the explanation for its heterogeneity will not be possible.

4.4. Results and discussions

This section presents the results and discussions pertaining to this chapter. As mentioned in the previous section, we examine the predictivity of score on five different outcomes which measures different aspects of the businesses. The first part presents the main results for the full sample and disaggregated sample by treatment status organized by the outcome variable, and in the latter sub-sections heterogeneity of results by types of competition and other aspects are explored.

4.4.1. Predicting firm entry and survival

Like many entrepreneurship development support programs, one of the objectives of the business plan competitions under study is to facilitate young potential entrepreneurs to setup their own business and/or help them survive and thrive. Consistent with this objective of the program, whether an applicant owned or were operating a business at the time of the follow-up survey is considered as my main outcome, which is also one of the measures of entrepreneurial success.

Before embarking on the full estimation results of the model for this outcome, I would like to explore the relationship between the standardized average score (score hereafter) and the predicted value of the outcome using binned scattered graphs. This will not only provide a quick preview of the relationship (positive, negative or zero) but also guides me about the choice of the functional form of the model I want to estimate. Therefore, graphical representation of the main finding will precede my regression tables throughout this chapter.

The relationship between probability of operating a business after one year of application to the competition and score is depicted in Figure 4.1 for the pooled sample. In this graph, the scattered plots (the black dots) are bin means computed with the same bin width of 4 and the sizes of the dots or circles are made to be proportional to the sample size within each bin. This also tells us the distribution of the sample over the entire support and the relative size of observations behind each bin mean. The scattered bins clearly suggest that the relationship is not linear, and the cubic specification is the best fit for this data. After running the cubic regression, we have conducted a linearity test and that was rejected at 1% with a p-value of 0.0024 (Table 4.A.1 in the appendix 4). The predicted probability was recovered from the linear probability model with linear and cubic specification and the former is included for comparison purpose. The fitted regression line from the linear model and the cubic models are represented by the dashed line and solid line, respectively.

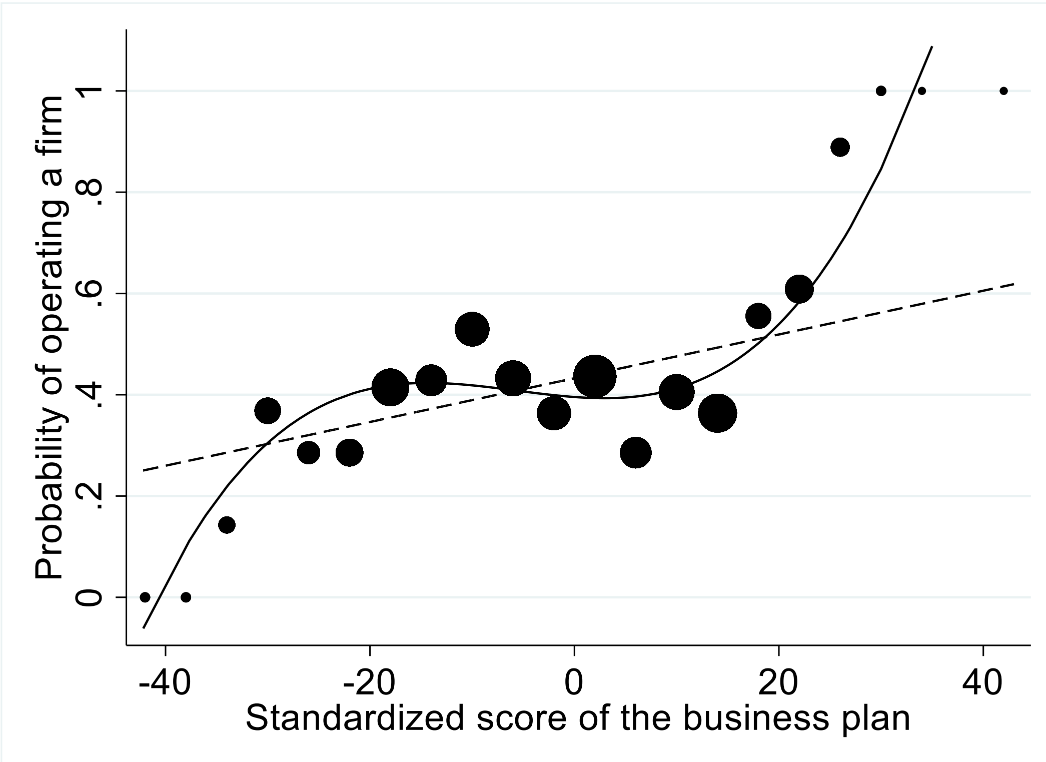
The positively sloped regression fitted lines in Figure 4.1 indicate that, in general, the probability of operating a business increase with score. That means the score from judges given at the time of application could predict future business entry and survival. A significant variation in the slope of the fitted regression line from my preferred model (cubic specification) shows the heterogeneity in the association between score and probability of operating a business over the entire distribution. The graph has a sharp positive slope in the bottom and the top of the distribution while it is almost horizontal around the middle of the distribution. This non-linear relationship suggest that the judges of business plan competitions were better in differentiating the most promising businesses or high-growth potential entrepreneurs and the least performing ones.

This result confirms the fragmented evidence regarding the relative performance of judges over the distribution of applicants in one experiment. For instance, while Fafchamps & Woodruff (2017) reported score was a better predictor of success in the bottom of the distribution, Scott et al. (2015) showed that judges were better in recognizing high-quality ideas as opposed to excluding non-series ones. However, it is only in the latter one that definition of my outcome is comparable.

In order to determine the predictive ability of the judges, we need to make sure that the correlation between score and outcome of interest should not be driven by the treatment. In this program, the main intervention that was accessible for many applicants was training. Though attendance to entrepreneurship training does not significantly affect the business outcomes as implied by out effective first-stage and reduced-form estimates in the previous chapter, we believe that the imprecisely measured coefficient is not the same as zero. Thus, controlling for the variation in training attendance status is essential to have a clean association between score and probability of owning a business. In addition to results for the full sample, we also provide results disaggregated by training status (trained or treated and non-trained or control applicants) for all outcomes considered throughout this chapter. Consistent with the previous chapter, an applicant is categorized as trained if she has ever taken any entrepreneurship training anywhere (in the program of interest or in any other similar program).

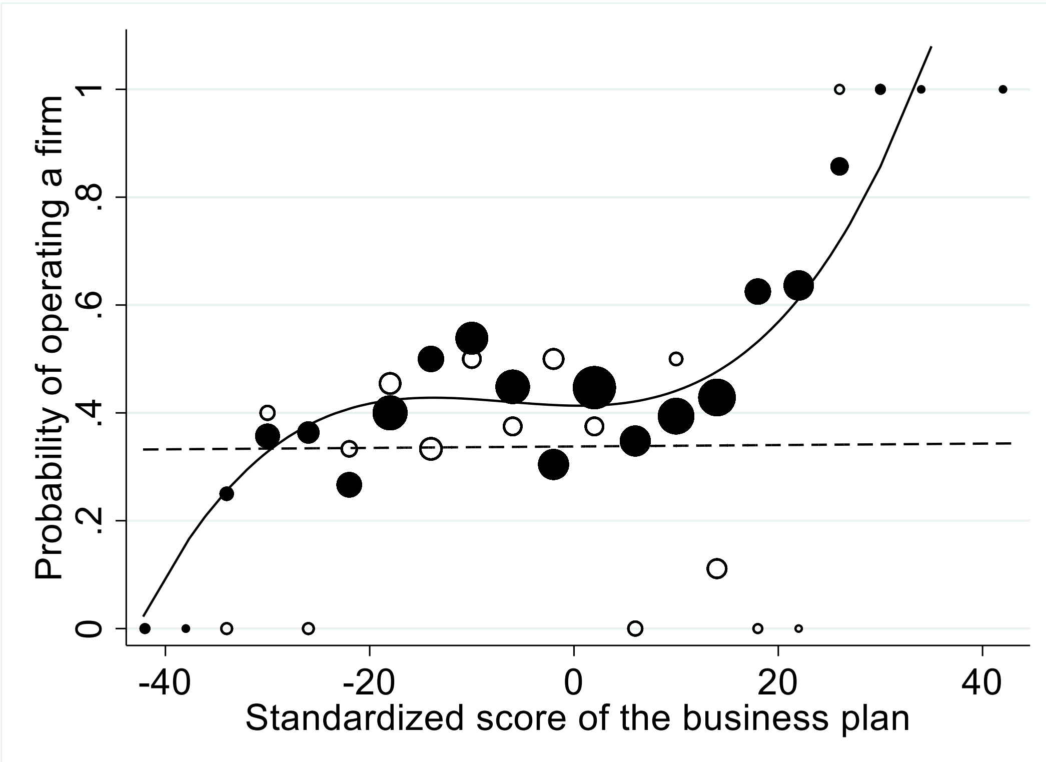
Figure 4.2 is an overlay binned scattered graph that depicts the relationship between score and probability of operating a firm for trained and non-trained sub-samples separately. For the trained sub-sample, the bin means are represented by black circle and the cubic fit drawn using the solid line is the best fit for this data as suggested by my test result, as the linearity hypothesis is rejected at 1% with a p-value of 0.0091(Table 4.A.1 in the appendix 4). For the non-trained sub-sample, however, the bin means are represented by the hallow circle and the dashed line is its linear fit. Like that of the full sample, score is a predictor of probability of operating a firm for the trained sub-sample. But the horizontal regression line fitted for the non-trained applicants' data shows that score is uncorrelated with the outcome for this sub-sample.

Figure 4. 1 Plots of score versus probability of operating a firm for the full sample



Notes: The dependent variable is a dummy variable whether the applicant was operating a business one year after the application and the model is estimated using the linear probability model for each specification. Scores are standardized at the cutoff. The scattered dots are bin means computed with a bin width of 4 and the size of the circle is proportional to sample size of each bin. The solid and dashed lines represent linear and cubic fits, respectively, for the full sample without other covariates.

Figure 4. 2 Plots of score versus probability of operating a firm for non-trained and trained sub-samples separately



Notes: Dependent variable in each group is a dummy variable whether the applicant was operating a business one year after the application and the model is estimated using the linear probability model. Scores are standardized at the cutoff. The scattered solid circles are bin means of the trained applicants while the hollow circles are for the non-trained applicants, and they were computed with a bin width of 4. The sizes of the circles are proportional to the sample size in each bin. The solid and the dashed lines represent cubic and linear fits for the trained and non-trained sub-samples, respectively, without other covariates.

Analysis of the marginal effects

It is important to check for non-linearities in the relationship between the variables of interest as the marginal effect varies across the distribution of the independent variables if non-linearity exists. I have demonstrated that the relationship between score and probability of operating a business is non-linear for the full sample and trained sub-sample. The linear and cubic model estimation results used to construct Figures 4.1 and Figure 4.2 are reported in appendix 4.1, Table 4.A.1.

Using coefficient estimates of those models, we have computed the change in probability of operating a business for a 10-percentage point change in scores at different positions in the distribution of scores for cubic specification, in comparison with results from the linear specification (Table 4.1).

For my outcome variable (let me say Y) and score (say X) which are related by a cubic function, we can drive the change in the probability of operating a business for any change in score (X) to be:

$$\Delta P(Y = 1|X) = \beta_1 \Delta X + \beta_2\left({X_2}^2 - {X_1}^2\right) + \beta_3\left({X_2}^3 - {X_1}^3\right) \qquad (4.3)$$

Where $\beta_1, \ \beta_2, \ and \ \beta_3$ are coefficients of the linear, quadratic and cubic terms, respectively. The corresponding standard errors can be computed from the resulting variance-covariance matrix and the given points of score data using a formula which can be easily derived from $var(\Delta P)$, in equation (4.3).[16]

As shown in column 1 of Table 3.1, a 10-percentage point increase in the score at the bottom of the distribution (-40 to -30) and at the top of the distribution (20 to 30) leads to a 27.7 and 30.7 percentage points, respectively, rise in probability of operating a business for the full sample and each of them are statistically significant at 1%. Similarly for the treated sub-sample (column 2), the cubic model yields a big change in the outcome variable for the same 10-percentage point change in score at the top and bottom of the distribution. However, in the middle of the distribution 10-percentage point change in score, say from -20 to -10, leads to almost no change in the

[16] Implementing the above ΔY formula in "*lincom*" command of Stata can also give us the required standard errors

probability of operating a business for both full sample and trained sub-sample. The result of this marginal analysis is quite consistent with what we observe in Figure 4.1 and Figure 4.2.

However, in the linear specification, we have the same marginal effect, about 0.43 percentage point for full sample and 0.48 percentage points for the trained sub-sample. This highlights how ignoring non-linearities in estimating such models would be misleading.

Table 4. 1 Simulating the effects of changes in standardized scores on probability of operating a firm for cubic specification in comparison with the liner model

Change in standardized score	Change in probability of operating a firm			
	Cubic model		Linear model	
	Full sample	Trained sample	Full sample	Trained sample
From -40 to -30	0.2767***	0.2307**	0.0432***	0.0485***
	(0.0855)	(0.0922)	(0.0140)	(0.0157)
From -20 to -10	0.0069	0.0084	0.0432***	0.0485***
	(0.0219)	(0.0246)	(0.0140)	(0.0157)
From 0 to 10	0.0022***	0.0275***	0.0432***	0.0485***
	(0.0462)	(0.0252)	(0.0140)	(0.0157)
From 20 to 30	0.3066***	0.2882***	0.0432***	0.0485***
	(0.0805)	(0.0795)	(0.0140)	(0.0157)
Observations	456	358	456	358
R-squared	0.0443	0.0474	0.0189	0.0239

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. These marginal effects were computed using the coefficients reported in Appendix 4.1, Table 4.A.1.

I have also formalized this heterogeneity in predicted probabilities over the distribution of the score using a quantile analysis. In appendix 4.1, Figure 4.A.1, the probability of operating a business is depicted against the score quintiles. I show that the average probability of operating a business is 54.4% in the fifth quintile while it is only 30.4% in the first quintile. Nonetheless, the unclear pattern of the relationship in the second, third, and fourth quintile implies that judges are unable to precisely rank applicants at the middle of the distribution while they precisely differentiate extreme performances. This results further bolsters the findings which what have been presented so far.

The main model results with and without covariates

The graphical analysis and related estimates presented so far did not include the fixed effects and other baseline covariates. Based the econometric model specified in equation 4.1, the full model results in comparison with and without covariates are reported in Table 4.2. While the results of Table 4.2 are estimates of the linear probability model, the marginal effects from the Probit and Logit model are also estimated as a robustness check and the result remain the same, as reported in Appendix 4.2, Table 4.A.2A and Table 4.A.2B. Table 4.2 summarizes results of different specification for three groups: the full sample, the non-trained sub-sample, and the trained sub-sample. Taking the full sample into account for instance, in column 1, the model includes only the variable of interest (score) as the only regressor since I am interested here to see predictive power of only judges' score without any covariates. In column 2, I included dummy for existing business at the time of application since near to 29% of the applicants had been operating businesses and I controlled for this variation. McKenzie & Sansone (2019) also use the same specification (or control) from which they gauge the predictiveness of score without survey variable and we reported results with this specification to facilitate comparison with their finding.

In column 3, I included additional controls, namely, gender (measured by female dummy); education of the entrepreneur which is categorized as high school or below (base category), Technical and vocational Education and Training (TVET) and any other non-degree college education, and undergraduate or graduate degree; sector fixed effect where it is defined as construction (the base category), agriculture, Information Technology (IT), manufacturing, and retail (includes trade and other services); panel of judges fixed effect, and regional fixed effect. This specification almost the same as what Fafchamps & Woodruff (2017) reported as results of the score without including survey variables.[17] Since the above two papers are the only empirical works available in low-income countries, but with contradicting results, having additional results with the same specification as these papers will have a paramount importance for the move towards generating a conclusive evidence in this area.

As shown in Table 4.2, column 1 to 3, score from judges is positively and significantly related to probability of operating a business. The result remains significant as we include other controls though the magnitude of the coefficient on score falls with some extent. Consistent with the previous discussion, score is significantly correlated with the outcome for the trained sub-sample (column 7 to 9). However, the coefficients of score in all specifications of the non-trained sub-sample (column 4 to 6) are statistically indistinguishable from zero which could be because I do not have enough statistical power for this group as the sample size is only 98.

The other variable which is worth to mention is existing dummy. In all specification of the full and trained samples, coefficient of existing dummy is large and statistically significant at 1%, implying that it is a strong predictor of probability of operating a business. It has also contributed a lot to the overall fitness of the model as the adjusted R-squared jumped when this variable is controlled for though the goodness of fit of the model generally remains low.[18] . The coefficient of existing dummy, for instance in column 2 of Table 4.2, is interpreted as applicants with young businesses in operation at the time of application are 22.6 percentage points more likely to be found operating

[17] See column 2 of Table 4 and the notes of the same table in Fafchamps & Woodruff (2017). Ability score, credit, and management practice as control of survey variables. However, these variables were not purely constructed from the baseline data (for example. credit) and using data collected during the follow-up survey is incorrect in such prediction exercise. As a result, I refrain from using such types of controls.
[18] I will return to the discussion of R-squared shortly.

a business after a year as compared to applicants who had just the business idea. This implies surviving the existing businesses is easier that launching the new ones. Further, level of education of the entrepreneur is also correlated with probability of operating a business.

Table 4. 2 Main results about the prediction of firm ownership and survival by panel of experts

	Full Sample			Non-trained			Trained		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Judges' Score	0.0043***	0.0033**	0.0030*	0.0001	0.0000	-0.0032	0.0049***	0.0036**	0.0038**
	(0.0014)	(0.0014)	(0.0016)	(0.0032)	(0.0034)	(0.0045)	(0.0016)	(0.0016)	(0.0018)
Existing firm		0.2257***	0.1933***		0.1519	0.1778		0.2347***	0.1960***
		(0.0508)	(0.0546)		(0.1205)	(0.1252)		(0.0562)	(0.0608)
Agriculture sector			0.0139			-0.0668			0.0375
			(0.1134)			(0.3227)			(0.1265)
IT sector			0.0248			-0.0274			0.0613
			(0.1051)			(0.2977)			(0.1184)
Manufacturing sector			0.0330			-0.2779			0.0877
			(0.1036)			(0.3030)			(0.1147)
Retail sector			-0.0986			-0.1679			-0.0537
			(0.1042)			(0.3034)			(0.1170)
Female			-0.0560			0.0674			-0.0896
			(0.0593)			(0.1521)			(0.0659)
TVET or some college			0.2194**			0.4745**			0.1891*
			(0.0894)			(0.1996)			(0.1004)
Undergraduate or graduate			0.0616			0.2471**			0.0382
			(0.0576)			(0.1230)			(0.0660)
Constant	0.4326***	0.3639***	0.2574**	0.3377***	0.3028***	0.1799	0.4510***	0.3765***	0.2360*
	(0.0234)	(0.0274)	(0.1122)	(0.0538)	(0.0601)	(0.3420)	(0.0261)	(0.0312)	(0.1233)
Regional FE	No	No	Yes	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	456	456	456	98	98	98	358	358	358
R-squared	0.0189	0.0609	0.103	1.63e-05	0.0180	0.159	0.0239	0.0700	0.125
Adjusted R-squared	0.0168	0.0568	0.0700	-0.0104	-0.00268	-0.00766	0.0212	0.0648	0.0842

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is having a business in operation one year after the application. Estimation is based on a linear probability model

4.4.2. Predicting employment

The issue of job creation has been a central agenda of the Ethiopian government since the economy is challenged by high urban unemployment. For instance, at the start of this in 2021, the urban youth unemployment rate of the country was as high as 23.1%(CSA, 2021). Given this situation, spurring job creation is not only the very objective of this program but also the sole reason for the existence of the jobs creation commission (JCC), the owner of *Bruh*. When judges were tasked to score business plans of the contenders, the job creation potential of businesses was considered and hence it would be appropriate to assess how judges are accurate in predicting the labor absorptive capacity of businesses.

To this effect, we measured employment by the total numbers of workers at the time of the follow-up visit and used it as an outcome while score is still my predictor variable. In this case, I defined the outcome to be unconditional on operating a business by coding level of employment to be zero for those who do not own businesses at the time of the follow-up survey. Doing so is important to handle the problem of sample selection. For the same reason, I have also implemented the similar technique and definition for the remaining other outcome variables which are presented in the next sub-sections.

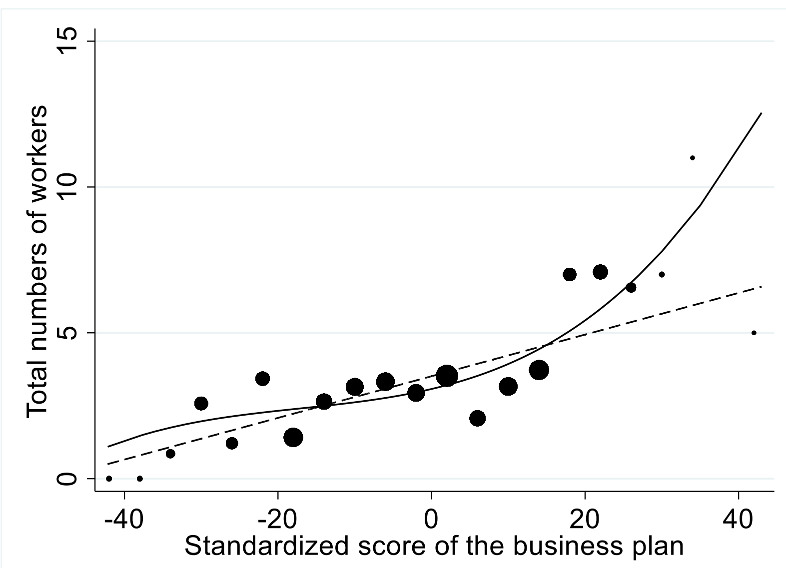
Like that of the previous section, I started with the binned scattered graph with a linear and cubic specification to visualize the relationship between score and level of employment for the full sample (Figure 4.3).[19] In this case, however, I fail to reject the linearity hypothesis with a p-value of 0.2510, 0.1631, and 0.5065 for the full sample, trained sub-sample, and non-trained sub-sample, respectively. Therefore, my preferred specification will be the polynomial of order 1 (dashed line)

[19] The way we constructed the graphs is the same as described in the previous section for the first outcome.

for this outcome though I also included the cubic regression fit line (solid line) in Figure 4.3, just for comparison purpose.

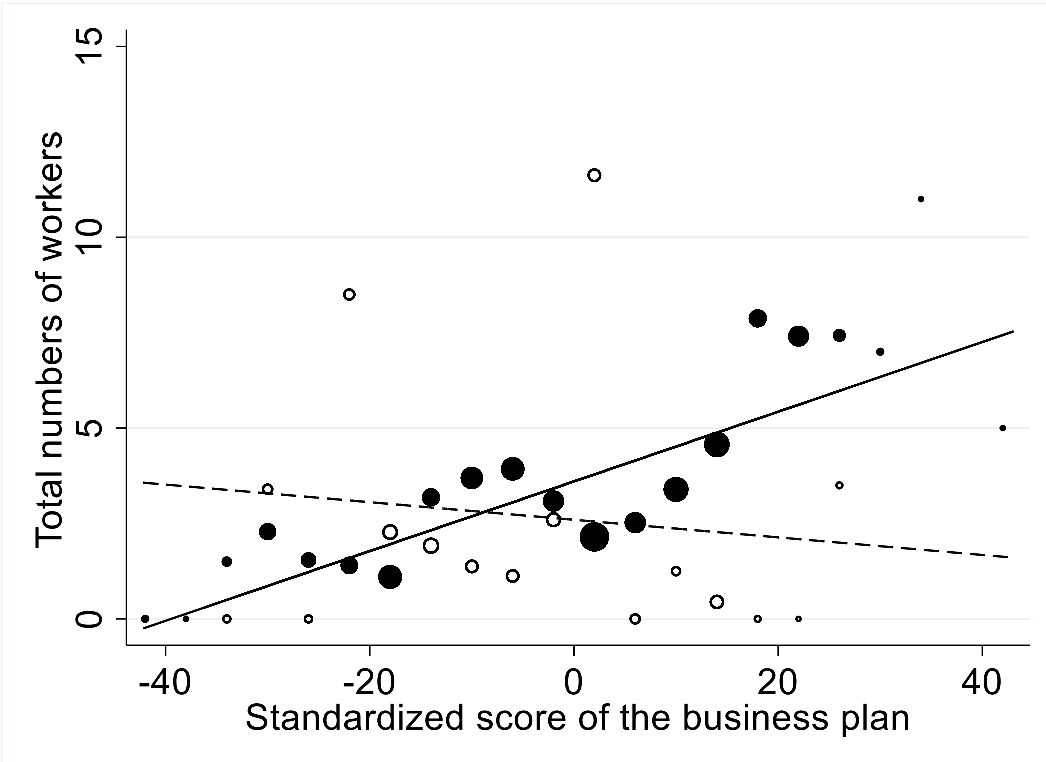
The upward sloping regression fit line in Figure 4.3 clearly shows that score is a predictor of future entrepreneurial success measured by level of employment. Disaggregating the sample by the training status yield different results for the trained and the non-trained applicants. Figure 4.4 presents an overlay scattered binned graphs for the disaggregated samples where the black circle and solid line are the scattered mean and linear regression fit for trained applicants while the hallow circle and the dashed line represent the same statistics for their non-trained counterparts. As shown in Figure 4.4, while level of employment increases with score for the trained sub-sample, it has somehow declined for non-trained sub-sample though the magnitude of the latter is small as one can tell from a near to horizontal slope of the dashed line.

Figure 4. 3 Binned scatter plots of score Versus level of employment for full sample



Notes: Dependent variable is a total number of workers measured a year after the application. Scores are standardized at the cutoff. The relationship is unconditional on operating a business which codes numbers of workers as zero for applicants who are not operating a firm. The scattered dots are bin means computed with a bin width of 4 and the size of the circle is proportional to the sample size in each bib. The solid line and dashed lines are cubic and linear fits, respectively, for the full sample without other covariates.

Figure 4. 4 Binned scattered plots of score versus level of employment for trained and non-trained sub-samples



Notes: Dependent variable is a total number of workers measured a year after the application. Scores are standardized at the cutoff. The relationship is unconditional on operating a business which codes numbers of workers as zero for applicants who are not operating a firm. The scattered solid circles are bin means of the trained applicants while the hollow circles are for the non-trained ones, and they were computed with a bin width of 4. The sizes of the circles are proportional to the sample size in each bin. The solid and dashed lines represent the linear fits for trained sub-sample and non-trained sub-sample, respectively, without other covariates.

The full regression result on employment

The full model result in comparison with and without controlling for covariates are presented in Table 4.3 for the full and disaggregated samples in the same manner that we did for the first outcome in the previous section. The model results shown in Table 4.3 revealed that business plan scores from juries have strongly predicted future employment level of the applicants. The result remains consistently significant at 1% almost in all cases despite the inclusion and exclusion of the baseline covariates as well as the judges, sector, and regional fixed effects for both full sample (column 1 to 3) and the trained sub-sample (column 7 to 9). As we have seen in the graphical representation, the estimates of coefficients for score remains statistically not different from zero in all the three specifications for the non-trained group. In addition, the estimated coefficients for the existing dummy imply that participants with the existing micro or small enterprise are more likely to expand and employ 2 to 3 additional workers as compared to new entrants.

The other worth mentioning issue in this exercise is the goodness of fit of the model, as measured by R-squared and adjusted R-squared. As shown in Table 4.3 the adjusted R-squared is generally low with a maximum of 12.5% in column 9. This implies that only a small proportion of the variation in employment is explained by the variation in scores even after controlling for the baseline covariates and the fixed effects. This not unique for this study as all other similar previous studies also reported small adjusted R-squared.

For instance, from a relatively larger sample experiment with the same definition of outcome and controls McKenzie & Sansone (2019) reported their adjusted R-squared to be 0.053 for treated group (grant winners in their case) in exactly the same specification as column 8 of Table 4.3 where my estimate is 0.0866. For the same sub-sample with control of additional covariates, like Column 9 in my case, their adjusted R-Squared is 0.072 while my adjusted R-squared is 0.1250. Again, for the control groups, their adjusted R-squared with and without additional covariates are 0.012 and 0.038, respectively, while mine in this regard is -0.0114 and 0.0055 as indicated in column 5 and 6 of Table 4.3. These similar findings reaffirm that the variation in entrepreneurial success is better explained by other factors than scores and baseline covariates.

Table 4. 3 Prediction of firm level of employment by panel of experts

	Full Sample			Non-trained			Trained		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Judges' Score	0.0714***	0.0564***	0.0479**	-0.0231	-0.0244	-0.0739	0.0913***	0.0735***	0.0529***
	(0.0214)	(0.0201)	(0.0199)	(0.0495)	(0.0509)	(0.0654)	(0.0228)	(0.0202)	(0.0161)
Existing firm		3.1778***	2.2883**		2.0223	1.2941		3.3338***	2.5244***
		(0.8494)	(0.9468)		(2.2475)	(2.3796)		(0.8591)	(0.8528)
Agriculture sector			-0.8377			0.4174			-1.4975
			(1.9478)			(5.5520)			(1.9628)
IT sector			-2.4705			-3.4855			-1.9959
			(1.5970)			(3.9616)			(1.7183)
Manufacturing sector			-0.9238			-5.7142			-0.3935
			(1.6094)			(4.3185)			(1.7239)
Retail sector			-1.6909			-1.2887			-1.5501
			(1.6409)			(4.5105)			(1.6958)
Female			-0.8299			-1.3622			-0.7235
			(0.8295)			(1.5946)			(0.9591)
TVET or some college			3.8224*			10.9931			2.3334
			(2.2024)			(7.7293)			(1.7374)
Undergraduate or graduate			0.5907			1.8768			0.7774
			(0.5480)			(1.1503)			(0.6639)
Constant	3.5119***	2.5446***	2.2576	2.5972**	2.1331	0.9649	3.6020***	2.5429***	1.7623
	(0.3746)	(0.3796)	(1.6098)	(1.0596)	(1.3239)	(4.5747)	(0.3782)	(0.3410)	(1.7015)
Regional FE	No	No	Yes	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	456	456	456	98	98	98	358	358	358
R-squared	0.0224	0.0584	0.119	0.00127	0.00946	0.170	0.0437	0.0917	0.164
Adjusted R-squared	0.0202	0.0543	0.0869	-0.00914	-0.0114	0.00551	0.0410	0.0866	0.125

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is total number of workers one year after the application.

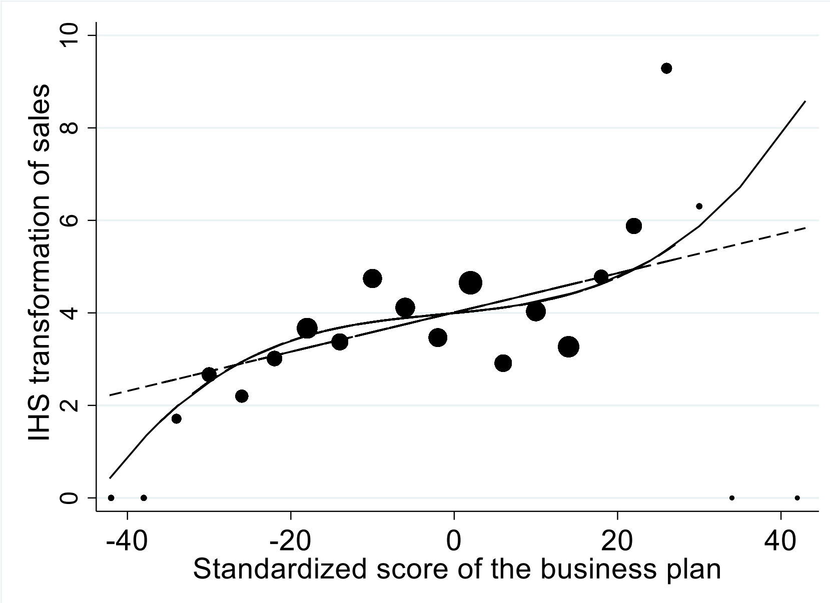
4.4.3. Predicting Sales

The other outcome we used to test the predictive ability of judges is monthly sales in Ethiopian Birr. Due to the righthand long tail in the sales data, I transformed it into inverse hyperbolic sine values. This outcome not only measures the expansion or size of the firm but also used to triangulate the self-reported status regarding business entry for the startups since respondents were also asked additional related information like the costs incurred and types of products they sold recently.

Further, sales and profit are relatively more difficult to precisely measure using self-reported survey method as compared to other outcomes like business entry and employment. Given this distinct nature of sales and profit, it would be interesting to test if judges' prediction becomes less precise when the outcome is not reasonably predictable (Kahneman & Klein, 2009) or the task character, which is business performance in a highly uncertain environment in this case, is less predictable (Shanteau, 1992).

Before embarking on the full model result for this outcome, we depicted the binned scattered plots of score against the IHS transformation of sales with linear and cubic fit in Figure 4.5. Again, for this outcome the linear specification is the best fit for the sales data as the we fail to reject the linearity hypothesis with a P-value of 0.5049. As shown in Figure 4.5, sales revenue of firms is, in general, positively correlated with initial evaluation of judges measured by average score. This positive relationship is also maintained when the sample is disaggregated by training status though it is stronger for the trained group as implied by the stepper regression fit line (solid line fitted for the black circles) in Figure 4.6.

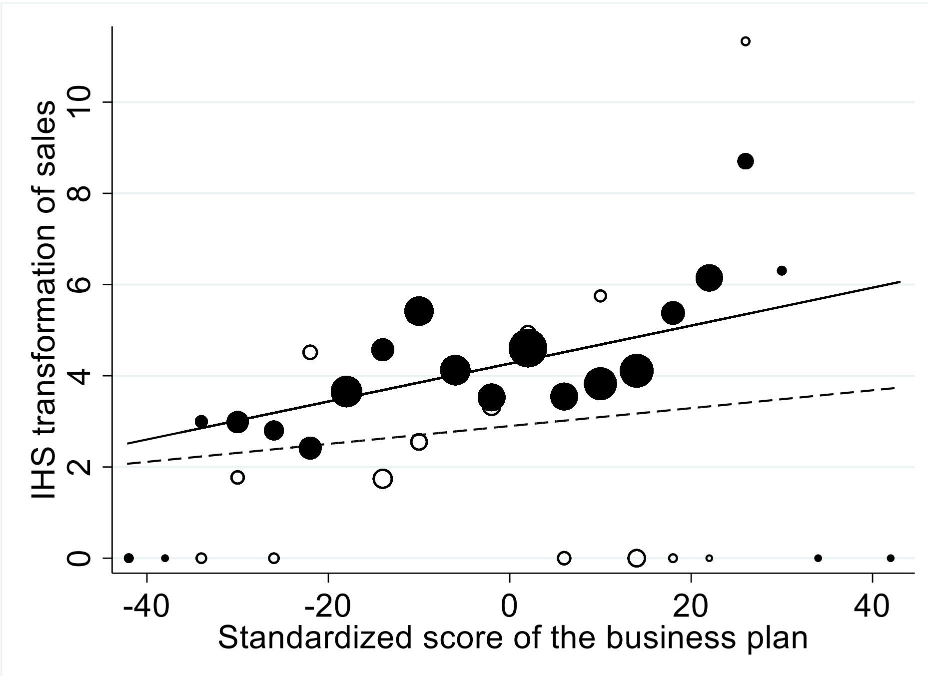
Figure 4. 5 Binned scatter plots of score versus the IHS transformation of sales for full sample



Notes: Dependent variable is the Inverse hyperbolic transformation of monthly sales in Birr. Scores are standardized at the cutoff. The relationship is unconditional on operating a business which codes sales as zero for applicants who are not operating firms. The dots are bin means computed with a bin width of 4 and their size are proportional to the sample size in each mean. The solid line and dashed line represent cubic and linear fits, respectively, for the full sample without other covariates.

Table 4.4 presents the results of the full model for sales. As indicated in this table, scores positively associated with sales and the estimated coefficient for the full sample is statistically significant at 1% when no control is included (column 1) and at 10% when existing dummy is added (column 2). Nonetheless, the correlation gets weaker and the statistical significance wither away when the baseline covariates and fixed effects are included (column 3). Again, in the disaggregated results presented from column 4 to column 9 of Table 4.4, statistically speaking the coefficients on score are not indistinguishable from zero in all case but one. The whole results for this outcome imply that judges are relatively less successfully in predicting sales performance. This finding confirms the theoretical prediction that the more difficult the outcome is to measure, the less accurate the experts' prediction would be.

Figure 4. 6 Binned scattered plots of score versus the IHS transformation of sales for non-trained and trained applicants



Notes: Dependent variable for each group is the Inverse hyperbolic transformation of monthly sales in Birr. Scores are standardized at the cutoff. The relationship is unconditional on operating a business which codes sales as zero for applicants who are not operating firms. The scattered solid circles are bin means of the trained applicants while the hollow circles are for the non-trained ones, and they were computed with a bin width of 4. The solid and dashed lines represent the linear fits for trained sub-sample and non-trained sub-sample, respectively, without other covariates.

On the other hand, though the correlation between score and sales is modest over full support, the judges are still successful in distinguishing highest and least performing business based on this measure too as shown by quintal results in appendix 4.4, Figure 4.A.3.

Table 4. 4 Prediction of firm sales by panel of experts

	Full Sample			Non-trained			Trained		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Judges' Score	0.0425***	0.0304*	0.0211	0.0196	0.0184	-0.0329	0.0416**	0.0271	0.0232
	(0.0154)	(0.0155)	(0.0176)	(0.0334)	(0.0343)	(0.0450)	(0.0175)	(0.0177)	(0.0202)
Existing firm		2.4518***	2.0743***		1.3096	2.1020		2.6287***	2.2075***
		(0.5808)	(0.6127)		(1.3413)	(1.3103)		(0.6436)	(0.6841)
Agriculture sector			0.5934			1.2094			0.5279
			(1.2305)			(2.5196)			(1.3999)
IT sector			0.0054			0.3997			0.2628
			(1.1014)			(2.1907)			(1.2736)
Manufacturing sector			0.3040			-1.6133			0.5075
			(1.0831)			(2.1675)			(1.2295)
Retail sector			-0.6936			-0.3788			-0.5682
			(1.0878)			(2.2725)			(1.2461)
Female			-1.0425*			-0.4979			-1.1377
			(0.6261)			(1.3385)			(0.7152)
TVET or some college			1.6250			4.8937**			0.9940
			(1.0495)			(2.3850)			(1.1800)
Undergraduate or graduate			0.2923			1.7843			0.0795
			(0.6163)			(1.1942)			(0.7096)
Constant	4.0099***	3.2730***	2.4858**	2.8978***	2.6054***	1.2623	4.2676***	3.4405***	2.4410*
	(0.2581)	(0.2932)	(1.1546)	(0.5764)	(0.6460)	(2.6797)	(0.2907)	(0.3351)	(1.2930)
Regional FE	No	No	Yes	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	452	452	452	97	97	97	355	355	355
R-squared	0.0156	0.0572	0.0899	0.00338	0.0156	0.151	0.0147	0.0626	0.105
Adjusted R-squared	0.0134	0.0530	0.0564	-0.00711	-0.00532	-0.0186	0.0120	0.0573	0.0629

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is the Inverse hyperbolic sine transformation of monthly sales in Birr.

4.4.4. Predicting profit

Profit is the other outcome considered in this study to test the judges' predictive ability in outcomes that is likely to be measured with self-reporting errors alike sales as I discussed in the previous section. In principle, profit for early stage-startup like my target group may not be a good measure of performance as some startups are expected to make various investments and may incur loss at their very stage with future expected returns. This is particularly true in advanced economies where external finance options such as angel investors, equity investment, and debt financing are widely available for startups.

However, in developing countries like Ethiopia external sources of formal finance for startups and small businesses are scant. For instance, my survey reveals that from the universe of applicants of this program, only 0.40% (just 2 of 494 entrepreneurs), 3.03%, and 4.05% had access to angel or equity investment, bank loan, and loan from Micro Finance Institutions (MFI), respectively. This implies that they apparently rely on either own source or finance from families and friends. Given this situation, I do not expect the startups considered in this study tolerate any loss for even one year, as opposed to startups in advanced economies that enter with a good financial stance. Therefore, profit cannot be ruled out even from being a measure of expansion or firm performance in this context.[20]

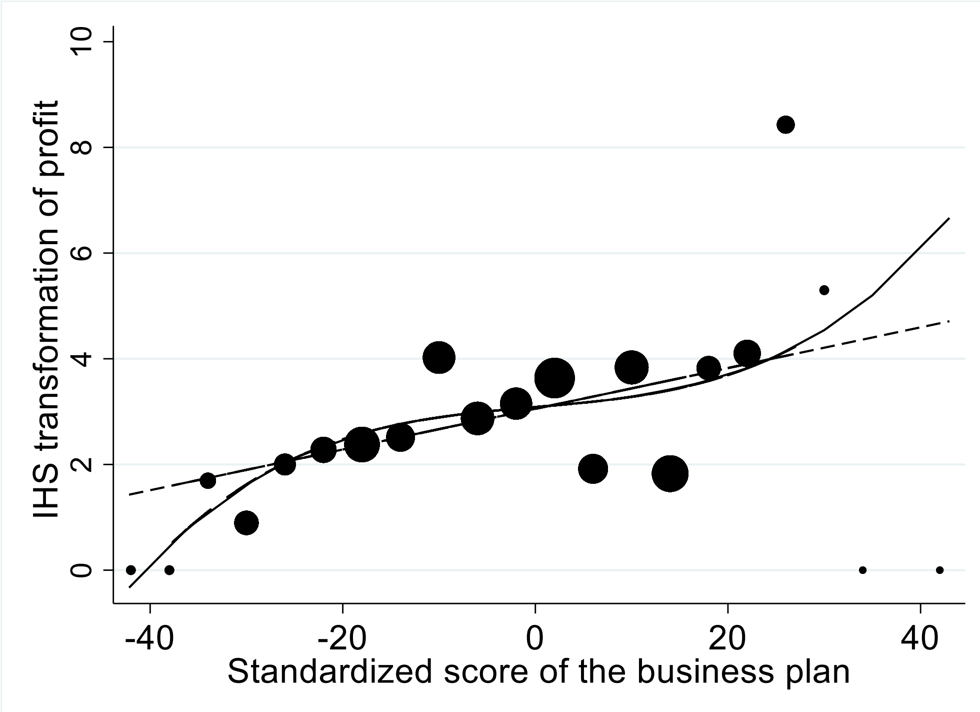
Now let's explore the performance of judges' prediction ability using profit as an outcome. Figure 4.7 depicts the binned scatter plots of score against the IHS transformation of monthly profit measured at the time of follow up survey in Ethiopian Birr. Like that of employment and sales, polynomial of order 1 function with score fits the data well as evident by the test statistics with a

[20] It could be because of a similar logic that other studies like McKenzie & Sansone (2019) used profit as a outcome in similar study where startups are included in their target group.

p-value of 0.5024 for the test for linearity. The positively sloped regression fit line in Figure 4.7 generally shows that score from business plan judges is positively correlated with profit. Further disaggregating the analysis by the training status in Figure 4.8 gives the same trend.

However, whether the slopes we view in these graphs are large enough to conclude about the score-profit nexus should be judged from the statistical significance of the coefficients estimated by include other covariates too. Table 4.5 presents the model results with and without covariates and fixed effects. Accordingly, the correlation between score and profit is statistically significant before including the full baseline covariates and fixed effect in the case of full sample (column 1 and column 2) and without any controls for the treated sub-sample (column 7). When the baseline covariates like gender and education as well as the fixed effects are included the coefficients of score become statistically insignificant. This result reaffirms the notion that prediction for some outcomes like profit and sales are more difficult than others like employment or business entry. In their out-of-sample prediction exercise from experts, models, machine learning techniques, McKenzie & Sansone (2019) also noted the difficulty of prediction particularly for sales and profit outcomes.

Figure 4. 7 Binned scattered plots of score versus the IHS transformation of profit for full sample

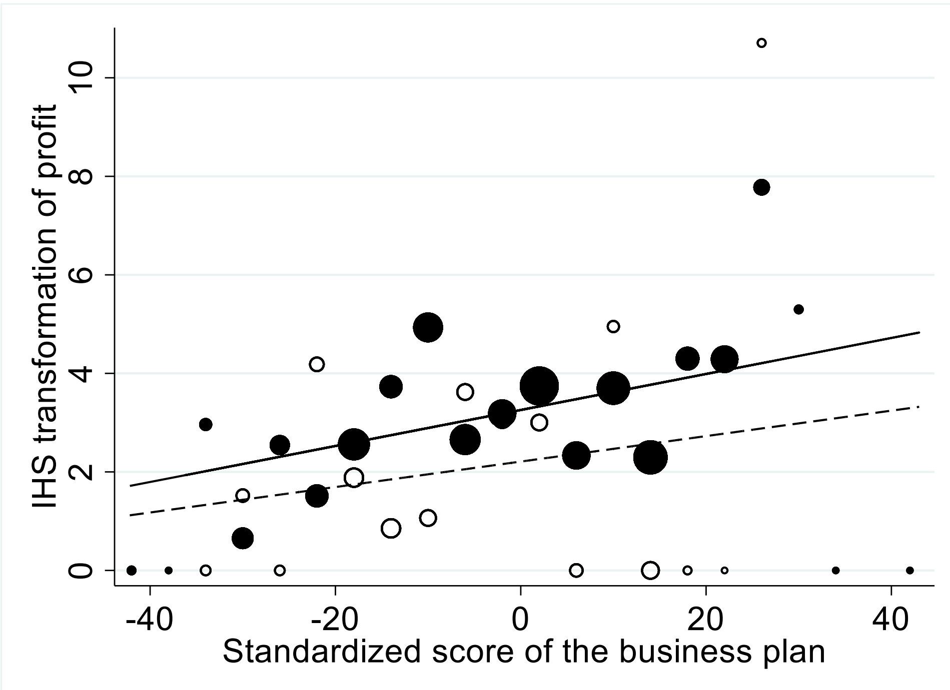


Notes: Dependent variable is the Inverse hyperbolic transformation of monthly profit in Birr. Scores are standardized at the cutoff. The relationship is unconditional on operating a business which codes profits as zero for applicants who are not operating a firm. The dots are bin means computed with a bin width of 4 and their size are proportional to the sample size in each mean. The solid line and dashed line represent cubic and linear fits, respectively, for the full sample without other covariates

Alike that of the other outcomes considered so far, only small fraction of the variation in the profit is explained by the variation in score and other covariates as implied by the estimated adjusted R-squared. Among the previous studies in this area, I can compare the finding regarding the goodness of fit of the model with what is reported in McKenzie & Sansone (2019) since they had unconditional profit with the same definition as an outcome. For instance, the adjust R-squared they reported for the treated sub-sample without controls of all covariates is 0.011while mine is 0.048, and when other controls are included their adjusted R-squared changed to 0.022 while it is

0.035 in my case.[21] Therefore, ending up with explaining only the small variation of such business outcomes seems a common feature of this area rather than being an exception. This is sometimes associated with the riskiness inherent in entrepreneurship particularly in highly volatile business environment like Ethiopia(Hall & Woodward, 2010).

Figure 4. 8 Binned scattered plots of score versus the IHS transformation of profit for non-trained and trained applicants



Notes: Dependent variable for each group is the Inverse hyperbolic transformation of monthly profit in Birr. Scores are standardized at the cutoff. The relationship is unconditional on operating a business which codes profit as zero for applicants who are not operating firms. The scattered solid circles are bin means of the trained applicants while the hollow circles are for the non-trained ones, and they were computed with a bin width of 4. The solid and dashed lines represent the linear fits for trained sub-sample and non-trained sub-sample, respectively, without other covariates.

[21] In another study on business plan competition, Fafchamps & Woodruff (2017) reported a relatively higher adjusted R-squared due to the fact that they studied only established firms and as results they were able to include baseline values of the outcomes among the controls. As pointed out by McKenzie & Sansone (2019), this is likely to exhibit persistence and yield higher Adjusted R-squared.

Table 4. 5 Prediction of firm profit by panel of experts

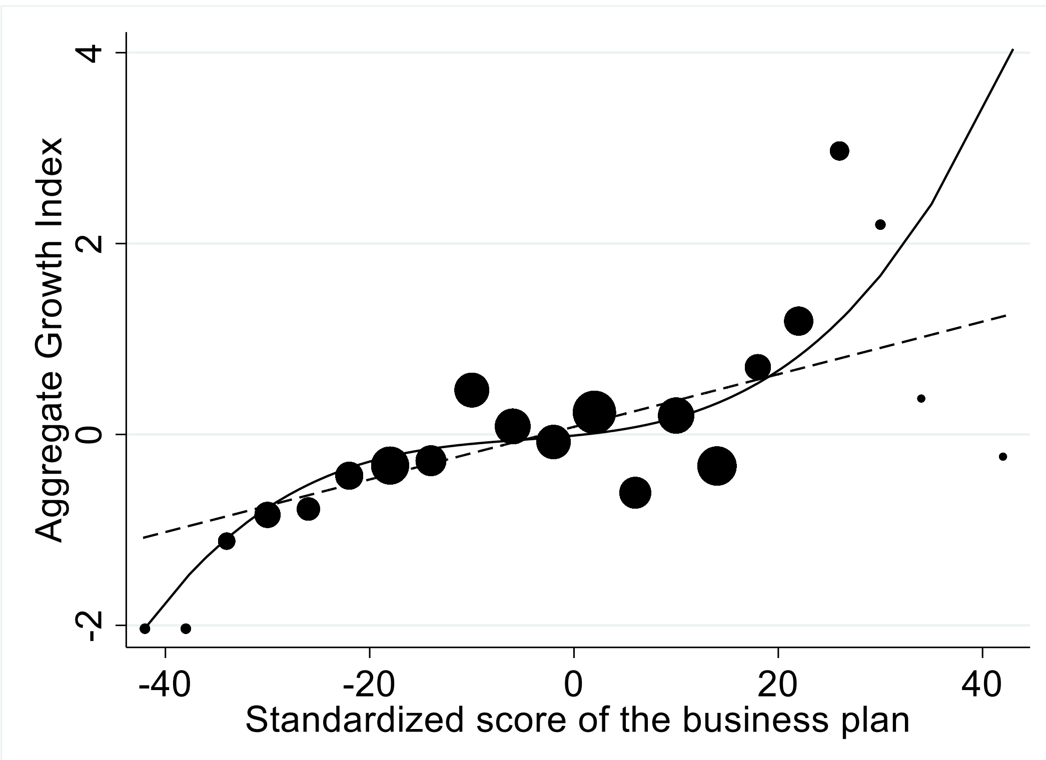
	Full Sample			Non-trained			Trained		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Judges' Score	0.0385**	0.0274*	0.0252	0.0259	0.0246	0.0049	0.0365**	0.0231	0.0226
	(0.0153)	(0.0153)	(0.0185)	(0.0304)	(0.0312)	(0.0434)	(0.0178)	(0.0177)	(0.0220)
Existing firm		2.2901***	2.0655***		1.2986	2.0201		2.4590***	2.2240***
		(0.5774)	(0.6197)		(1.2200)	(1.2561)		(0.6523)	(0.7172)
Agriculture sector			-0.1871			-1.1458			-0.0036
			(1.1182)			(2.6578)			(1.2593)
IT sector			-0.4722			-1.0905			-0.2325
			(0.9844)			(2.4517)			(1.1389)
Manufacturing sector			0.1488			-2.1240			0.3516
			(0.9549)			(2.3416)			(1.0795)
Retail sector			-0.7370			-1.7141			-0.5419
			(0.9444)			(2.4129)			(1.0764)
Female			-1.1601*			-1.2567			-1.0803
			(0.6146)			(1.2634)			(0.7109)
TVET or some college			1.4797			4.1501*			0.8064
			(1.0768)			(2.2051)			(1.2583)
Undergraduate or graduate			0.2621			1.3983			0.0205
			(0.6186)			(1.1656)			(0.7421)
Constant	3.0538***	2.3650***	2.2208**	2.2101***	1.9201***	2.6269	3.2574***	2.4828***	2.1344*
	(0.2553)	(0.2894)	(1.0517)	(0.5269)	(0.5957)	(2.9119)	(0.2920)	(0.3352)	(1.1892)
Regional FE	No	No	Yes	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	451	451	451	97	97	97	354	354	354
R-squared	0.0131	0.0502	0.0741	0.00665	0.0203	0.141	0.0112	0.0530	0.0789
Adjusted R-squared	0.0109	0.0460	0.0400	-0.00380	-0.000554	-0.0308	0.00843	0.0476	0.0351

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is the inverse hyperbolic sine transformation of monthly profit in Birr.

4.4.5. Predicting Aggregate growth

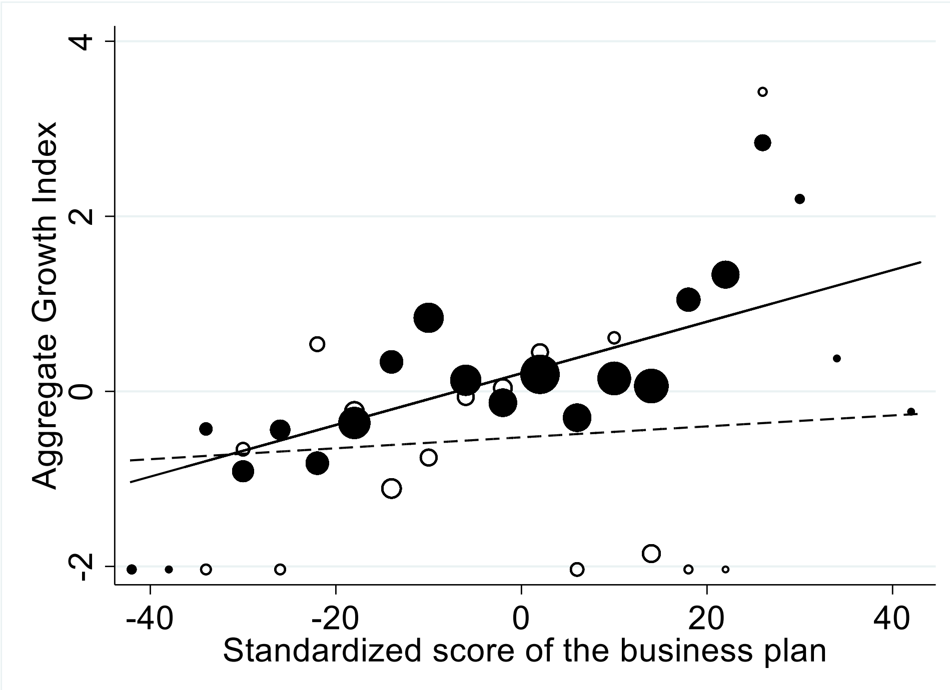
In order to summarize the overall success of judges in predicting future entrepreneurial success using a more aggregate growth measure, I have constructed aggregate growth index from employment, sales, and profit data as described in section 4.3. Having this outcome in this study is not only useful for showing the overall picture of business operations of applicants but also to compare out results with the previous studies like Fafchamps & Woodruff (2017) that used aggregate growth to draw a conclusion about the same research question.

Figure 4. 9 Binned scattered plots of score against aggregate growth for the full sample



Notes: Dependent variable is aggregate growth which is a sum of standardized values of the Inverse hyperbolic transformations of unconditional sales, unconditional profit and unconditional employment. Scores are standardized at the cutoff. The relationship is unconditional on operating a business which codes outcomes that form aggregate growth as zero before their transformation for applicants who are not operating a firm. The dots are bin means computed with a bin width of 4 and their size are proportional to the sample size in each mean. The solid line and dashed line represent cubic and linear fits, respectively, for the full sample without other covariates

Figure 4. 10 Binned scattered plots of score against aggregate growth for non-trained and trained applicants



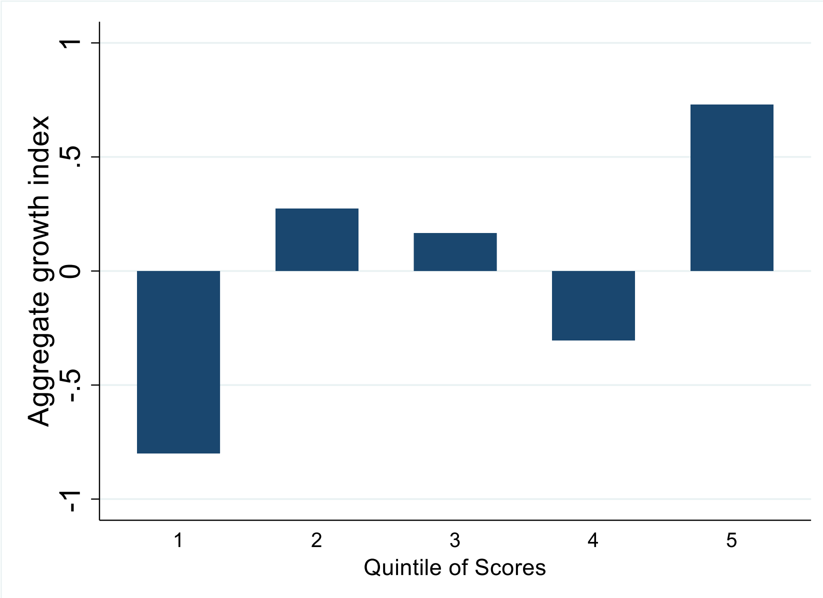
Notes: Dependent variable for each group is aggregate growth which is a sum of standardized values of the inverse hyperbolic transformations of unconditional sales, unconditional profit and unconditional employment. Scores are standardized at the cutoff. The relationship is unconditional on operating a business which codes outcomes that form aggregate growth as zero before their transformation for applicants who are not operating a firm. The scattered solid circles are bin means of the trained applicants while the hollow circles are for the non-trained ones, and they were computed with a bin width of 4. The solid and dashed lines represent the linear fits for trained sub-sample and non-trained sub-sample, respectively, without other covariates.

The association between score and aggregate growth is depicted using a binned scattered plots for the full sample (Figure 4.9) as well as trained and non-trained sub-samples (Figure 4.10). The linearity test conducted for aggregate growth index fails to reject the linearity hypothesis and thus the linear model is used as my preferred specification. The results from the graphical analyses reveal that business plan score from judges is a predictor of aggregate growth, particularly for the full sample as implied by the dashed regression fit line in Figure 4.9 and for the trained sub-sample as represented in the solid stepper regression fit line in Figure 4.10.

This result is further confirmed by the full regression results presented in Table 4.6 as the estimated coefficients for score remains positive and statistically significant in all alternative specifications, both in the pooled sample (column 1 to 3) and trained sub-sample (column 7 to 9). Consistent with our observation from the near to horizontally slopped regression fit line (the dashed line) in Figure 4.10, however, score is not significantly correlated to aggregate growth in the non-trained sub-sample as shown in column 4 to 6 of Table 4.6.

The other related result which is worth presenting here as also a summary of all other outcomes presented in previous sections is about the question whether judges can differentiate less-promising businesses (business ideas) sorted to the bottom of the distribution or high growth-potential ones at the top of the distribution. I can answer this question from the relationship between actual aggregate growth index observed a year the prediction about their growth potential was made and score quintile. The result depicted in Figure 4.11 clearly shows that judges in the competition were successful to easily identify least performing applicants as implied by the aggregate growth of applicants in the first quintile and top performing ones (or gazelles) as indicated by the high level of aggregate growth achieved by the applicants in the fifth quintile. This result reaffirms that the scores are not only correlated with future business outcomes over the entire distribution on average, but also provide suggestive evidence that business plan can help identify gazelles.

Figure 4. 11 Aggregate growth index by score quintile (full sample)



In sum, the overall result for this aggregated measure of actual business performance unequivocally bolsters the finding that business plan scores from panel of experts can successfully predict future entrepreneurial success. Prediction of entrepreneurial success is better off when we combine scores from judges with entrepreneurial traits and other baseline covariates than using each separately, as indicated in the results presented so far and results reported in the appendix Table 4.A.3. This implies that scores from judges add to the predictive power of econometric models. In general, my finding supports most of the previous studies mainly Fafchamps & Woodruff (2017), Åstebro & Elhedhli (2006) and Scott et al. (2020) while we contradict with McKenzie & Sansone (2019).

Table 4. 6 Prediction of firm aggregated growth by panel of experts

	Full Sample			Non-trained			Trained		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Judges' Score	0.0276***	0.0205***	0.0170*	0.0063	0.0054	-0.0137	0.0295***	0.0212**	0.0188*
	(0.0077)	(0.0077)	(0.0086)	(0.0166)	(0.0173)	(0.0232)	(0.0087)	(0.0087)	(0.0098)
Existing firm		1.4397***	1.2292***		0.8560	1.1286*		1.5219***	1.2949***
		(0.2994)	(0.3136)		(0.6882)	(0.6720)		(0.3313)	(0.3489)
Agriculture sector			-0.0880			-0.4515			-0.0166
			(0.6352)			(1.3057)			(0.7351)
IT sector			-0.2888			-0.5563			-0.0967
			(0.5753)			(1.1985)			(0.6699)
Manufacturing sector			0.0192			-1.4367			0.1949
			(0.5654)			(1.2215)			(0.6446)
Retail sector			-0.5539			-0.8082			-0.4156
			(0.5680)			(1.2587)			(0.6510)
Female			-0.5358*			-0.2850			-0.5835
			(0.3156)			(0.6199)			(0.3657)
TVET or some college			1.0294*			2.8171**			0.6850
			(0.5368)			(1.2194)			(0.5985)
Undergraduate or graduate			0.1851			0.9851*			0.0830
			(0.3041)			(0.5281)			(0.3591)
Constant	0.0799	-0.3532**	-0.5914	-0.5250*	-0.7162**	-0.8770	0.2059	-0.2735	-0.6802
	(0.1321)	(0.1455)	(0.5999)	(0.2812)	(0.3153)	(1.4375)	(0.1495)	(0.1676)	(0.6781)
Regional FE	No	No	Yes	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	451	451	451	97	97	97	354	354	354
R-squared	0.0248	0.0794	0.117	0.00138	0.0224	0.154	0.0276	0.0878	0.133
Adjusted R-squared	0.0227	0.0753	0.0841	-0.00913	0.00163	-0.0153	0.0248	0.0826	0.0919

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is aggregate growth which is a sum of standardized values of the Inverse hyperbolic sine transformations of sales, profit and employment.

4.4.6. Heterogeneities in predicting success

In the previous sections, I have shown that prediction ability of experts varies by the types of outcomes being predicted, the relative position of applicants in the distribution of the score, and the availability of enough power in the estimation. In this section, explore two additional potential sources of heterogeneity for success in predicting entrepreneurial success have been explored and possible explanations for the variations have been provided. I will start with the types of competition (EDC Versus Bruh) and then move to business status of applicants at the baseline (Existing Versus New business). Each of these issues are discussed as in next sub-sections.

4.4.6.1. Results disaggregated by competition type (Bruh and EDC)

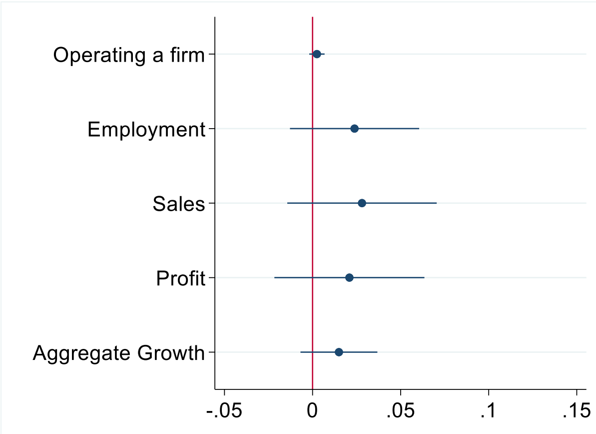
As described in chapter 2, the data of this study is pooled from two different business plan competitions- Bruh and EDC- and results presented so far in this chapter were for the pooled sample. In order to examine the effectiveness of judges in predicting future entrepreneurial success separately in the two competitions, I have estimated the models for each case separately. The estimation results of the 5 outcome variables considered in this study are presented in Appendix 4.3 from Table 4.A.4 to Table 4.A.8. Estimates reported in these tables about the parameter of interest, coefficients on score, and their respective 95% confidence intervals are summarized in Figure 4.13.

Panel A and panel B of Figure 4.13 depicts the estimates from Bruh data without and with additional covariates, respectively. In both specifications, the coefficient on score is not statistically different from zero for all outcomes. This implies that judges' score in Bruh business plan competition is uncorrelated with future business outcomes of applicants. On the other hand, the result presented in panel C and panel D of Figure 4.13 show a different result for EDC sub-sample. All the estimated coefficients of score without controlling for baseline covariates and fixed effect (Panel C) are positive and statistically different from zero for all outcomes. When we include baseline covariates and sector, panel, and regional fixed effects, score remains a significant predictor in all outcomes but sales and profit (Panel D). The correlation between score and outcomes do vary by types of cases (competitions) since judges were successful to predict entrepreneurial success in EDC but not in Bruh.

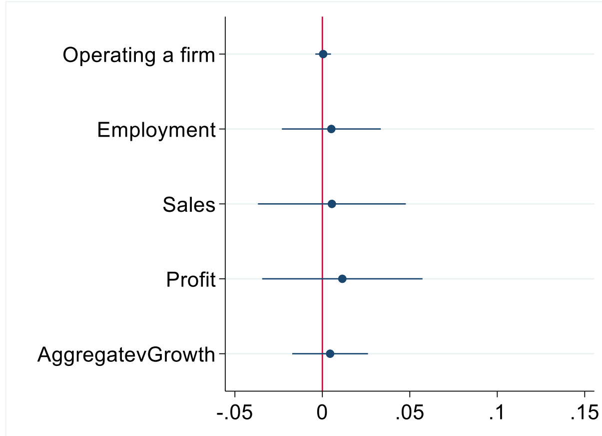
The variation of results between Bruh and EDC is not only in terms of the overall correlation between score and outcomes of interest but also in differentiating applicants in the bottom and top of the distribution, as shown in Appendix 4.4 Figure 4.A.6 to Figure 4.A.9.

Figure 4. 12 Estimated coefficients of score in comparison with Bruh versus EDC

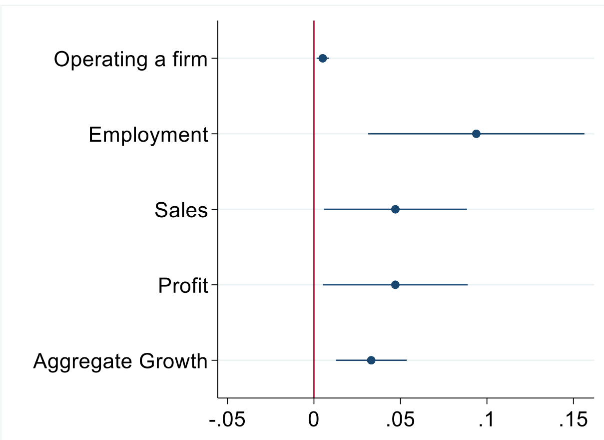
Panel A: Bruh Without covariates



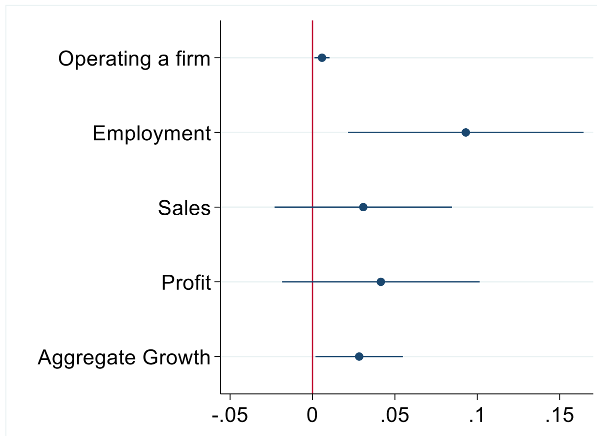
Panel B: Bruh with covariates



Panel C: EDC without covariates



Panel D: EDC with covariates



Notes: These graphs show the coefficient plots of score and their 95% confidence intervals (X-axis) for the five outcomes (Y-axis). Panel A and B are for estimates for Bruh and panel C and D are for EDC data separately.

The results of Bruh and EDC presented so far also serves to further bolster the argument that correlation between score and outcomes in the pooled sample is nor driven by the treatment (training in this case). All the disaggregated results by types of competition (Bruh versus EDC) show that score is a significant predictor of success in EDC sub-sample but not in Bruh. If score affects outcome of interest through training, we must observe a stronger relationship between score and training probability in EDC than that of Bruh. However, the result reported in the Appendix 4.3, Table 4.A.9, clearly shows the opposite. That means score is a significant predictor of training in Bruh but score is uncorrelated with all the outcomes in this sub-sample. In case of EDC, even if score is not a significant predictor of training probability, it is significantly correlated with outcome. These results suggest that the effect of score on outcomes is not through training.

Further, results of my previous chapter demonstrated that training do not a significantly affect business outcomes. Even in the presence of significant effects of training intervention, training does not necessarily change ‘subpar ideas into high-growth firms’ as pointed out by González-Uribe & Reyes (2021) which implies working with gazelles (i.e selection effect) also matters.

4.4.6.2. Possible explanations for heterogeneity in Bruh versus EDC results

Bruh and EDC business plan competitions are natural experiments of the same setting. Nonetheless, I found different results for these cases regarding the predictive ability of judges. Why Score predicts success in EDC but not in Bruh? Answering this question will have a paramount importance to explain the sources of heterogenous results of the previous literature beyond the variation in settings since we are dealing here with two different results occurred in the same setting. Similarly, the variation in statistical power cannot be the reason for this as both cases have comparable sample size. I tried to closely study the qualitative phenomena concerning the design and implementation of each competition and provided preliminary reasons behind the variation of the results between Bruh and EDC and presented as follows.

a) Clarity and depth of the evaluation criteria

One of the differences we witness in the implementation of the competitions that could cause the variation in results between Bruh and EDC is the evaluation criteria used in the first screening. While the first-round evaluation criteria given for the jury was broader and more aggregated with only 5 categories in case of Bruh, the criteria used by EDC were 5 major components further disaggregated into 20 specific cues (sub-components) that makes the total score, as shown in appendix 2.1. The more specific the criteria are the better the evaluation would be. Because if the evaluation criteria clearly included many dimensions with proper weights, it would be easier for experts to score the viability of businesses on various perspectives and then the average score is likely to be a true reflection of the potentials of contenders. Similarly in Canada where experts succeed to predict success, individual experts had 37 specific criteria to form the total score of 100% and the consistent use of those criteria across proposals led to a better prediction (Åstebro & Elhedhli, 2006).

b) Experience/ expertise

Another possible reason that could explain the variation in prediction success between Bruh and EDC is the prior experience in doing similar tasks. As a newly established government organization, Bruh was the first business plan competition for JCC as an organization and it did not have any experience in doing similar tasks before that. On the other hand, EDC has been operating in the market for near to a decade dealing with entrepreneurship training, business development support provider, and business counselor particularly for SMEs, to name the few. EDC had also a reasonable experience in organizing business plan competitions prior to this competition. For example, EDC had been organizing business plan competitions exclusively for established businesses and supporting the business idea competitions held in various universities in Ethiopia. Therefore, EDC has accumulated good organizational expertise which is likely to be helpful in managing the competition under evaluation.

In addition, its staff, who are mostly certified entrepreneurship trainers and did the first screening of applicants for this competition, are experienced in scoring business plans and judge viability of businesses in various scenarios. They are more acquainted with the business environment particularly for startups and small businesses. In contrast, renowned entrepreneurs or large business owners are usually assigned as members of juries in business plan competitions of small and young businesses. However, there is a risk that such big businesspersons may not understand the business environment that startups face. They could confuse the large business situations with the startup ones, which are two distinct features, in performing the evaluation. Success in entrepreneurship does not guarantee success in making accurate prediction as the latter one needs to be serious on paper works which requires proper extraction and utilization of relevant information from the business plans and application forms. It may be because of such differences that Kerr et al (2014b) found that venture capitalists, who are large investors and entrepreneurs, were unable to differentiate between successful and unsuccessful investment.

Nonetheless, experts dealing with startups every day will have a clear idea about their situation and likely make proper imagination during their evaluation. Given their ample experience in a similar context for several years, experts in EDC are also likely to consistently use the evaluation criteria and also make dynamic comparisons when they face new business ideas in their evaluation exercise, both of them are crucial to success in prediction as indicated in Åstebro & Elhedhli (2006) and Åstebro & Koehler (2007).

Further, expert of EDC who were responsible for scoring the business proposals in the first screening are not only experienced in doing similar tasks before but also more likely to get feedbacks about the accuracy of the judgments as they have been operating in their respective regions where they can easily witness to receive outcomes of their judgment. This practical situation allows the fulfilment of 2 of the 3 theoretical conditions stated by Kahneman & Klein (2009) which are required to make accurate prediction. Therefore, both individual and organizational expertise in making similar tasks (evaluation) matter.

c) Availability of conditions prone to bias in evaluation

Success in prediction partly stemmed from objective and impartial evaluation of applicants based on just the growth potentials of businesses. When the environment with in which the scoring is made is prone to bias judgment, judges would knowingly or unknowingly deviate from their objectivity. One potential source of bias that could lead an erroneous scoring of applicants we witness in these competitions is the variation in size of applicants.

In Bruh a jury had to screen 345 applications in short period of time since all applications were collected and evaluated centrally using one committee. Each member of the jury had to evaluate each applicant which implies that a person had to read 345 business plans including those excluded as illegible and score 277 eligible ones. This is a huge burden for the evaluators and more likely to compromise quality of evaluation. On the other hand, in EDC, competition was clustered by its regional offices (four centers). Each region had relatively small number of applicants and this situation allows experts to critically review the applications with a reasonable time to give a deserved scores for each contestant.

I had also a chance to see the business ideas of each applicant submitted to both competitions and I witnessed huge variations among applicants in the skill and completeness of presenting their ideas. Many applicants write the right and crucial information in the wrong place within the proposal. For instance, in the topic where they are asked to discuss the expected gains from the program, they present irrelevant ideas like their prior success, milestones, or sustainability of their business. While such data are crucial by themselves for the evaluation, evaluators cannot find them in their respective sections. If an evaluator gives more time and read all sections of a business plan before start scoring, she/he is more likely to better distinguish good and bad ideas no matter how the idea is presented. Otherwise, more promising businesses are likely to be left out in the screening if a busy evaluator only considers information presented in the relevant sections of the business plan format and give marks for a respective criterion. Therefore, assigning reasonable numbers of applicants per jury and give reasonable time to review each application is important.

Another factor worth considering in examining the evaluation environment is the immediate incentive for winner. In some business plan competitions winners are awarded with a big prize money, like 50,000 USD for each winner with a total budget of 34 million USD, being qualified in the fist-screening increase the probability of winning the grant by 20% in case of YouWin business plan competition in Nigeria as reported in McKenzie (2017) for example.[22] In such condition, some evaluators could knowingly bias the evaluation for some applicants in association with some group ties, if not personal, or other interest particularly if the identity of applicants are either fully or partly disclosed for the evaluators. If the environment allows such manipulation, we should not expect a significant correlation between score and future business outcomes.[23]

Zooming in the Bruh and EDC evaluation circumstance in this regard, we see some variation that could have partly contributed for the heterogeneity of the results. In case of Bruh, both applicants

[22] I calculated this probability from the ratio of total number of winner (1204) to total numbers of applicants who passed the first screening (6000).

[23] I cannot tell if the poor performance of judges in predicting entrepreneurial success in case of YouWin reported in McKenzie & Sansone (2019) is associated with this.

and judges were aware that qualification of the first screening (the score used for this study) increases the probability of winning 5000 USD as prize money to 28.6% for a show-up complier since 20 of the 70 applicants who are admitted to the bootcamp were granted to win the grant.[24] Here, financial incentive could be a good reason for the applicants to be in this competition and possibly the availability of bias in judgments cannot be ruled out.

In case of EDC, on other hand, only 6 applicants were selected as national winners and awarded 3000 to 5000 USD each, where qualifying for the next stage in the first screening will only lead to a 4.1% probability of winning the financial award. For EDC applicants, getting the business development services is the main incentive for applying to the competition in the first place as well as the main award for advancing to the next stage. Since free entrepreneurship training and business development supports are ubiquitous as confirmed in this study, there would not be any reason for experts to be deliberately bias in their assessment. Further, as both applicants and evaluators live in the same region, the same environment, and with a more homogeneous background, experts in EDC are likely to score business plans solely based on their viability.

d) Depth of information about the idea

Bruh and EDC also vary by the depth of information that the business plans contain about each applicant and their proposed businesses. The business plan format used by EDC allows to collect more detailed information about the entrepreneurs and their enterprises or proposed business ideas than that of Bruh. In addition, the fact that the EDC contest was conducted in respective regions with the evaluator also in the same geographical area enables the judge better undertand the context of the proposed business and easily guage its viability though reducing the problem

[24] If we consider applicants who did not comply to the bootcamp in the process of filling the quota of 70 applicants, this probability reduces to 19.4%, still large.

of information assymetry. My result suggests that the more information the judges have, the better the prediction accuracy would be. This is contrary to Zacharakis & Meyer (2000) who found that the more information were provided to the venture capitalists (that are evaluators), the less accurate were their predition on outcomes.

e) Other factors

There are other factors that could explain the weak correlation of score and outcomes in Bruh case. First, female applicants were given additional 3 percentage points since JCC wanted to promote gender equity where male applicants will earn zero at this criterion. Given this criterion, scores do not purely reflect the potential of businesses.

Second, there seems a mismatch between outcomes measured by researchers and evaluation criteria used by competition organizers. For instance, competition organizers look for innovativeness, creativity, social value adding, environmental sustainability, and other desirable features of the businesses, whereas researchers measure profit, sales, employment, or survival (which are private returns) regardless of whether the money was made from innovative, unique, socially responsible business or the customary or self-centered firms. A business that is scored low due to duplicating customary businesses or low benefit to society still can give higher return for the owner. This leads to a poor correlation between score and outcome.

Finally, my data shows that both applicants and their enterprises are younger in Bruh than in EDC as the former one followed a more strict and clearer criteria to exclude older applicants. Too young entrepreneurs are usually unstable and observed to be undecided whether continue business or schooling. That weakens the relationship between score and outcomes.

To sum up, the qualitative information I provided in this section about each competition could be taken as preliminary explanations for hetrogenous results found in this study as well as we witness in the literature. However, a more structured studies, ideally using experimental methods, is required to genrate adequate evidence on this area and my explanations could be used to build up from.

4.4.6.3. Prediction for existing Versus new firms

As stated before, my sample consists of both exising and new businesses at the time of application. I utilized this feature to test the hypothesis that prediction of outcome for new ideas or busineses is more difficult than that of exisitng ones at the point of evaluation. To this effect, I dissagregated the analyses by the business status of applicant (existing vs new). The estimated results for all the five outcome variables without controlling for other covariates are presented in Table 4.7. As shown in this table, the correlation between score and busiess outcomes are stronger for existing businesses than in their new counterparts. The result remains the same when baseline covariates and fixed effects are controlled for though the statistical significance goes away for some outcomes (Table 4.A.10 in the appendix). This finding confirms the notion that new ideas are difficult to evaluate (Arrow, 2012).

Table 4. 7 Prediction of success by score from panel of experts disaggregated by firm type (without controls)

	Owning a firm		Employment		Sales		Profit		Aggregate growth	
VARIABLES	Existing	New	Existing	New	Existing	New	Existing	New	Existing	New
Judges' Score	0.0049*	0.0025	0.1195**	0.0262	0.0505*	0.0209	0.0462*	0.0184	0.0352**	0.0136
	(0.0025)	(0.0018)	(0.0509)	(0.0169)	(0.0297)	(0.0181)	(0.0274)	(0.0185)	(0.0153)	(0.0087)
Constant	0.5875***	0.3605***	5.6391***	2.4139***	5.6923***	3.2319***	4.6246***	2.3268***	1.0628***	-0.3829***
	(0.0425)	(0.0279)	(0.7662)	(0.3822)	(0.4994)	(0.2979)	(0.4982)	(0.2940)	(0.2606)	(0.1473)
Observations	133	323	133	323	130	322	130	321	130	321
R-squared	0.0268	0.00609	0.0445	0.00388	0.0211	0.00401	0.0178	0.00314	0.0368	0.00672
Adjusted R-squared	0.0193	0.00299	0.0372	0.000776	0.0134	0.000895	0.0101	1.06e-05	0.0293	0.00361

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variables are as defined in previous tables

4.4.7. Robustness checks

So far, I provided evidence about the positive correlation between scores from the business plan competition judges and future entrepreneurial success. I have also demonstrated that the result is not driven by the main intervention of the program, that is training. Nonetheless, the possible effect of the grant, another intervention of the program, was ignored since I thought that the number of participants who won and received the cash grant are too small to affect the overall result. However, it is worth testing if this expectation is true. That is what the robustness check exercise of this section does.

The disaggregated estimation results presented in appendix Table 4.A.4 to Table 4.A.8, which included the grant winners, consistently show that for Bruh sub-sample the correlation is negligible even without turning off the possible effect of the grant. Therefore, the inclusion of grant winners in the sample will not be an issue in case of Bruh. For EDC, however, there is a strong positive correlation between score and entrepreneurial success, and the same results found for the full sample is also driven by the EDC result. Hence, for this sub-sample, it is reasonable to suspect that the positive correlation may have been caused by the grant effect. And a positive correlation driven by the grant weakens the argument that business plan competition judges can help predict future business success.

Therefore, I excluded the grant winners and re-estimated all the models for EDC sub-sample to check if excluding the grant winners will fade away the positive correlation, and the results are reported in Appendix Table A.4.11 and Table A.4.12. The results here again reaffirm that score is a predictor of entrepreneurial success even after the removing the possible effects of the grant. Comparing the results with and without grant winners (last three columns of Table 4.A.4 to 4.A.8 Versus Table A.4.11 to Table A.4.12) clearly shows that the estimated coefficients are very comparable. Similarly, statistical significance of the parameter of interest has been preserved in all the cases except in the second specification of profit model where we have some attenuations as the 10% of significance we had with grant winners goes away when grant winners are excluded.

In sum, this robustness check exercise suggests that the argument of the small number of grant winners will lead to a negligible effect on overall results of the study is valid and thus the effect of score and business outcomes is direct, not through the intervention of the program.

4.5. Concluding remarks

This paper examines the link between the prediction of growth potentials for business plan competition applicants in Ethiopia measured by business plan scores from panel of judges and actual entrepreneurial success observed a year after the predictions were made. The probability of operating a firm after a year, level of employment, sales, profit, and aggregate growth were used to measure entrepreneurial success of applicants of Bruh and EDC startups' business plan competitions.

I found that score from judges is a significant predictor of future entrepreneurial success in general. The correlation is much stronger in outcomes that are less susceptible to measurement error. I provided compelling evidence in this study that judges were particularly effective to differentiate both high-growth potential businesses at the top of distribution and less-promising ones at the bottom of the distribution. This finding suggests that business plan competition is an effective policy option that help identify gazelles.

This study confirmed the previous literature in the fact that judges' score and other baseline covariates explain only small fraction of the variation in business outcomes as we did not find the adjusted R-squared that exceed from 13% in all the models we have estimated throughout the chapter. The finding also revealed that success in prediction varies by the nature of the outcome we are trying to predict, the availability of enough statistical power, the nature of applicants being evaluated (existing versus new; extreme ideas versus medium), and the way that a business plan competition is designed and implemented (as we learned from the comparison of Bruh versus EDC). Particularly the last point is associated with the scoring environment.

In a business plan competition where the scoring environment is enabling to make objective and serious evaluation, experts can predict entrepreneurial success. One source of success in prediction comes from having the right experts to score the business plans. Right experts in this case are those professionals who are knowledgeable about the unique (as opposed to general) entrepreneurship landscape or business environment for the target groups being evaluated, experienced in making similar evaluation, adhere to the given criteria, and consistently use the given criteria in making the judgements. Another source of success in making good prediction is associated with the (un)availability of situation that compromise the unbiasedness of the evaluation. The workload of the expert (number of applicants and time given to make the evaluation), the availability of prize money with a good chance for each applicant to win it, and the depth of information available to the judges at the time of evaluation are possibly among the factors determining the accuracy of score which, in turn, derives the accuracy of predicting entrepreneurial success.

The bottom line of the finding of this chapter is a properly managed business plan competition can help predict future entrepreneurial success and potentially identify gazelles. This study confirms that the business plan competitions in Ethiopia generally performed well at least in achieving one (the first) of their dual purposes they intended to do when designed: identify potential gazelles and intervene to overcome their constraints.

CHAPTER 5

5. Conclusions and Policy Implications

5.1. Conclusions

This dissertation was designed to study new approaches of entrepreneurship development programs in developing countries by pooling data from two business plan competitions in Ethiopia which target young entrepreneurs with innovative businesses (ideas). In line with the intention of any business plan competition, the study evaluates the program from two perspectives: its impact in nurturing entrepreneurship through its direct intervention, and its effectiveness in targeting the right entrepreneurs that the program intended to address.

About a year after the opening of the program, I traced about 500 (potential) entrepreneurs who applied for the program and conducted a carefully designed survey to address both issues. I employed a fuzzy regression discontinuity design to identify the short-term impact of the training intervention on business entry and expansion. However, the business performance of the training beneficiaries of the program was not better than the rejected applicants. One possible reason for this is because rejected applicants were able to get similar trainings in other programs and thus they cannot be considered as pure controls. Potentially, this could be one of the main reasons behind the negligible impacts of entrepreneurship training program widely reported by many studies around the world.

Though this study is not able to infer about the effectiveness of the program, the business plan scores from judges were predictive of success and they were able to differentiate high and low potential businesses. This implies that business plan competition is a successful policy option to identify the types of enterprises based on their growth potential, which is a key for proper targeting as well as for design and implementation of tailored policies. However, this success is not automatically guaranteed for any business plan competition. The way a competition is designed and implemented matters for experts to succeed in predicting future entrepreneurial success.

5.2. Policy Implications

Important policy implication can be drawn from the findings of this study which could be utilized by policy makers, program implementers including business plan competition organizers, and researchers.

For policy makers, it is worth considering business plan competition as one of the innovative approaches for fostering entrepreneurship. It helps at least to differentiate promising businesses from the mass and could facilitate financing by serving as a bridge between interested investors and constrained gazelles.

The implication of supporting business plan competitions and award winners go beyond the private return for the recipient firms. It is due to the huge social benefit including formation of sustainable businesses, productive employment, enhancing competitiveness and creativity, and improve resource allocations that public policy should target businesses with growth potentials. As stated in Shane (2009) "getting economic growth and jobs creation from entrepreneurs is not a numbers game. It is about encouraging high quality, high growth companies to be founded." However, it should also be noted that at low level of economic development like Ethiopia, having large number of typical start-ups with a negligible role beyond providing temporary employment for owners is inevitable. In such context, as supporting these enterprises could be totally unavoidable from poverty reduction or political perspectives, it is important for policy makers to have a clear understanding of the expected outcomes of such policies and make the right balance between policies favoring gazelles versus policies supporting the survivalists.

For business plan competition organizers, there are some lessons obtained from this study. The first source of success for startups' intervention stems from the ability to properly hunt talents. I witnessed considerable variations among applicants of the business plan competitions in communicating their ideas. Some applicants misplace important ideas within their business plan, others unnecessarily try to write it in English but fail to communicate clearly, and some others are not good in organizing their ideas and make it marketable. This may mask the real potential of applicants from being detected by the evaluators. I suggest two solutions for this problem. First, it could be a good idea to have a briefing session on how to present businesses ideas just before or during the application period. If this is not feasible, a brief video explaining the formats and what to present in each heading could be prepared and released with the call for application in various media. Second, it would be important for the evaluators to allow more time to check every content of the plans and judge the potential of the business regardless of how the ideas are presented.

The other crucial task of the competition is the proper scoring of the business plan. Having the right experts, enabling scoring environment that promotes objective evaluation, allowing sufficient time and minimizing burden of each member of the jury, and developing more disaggregated criteria to properly assess the various aspects of the business are among the steps one can take to improve the screening accuracy.

As I learned from the motivation of applicants, unlocking the potentials of startups through overcoming their skill constraints requires going beyond entrepreneurship training. Many people also look for hard skill and technical trainings whereas the trainings given through the process of the competitions are usually limited to business trainings and development of soft skills. Probably, organizing business plan competition by sectors or themes (like Energy, IT or software, Hardware, services, etc) and including technical supports could be helpful.

In Ethiopia, I witnessed that many training programs run in the market. In this situation, new programs like Bruh and EDC business plan competitions may not be able to attract more competent participants, which makes success difficult with small push. To attract talented and high-growth potential applicants for the program to succeed, it is essential to improve the program design in a way that creates incentive for competent applicants to enroll. Providing unique interventions (more relevant trainings which are not easily available in the market), using more innovative delivery methods, and increasing the numbers of winners and amount of the cash grant are some of the actions that help attract potential gazelles in the competition. In a situation where small amount of grant is a reward for participating and winning a long waiting business plan competition, the program will end up attracting typical startups with low potential to growth. Note that such programs only help innovative ideas to flourish, it is not a magic bullet to change bad ideas to best ones.

The main message of this study for researchers is associated with ensuring the validity of counterfactual while conducting impact studies. My findings about substantial take-up of substitute treatments poses a question on many impact studies. The cleanness of the control group for any design cannot be confirmed unless data prove that they stayed away from substitute programs. Therefore, it is important to consider it as an important task while designing follow-up surveys.

5.3. Limitations and future research

Though attempt is made to carefully design this study and thoroughly analyze the data throughout the dissertation, it is also subject to some limitations. First, alike almost all the studies I reviewed in this area, this study also suffers from problem of small sample size. It is important to conduct more study using larger sample size whenever there are opportunities to get such natural experiments with large numbers of applicants. Second, this study provides preliminary explanations regarding the drivers of prediction accuracy in the context of entrepreneurship competition. Nevertheless, these explanations are based on qualitative information. Future researchers could capitalize on this by testing these factors using randomized control trials.

References

Abebe, G., Gebreeyesus, M., Assefa, B., & Getahun, T. (2017). *Who are growth-oriented entrepreneurs ? Evidence from a large scale survey of micro and small enterprises*.

Arrow, K. J. (2012). The Economics of Inventive Activity over Fifty Years. In *Lerner J, Stern S, eds.The Rate and Direction of Inventive Activity Revisited* (Issue March, pp. 43–48). https://doi.org/10.7208/chicago/9780226473062.003.0004

Åstebro, T., & Elhedhli, S. (2006). The effectiveness of simple decision heuristics: Forecasting commercial success for early-stage ventures. *Management Science*, *52*(3), 395–409. https://doi.org/10.1287/mnsc.1050.0468

Åstebro, T., & Koehler, D. (2007). Calibration Accuracy of a Judgmental Process that Predicts the Commercial Success of New Product Ideas. *Journal of Behavioral Decision Making*, *20*(February), 381–403. https://doi.org/10.1002/bdm

Ayyagari, M., Demirguc-Kunt, A., & Maksimovi, V. (2011). Small vs. Young Firms across the World: Contribution to Employment, Job Creation, and Growth. In *World Bank Policy Research Working Paper 5631*. https://www.researchgate.net/publication/228319883_Small_vs_Young_Firms_Across_the_World_Contribution_to_Employment_Job_Creation_and_Growth

Beck, T., & Demirguc-Kunt, A. (2006). Small and medium-size enterprises: Access to finance as a growth constraint. *Journal of Banking & Finance*, *30*(11), 2931–2943. https://doi.org/10.1016/J.JBANKFIN.2006.05.009

Berner, E., Gomez, G., & Knorringa, P. (2012). "Helping a large number of people become a little less poor": The logic of survival entrepreneurs. *European Journal of Development Research*, *24*(3), 382–396. https://doi.org/10.1057/ejdr.2011.61

Bigsten, A., & Gebreeyesus, M. (2007). The small, the young, and the productive: Determinants of manufacturing firm growth in Ethiopia. *Economic Development and Cultural Change*, *55*(4), 813–840. https://doi.org/10.1086/516767

Bigsten, A., & Söderbom, M. (2006). What have we learned from a decade of manufacturing enterprise surveys in Africa? *World Bank Research Observer*, *21*(2), 241–265. https://doi.org/10.1093/wbro/lkl003

Blattman, C., & Dercon, S. (2018). The Impacts of Industrial and Entrepreneurial Work On Income and Health: Experimental Evidence from Ethiopia. *American Economic Journal: Applied Economics*, *10*(3), 1–38.

Blattman, C., Dercon, S., & Franklin, S. (2022). Impacts of industrial and entrepreneurial jobs on youth: 5-year experimental evidence on factory job offers and cash grants in Ethiopia. *Journal of Development Economics*, *156*, 1–15. https://doi.org/10.1016/j.jdeveco.2021.102807

Blattman, C., Fiala, N., & Martinez, S. (2014). GENERATING SKILLED SELF-EMPLOYMENT IN DEVELOPING COUNTRIES : EXPERIMENTAL EVIDENCE FROM UGANDA. *The Quarterly Journal of Economics*, *129*(2), 697–752. https://doi.org/10.1093/qje/qjt057.Advance

Bone, J., Gonzalez-Uribe, J., Haley, C., & Lahr, H. (2019). The impact of business accelerators and incubators in the UK. In *BEIS Research Paper Number 2019/009*. https://www.gov.uk/government/organisations/department-for-business-energy-and-industrial-strategy

Brinckmann, J., Grichnik, D., & Kapsa, D. (2010). Should entrepreneurs plan or just storm the castle? A meta-analysis on contextual factors impacting the business planning–performance relationship in small firms. *Journal of Business Venturing*, *25*(1), 24–40. https://doi.org/10.1016/J.JBUSVENT.2008.10.007

Bruhn, M., Karlan, D., & Schoar, A. (2010). What Capital is Missing in Developing Countries ? *The American Economic Review*, *100*(2), 629–633.

Buera, F. J., Kaboski, J. P., & Shin, Y. (2020). Taking stock of the evidence on microfinancial interventions. *Federal Reserve Bank of St. Louis Review*, *102*(2), 173–202. https://doi.org/10.20955/r.102.173-202

Carrillo, P., Pomeranz, D., & Singhal, M. (2017). Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement. *American Economic Journal: Applied Economics, 9*(2), 144–164. https://doi.org/10.1257/APP.20140495

Cattaneo, M. D., Jansson, M., & Ma, X. (2020). Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association*, *115*(531), 1449–1455. https://doi.org/10.1080/01621459.2019.1635480

Cattaneo, M. D., Jansson, M., & Ma, X. (2021a). Local regression distribution estimators. *Journal of Econometrics*, *xxxx*, 1–18. https://doi.org/10.1016/j.jeconom.2021.01.006

Cattaneo, M. D., Jansson, M., & Ma, X. (2021b). Local regression distribution estimators. *Journal of Econometrics*. https://doi.org/10.1016/j.jeconom.2021.01.006

Cattaneo, M. D., Jansson, M., & Ma, X. (2022). lpdensity: Local Polynomial Density Estimation and Inference. *Journal of Statistical Software*, *101*(2), 1–25. https://doi.org/10.18637/jss.v101.i02

Chiang, I.-C. A., Jhangiani, R. S., & Price, P. C. (2015). Research Methods in Psychology Research Methods in Psychology. Canadian Edition. In *Oxford University Press*. http://www.saylor.org/site/textbooks/Research Methods in Psychology.pdf

Cotter, J. (1996). Distinguishing between poverty alleviation and business growth. *Small Enterprise Development*, *7*(2), 49–52. https://doi.org/10.3362/0957-1329.1996.016

CSA (Central Statistical Agency, E. (2021). *2021 LABOUR FORCE AND MIGRATION*

SURVEY (Issue August). https://www.statsethiopia.gov.et/wp-content/uploads/2021/08/Final-2021-LABOUR-FORCE-AND-MIGRATION-SURVEY_Key-finding-Report-.17AUG2021.pdf

de Mel, S., McKenzie, D. J., & Woodruff, C. (2009). Measuring microenterprise profits: Must we ask how the sausage is made? *Journal of Development Economics*, *88*(1), 19–31. https://doi.org/10.1016/j.jdeveco.2008.01.007

Decker, R., Haltiwanger, J., Jarmin, R., & Miranda, J. (2014). The role of entrepreneurship in us job creation and economic dynamism. *Journal of Economic Perspectives*, *28*(3), 3–24. https://doi.org/10.1257/jep.28.3.3

Fafchamps, M., & Quinn, S. (2017). Aspire. *The Journal of Development Studies*, *53*(10), 1615–1633.

Fafchamps, M., & Woodruff, C. (2017). Identifying gazelles: Expert panels vs. surveys as a means to identify firms with rapid growth potential. *World Bank Economic Review*, *31*(3), 670–686. https://doi.org/10.1093/wber/lhw026

Fritsch, M. (2008). How does new business formation affect regional development? Introduction to the special issue. *Small Business Economics*, *30*(1), 1–14. https://doi.org/10.1007/s11187-007-9057-y

Gebreeyesus, M., Ambachew, A., Getahun, T., Assefa, B., Abebe, G., Hassen, S., & Medhin, H. (2018). *Main Features of Micro and Small Manufacturing Enterprises in Ethiopia*. EDRI.

González-Uribe, J., & Reyes, S. (2021). Identifying and boosting “Gazelles”: Evidence from business accelerators. *Journal of Financial Economics*, *139*(1), 260–287. https://doi.org/10.1016/j.jfineco.2020.07.012

Gries, T., & Naudé, W. (2010). Entrepreneurship and structural economic transformation. *Small Business Economics*, *34*(1), 13–29. https://doi.org/10.1007/s11187-009-9192-8

Grimm, M., Knorringa, P., & Lay, J. (2012). Constrained Gazelles: High Potentials in West Africa’s Informal Economy. *World Development*, *40*(7), 1352–1368. https://doi.org/10.1016/j.worlddev.2012.03.009

Hall, R. ., & Woodward, S. (2010). The Burden of the Nondiversifiable Risk of Entrepreneurship. *American Economic Review*, *100*(3), 1163–1194.

Hallberg, K. (2000). A Market-Oriented Strategy for Small and Medium Scale Enterprises. In *A Market-Oriented Strategy for Small and Medium Scale Enterprises*. The World Bank. https://doi.org/10.1596/0-8213-4727-6

Henrekson, M., & Johansson, Æ. D. (2010). Gazelles as job creators : a survey and interpretation of the evidence. *Small Business Economics*, *35*, 227–244. https://doi.org/10.1007/s11187-009-9172-z

Hessels, J., & Naudé, W. (2019). the Intersection of the Fields of Entrepreneurship and Development Economics: a Review Towards a New View. *Journal of Economic Surveys*, *33*(2), 389–403. https://doi.org/10.1111/joes.12286

Higuchi, Y., Mhede, E. P., & Sonobe, T. (2019). Short- and medium-run impacts of management training : An experiment in Tanzania. *World Development*, *114*, 220–236. https://doi.org/10.1016/j.worlddev.2018.10.002

Higuchi, Y., Nam, V. H., & Sonobe, T. (2015). Sustained impacts of Kaizen training. *Journal of Economic Behavior and Organization*, *120*, 189–206. https://doi.org/10.1016/j.jebo.2015.10.009

House, W. J. . (1984). Nairobi ' s Informal Sector : Dynamic Entrepreneurs or Surplus Labor? *Economic Development and Cultural Change*, *32*(2), 277–302.

Kahneman, D., & Klein, G. (2009). Conditions for Intuitive Expertise: A Failure to Disagree. *American Psychologist*, *64*(6), 515–526. https://doi.org/10.1037/a0016755

Kerr, W. R., Lerner, J., & Schoar, A. (2014). The consequences of entrepreneurial finance: Evidence from angel financings. *Review of Financial Studies*, *27*(1), 20–55. https://doi.org/10.1093/rfs/hhr098

Kerr, W. R., Nanda, R., & Rhodes-kropf, M. (2014). Entrepreneurship as Experimentation. *The Journal of Economic Perspectives*, *28*(3), 25–48.

Kirkwood, J. (2009). Is a lack of self-confidence hindering women entrepreneurs? *International Journal of Gender and Entrepreneurship*, *1*(2), 118–133. https://doi.org/10.1108/17566260910969670/FULL/XML

Klinger, B., & Schündeln, M. (2011). Can Entrepreneurial Activity be Taught? Quasi-Experimental Evidence from Central America. *World Development*, *39*(9), 1592–1610. https://doi.org/10.1016/j.worlddev.2011.04.021

Kumler, T., & Verhoogen, E. (2020). Enlisting Employees in Improving Payroll Tax Compliance: Evidence from Mexico. *The Review of Economics and Statistics*, *102*(5), 881–896. https://doi.org/10.1162/REST_A_00907

Lall, S. A., Chen, L. W., & Roberts, P. W. (2020). Are we accelerating equity investment into impact-oriented ventures? *World Development*, *131*. https://doi.org/10.1016/j.worlddev.2020.104952

Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics*, *142*(2), 675–697. https://doi.org/10.1016/j.jeconom.2007.05.004

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*, 281–355. https://doi.org/10.1210/endo-127-4-1934

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design:

A density test. *Journal of Econometrics*, *142*(2), 698–714. https://doi.org/10.1016/j.jeconom.2007.05.005

McKenzie, D. (2017). Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition. *American Economic Review*, *107*(8), 2278–2307. https://doi.org/10.1257/aer.20151404

McKenzie, D., & Sansone, D. (2019). Predicting entrepreneurial success is hard: Evidence from a business plan competition in Nigeria. *Journal of Development Economics*, *141*(July). https://doi.org/10.1016/j.jdeveco.2019.07.002

McKenzie, D., & Woodruff, C. (2014). What are we learning from business training and entrepreneurship evaluations around the developing world? *World Bank Research Observer*, *29*(1), 48–82. https://doi.org/10.1093/wbro/lkt007

Mckenzie, D., Woodruff, C., Bjorvatn, K., Bruhn, M., Cai, J., Gonzalez-uribe, J., Quinn, S., Sonobe, T., & Valdivia, M. (2020). Training Entrepreneurs. *VoxDevLit*, *1*(1), 1–23.

Mead, D. C., & Liedholm, C. (1998). The dynamics of micro and small enterprises in developing countries. *World Development*, *26*(1), 61–74. https://doi.org/10.1016/S0305-750X(97)10010-9

Mottaleb, K. A., & Sonobe, T. (2012). The role of human capital in firm performance : evidence from the engineering industry in Bangladesh Khondoker Abdul Mottaleb * Tetsushi Sonobe. *Int. J. Entrepreneurship and Innovation Management*, *16*(3/4), 245–262.

Nanda, R. (2016). Financing high-potential entrepreneurship. In *IZA World of Labor* (Issue April). https://doi.org/10.15185/izawol.252

NBE. (2022). *2020-21 Annual Report*. https://nbebank.com/wp-content/uploads/pdf/annualbulletin/Annual Report 2020-2021/2020-21 Annual Report.pdf

Nichter, S., & Goldmark, L. (2009). Small Firm Growth in Developing Countries. *World Development*, *37*(9), 1453–1464. https://doi.org/10.1016/j.worlddev.2009.01.013

Noseleit, F. (2013). Entrepreneurship, structural change, and economic growth. *Journal of Evolutionary Economics*, *23*(4), 735–766. https://doi.org/10.1007/s00191-012-0291-3

OECD. (2004). *PROMOTING ENTREPRENEURSHIP AND INNOVATIVE SMEs IN A GLOBAL ECONOMY: TOWARDS A MORE RESPONSIBLE AND INCLUSIVE GLOBALISATION* . www.copyright.com.

Ohno, I., Amatsu, K., & Hosono, A. (2022). *Policy Learning For Industrial Development And The Role Of Development Cooperation* (Vol. 1, Issue February). https://www.jica.go.jp/jica-ri/ja/publication/booksandreports/tfpeil0000002pxe-att/policy_20220304.pdf

Pomeranz, D. (2015). No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax. *American Economic Review*, *105*(8), 2539–2569.

https://doi.org/10.1257/AER.20130393

Scott, E. L., Shu, P., & Lubynsky, R. (2015). Are “Better” Ideas More Likely to Succeed? An Empirical Analysis of Startup Evaluation. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2638367

Scott, E. L., Shu, P., & Lubynsky, R. M. (2020). Entrepreneurial uncertainty and expert evaluation: An empirical analysis. *Management Science*, *66*(3), 1278–1299. https://doi.org/10.1287/mnsc.2018.3244

Shane, S. (2009). Why encouraging more people to become entrepreneurs is bad public policy. *Small Business Economics*, *33*(2), 141–149. https://doi.org/10.1007/s11187-009-9215-5

Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational Behavior and Human Decision Processes*, *53*(2), 252–266. https://doi.org/10.1016/0749-5978(92)90064-E

Sims, M. A., & O’Regan, N. (2006). In search of gazelles using a research DNA model. *Technovation*, *26*(8), 943–954. https://doi.org/10.1016/j.technovation.2005.07.001

Sonobe, T., & Otsuka, K. (2006). Cluster-based industrial development: An East Asian model. In *Cluster-Based Industrial Development: An East Asian Model*. https://doi.org/10.1057/9780230596061

Sonobe, T., & Otsuka, K. (2011). Cluster-Based Industrial Development: A Comparative Study of Asia and Africa. In *Cluster-Based Industrial Development: A Comparative Study of Asia and Africa*. https://doi.org/10.1057/9780230295124

Vandenberg, P., Chantapacdepong, P., & Yoshino, N. (2016). SMEs in Developing Asia: New Approaches to Overcoming Market Failures. In *Asian Development Bank Institute*. https://www.adb.org/sites/default/files/publication/214476/adbi-smes-developing-asia.pdf%0Ahttps://www.adb.org/sites/default/files/publication/214476/adbi-smes-developing-asia.pdf#page=41

World Bank. (n.d.). *Ethiopia Overview: Development news, research, data | World Bank*. Retrieved June 30, 2022, from https://www.worldbank.org/en/country/ethiopia/overview

World Bank. (2008). *Finance for All? Policies and pitfalls in expanding access*.

World Bank. (2015). SME Finance in Ethiopia : Addressing the Missing Middle Challenge. In *World Bank Group*.

Zacharakis, A. L., & Meyer, G. D. (2000). The potential of actuarial decision models: Can they improve the venture capital investment decision? *Journal of Business Venturing*, *15*(4), 323–346. https://doi.org/10.1016/s0883-9026(98)00016-0

APPENDICES

Appendix 2. Additional materials for chapter 2

Appendix 2.1. Evaluation criteria used to screen applicants

Appendix 2.1.A.1: Criteria used in the final pitch competition as well as the first screening in Bruh

- Innovative aspect of the proposed solution/product (25 point)
- Sustainability & Scalability (25 points)
- Business Feasibility & Traction (30 points)
- Professionalism and Presentation of Business (5 points)
- Job creation potential and inclusiveness (15 points)

NB: The average score ranges from 58 to 92.5% and the top to won the financial award with the exogenous cut off point=79

Appendix 2.1.A.2: criteria used in the second screening for Bruh

- Innovative Aspect (15)
- Market Potential (15)
- Financial Feasibility (10)
- Technical Feasibility (10)
- Marketing and Sales (10)
- Sustainability and Scalability (10)
- Social and Economic (10)
- Financial and HR needed (10)
- Understanding of the Business (5)
- Team Excellence (5)

Appendix 2.1.A.3: criteria used in the third screening for Bruh

- Understand the ideas (10)
- Market mix (20)
- Target market identification (10)
- Macro Environment (10)
- Meso Environment (10)
- Microenvironment (10)
- Profitability (15)
- Economic Impact (15)

Appendix 2.1.A.5 EDC first-screening evaluation criteria

The idea

- Originality/innovativeness of the idea
- How well is the idea described? Is it feasible?
- Is the social aspect clear
- Is the idea relevant in social context? Does it address a social challenge?
- Clarity of vision and goal

Scalability

- Is the idea scalable in terms of growth and impact? Idea for incubation?
- Will the solution be implemented and has growth potential?
- Does it have a clearly defined long term plan of activities?
- Are required activities and resources clearly described?

Impact research

- Gained proof of the need for proposed solution
- Done proper research to testify the impact of their idea
- Conducted research in their community for th need for the solution, pricing , customer, competition, etc,? used the research outputs to change their business strategies?
- Potential risks are determined and strategies are drawn

The types of business, skill, and expertise

- Is has clearly considered the form of the company? Or forms of operation?
- Has relevant skill/industry knowledge to implement the idea? Ideal team composition?
- Passionate and committed to meet the objectives, vision and goal?

Finance

- Has the business clearly indicated its source of income?
- Innovative in their way of achieving revenue? Are revenue strategies feasible?
- Has the team clearly indicated financial projections (cost/revenue/profit/loss)
- Has a clear capita projection?

Appendix 2.1.A.6. EDC Second round screening criteria

Incubation Program 2020/21 - Scoring Sheet **Presenter Code:**	**Very Poor (1-3)**	**Average (4-6)**	**Very Good (7-10)**
THE IDEA			
• ***Innovativeness of the idea? Does it address a real social challenge?***			
• *How well is the idea described? Is it feasible?*			
• ***Is the business model clear?***			
• *Is the product/service offering clear?*			
SCALABILITY			
• *Is the idea scalable in terms of growth and impact? Ideal for incubation?*			
• *Does it have a clear defined vision, long-term plan of activities?*			
• *Are required activities and resources clearly stipulated?*			
• *How well on track are their growth metrics?*			
• ***Has relevant skills/industry knowledge to implement the idea? Ideal team composition?***			
MARKET ANALYSIS			
• *Gained proof of the need for the proposed solution?*			
• *Knowledgeable about their market, its size and competitors?*			
• *Conducted market research and used it in their business strategy?*			
• *The plan for overcoming competition is viable*			
• *Potential risks are determined and strategies are drawn?*			
• *Passionate & committed to meet the objectives, vision and goals?*			
FINANCIAL EVALUATION			
• *Has the business clearly indicated its source of income?*			
• *Innovative in their ways of achieving revenue? Are revenue strategies feasible?*			
• *Has the team clearly indicated financial projections (cost/ revenue/ profit/ loss)*			
• *How balanced are the finances?*			
• *Has a clear capital projection?*			
ELEVATOR PITCH			
• *Pitch was clear and comprehensive*			
• *Time management*			
• *Questions were answered clearly*			

Appendix 2.2: Timeline and steps of the project

Figure 2.A.1: Timeline and key procedures of the program as a whole

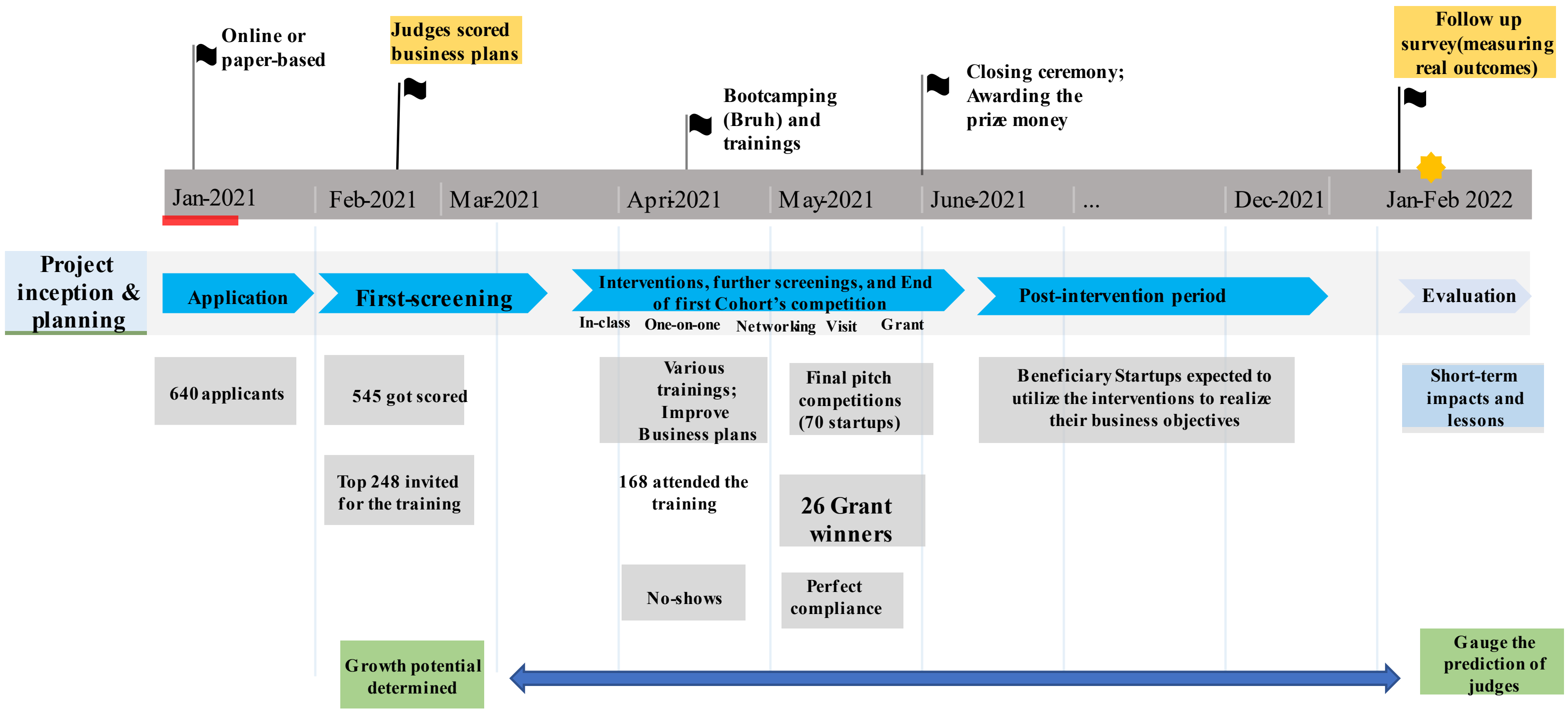


Table 2.A.1: Summary of the first version of the *Bruh* Business plan competition, 2021

S.No.	Activity	Participants	Timeline	Remarks
1.	Preparation (documents, partners, agreements, etc)	JCC and its partners organizations	Oct-Dec, 2020	
2.	Registration of applicants	Any individual aged 15-29 Years old or any business <2 years	January 1, 2021- Februaru 7, 2021	Advertised in various media
3.	Pre-screening	345 applicants and JCC staff		
4.	First round screening	277 applicants and 7 technical committee members		70 businesses selected
5.	Bootcamp entry and orientation about their stay	60 business represented by 112 members and facilitators	March 22/2021	There are unattended members within the participating businesses, the rest did not show up because of personal reasons and merger
6.	Experience sharing forums	60 business represented by 112 members, one invited guest, and facilitators	March 23/2021	Trainees allocated into two training rooms, 30 each.
7.	Training on "entrepreneurship competency"	60 business represented by 112 members, 8 certified trainers, and facilitators	March 24-27/2021	Fun games included
8.	Training on "Holistic Business idea development"	60 business represented by 112 members, 8 certified trainers, and facilitators	March 29- April 3/2021	Game included
9.	Second round screening (1st on-training screening within the bootcamp)	60 business represented by 112 members; judges composed of the trainers	April 3/2021	Based on 15 minutes presentation for each team and scored based on 10 criteria
10.	Elimination of unsuccessful team	Least 20 teams consisting of 39 members left the bootcamp	April 4/2021	Top 40 businesses continued
11.	Training on "visual prototype and product development"	40 businesses represented by 73 members, 8 trainers, facilitators	April 5-8/2021	Games included; prototype presentation by invited guest; industry visit conducted
12.	Training on Market research and unique selling proposition (USP)	40 businesses represented by 73 members, 8 trainers, facilitators	April 9-12/2021	Games included
13.	Third round screening (2nd on-training screening within the bootcamp)	40 businesses represented by 73 members and Judges composed of the trainers	April 13/2021	10 minutes presentation for each team; the least 5 businesses consisting of 11 members have been eliminated; 35 projects with 62 members passed.
14.	Training on "Legal Business Setup"	35 businesses with 62 members, trainers and facilitators	April 14-15/2021	
15.	Training on Business plan preparation	35 businesses with 62 members, trainers and facilitators	April 16-19/2021	
16.	Preparation for the final pitch competition	35 businesses and their members	April 20-21/2021	
17.	Final pitch competition	35 businesses and their members; 4 external judges	April 22-23/2021	Conducted in Kuriftu resort at Bishoftu; 20 minutes presentation for each team: top 20 selected for the financial award (5000 USD each)
18.	The closing and award ceremony	20 winner businesses, invited guests, JCC staff and leaders, media	May 27/2021	Winners handed 200,000 ETB (5000 USD cheque)

Appendix 2.3. Survey Questionnaire

The National Graduate Institute for Policy Studies (GRIPS)
FDRE Policy Studies Institute (PSI)

Questionnaire for Startups’ business operation and youth entrepreneurship research

1) **Objective and confidentiality**

This questionnaire is designed to assess the state of youth entrepreneurship and various forms of interventions being made to promote start-ups in low-income countries taking Ethiopia as a case. You are among the potential entrepreneurs/youth randomly selected for this study. The result of this study will play a crucial role in fostering the entrepreneurship and small business development through generating credible and up-to-date information that could be utilized by policy makers, the government, entrepreneurs, and other stakeholders. To this effect, your candid response to the questions provided is crucial. You have the right to refuse to answer any question, but answering is much appreciated. All the answers you provide here will remain confidential and it will be used only for the research’s purpose and information that identify you or your enterprise will not be published in any form. The questionnaire does not consume more than 15 minutes and upon the finalization of the interview questions, you will be provided with a 50 Birr worth mobile card (airtime top-up) as a compensation to participate in this survey. Therefore, you are kindly requested to provide correct information for all questions.

Thank you in advance!

Are you willing for the interview? 1) Yes 2) No

ይህ መጠይቅ በኢትዮጵያ ውስጥ የወጣቶች የስራ ፈጠራ እንቅስቃሴና ለአዳዲስ ኢንተርፕራይዞች የሚያስፈልጉ ወይም እየተደረጉ ያሉ ድጋፎችን በመለየት ለጀማሪና አነስተኛ ኢንተርፕራይዞች ምስረታና እድገት እንቅፋት የሆኑ ጉዳዮችን ማቃለል የሚያስችል የፖሊሲ ሃሳብ ለማመንጨት የሚያግዝ ጥናት ለማጥናት በኢፌዴሪ የፖሊሲ ጥናት ኢንስቲቱት አማካኝነት የተዘጋጀ ነው። ይህም ጥናት አላማ ያደረገው ባልፉት ጥቂት አመታት ውስጥ የራሳቸውን ስራ የፈጠሩ ወይም ለመፍጠር ጥረት ያደረጉና የተለያዩ መንግስታዊና መንግስታዊ ካልሆኑ ተቋሞች ጋር ግንኙነት ማድረግ የቻሉ ወጣቶችን ነው። እርስዎም ከነዚህ ወጣቶች መካከል ለጥናቱ ሃሳብ ለመስጠት በናሙናነት ከተመረጡት ወጣቶች ውስጥ አንዱ በመሆንዎና የእርስዎ ሃሳብ ጥሩ ግብዓት ስለሚሆን ትብብርዎን እንጠይቃለን። ቃለ መጠይቁ የሚፈጀው 15 ደቂቃ ያክል ሲሆን የሚሰጡት ሃሳብ ለጥናቱ አላማ ብቻ የሚውል፣ ለሌላ አካል ተላልፎ የማይሰጥ፣ ሚስጥራዊነቱ የተጠበቀ መሆኑን አረጋግጥልዎታለሁ። ይህም ሆኖ የትኛውንም ጥያቄ ያለመመለስ መብት አለዎት። የዚህን ጥናት ጥቅም ተረድተው ስለሚያደርጉልን ትብብር እያመሰገን፣ ለዚሁ ቀና ትብብርዎም የ50 ብር ካርድ ስጦታ ያዘጋጀንልዎት መሆኑን እንገልጻለን።

ለቃለ መጠይቁ ፈቃደኛ ነዎት? 1) አዎ 2) አይደለሁም

2) **Calendar: use the Ethiopian calendar year throughout.**
3) Codes: use -77 = for “Not applicable”, -99 = for “Do not know”, and -88 = for “refusal”

Name of the interviewer ____________________________

Numbers of attempts made to do the interview __________

<u>Contents</u>

Section A: General Information

Section B: Characteristics of the entrepreneurs

Section C: Entrepreneurship activity, other measures of entrepreneurship

Section D: Access to finance, substitute entrepreneurship training, and other services

Section E: Wage employment Conditions

Section F: Miscellaneous Questions

Q.code	Question and instructions	Answer codes	Answers
Section A: General information			
a1	Date of interview (dd/mm/year) in E.C		..…/…..../…. ….
a2	Entrepreneur’s ID		
a3	Respondents Name		
a3_1	Respondent’s **position**	1 = Target respondent (owner/manager or team leader) 2 = owner/team member, 3 =employee 4=family member of any owner (Spouse, father, sister, brother etc) 5 =friend 6= Others	|______|
	Section B: Characteristics of the entrepreneur		
b1	Sex	1 = Male 2 = Female	|______|
b2	Highest completed education ያጠናቀከው **የትምህርት ደረጃ**	1= No formal education 2= Some primary school (not completed) 3=primary school 4= Highschool (9-10) 5= Preparatory school (11-12) 6= TVET 7= Diploma (non-vocational) 8= First Degree 9= Masters/MD/VDM 10= PhD	|______|
	Section C: Entrepreneurship activity and other measures of entrepreneurship		
c1.	Do you derive an income from activities other than wage employment, that is, are you **self-employed** at this time? በቅጥር ከሚገኝ ደሞዝ ውጭ ገቢ የሚያስገኝ የግል ስራ አለህ?	1= Yes 2= No	|______|
c2.	Do you **own a micro or small business** (formal or informal, individually or in group) at this time? በአሁኑ ሰዓት በግልህ ወይም ከሌሎች ጋር በማህበር ባለቤት የሆንክበት አነስተኛ ወይም ጥቃቅን **ኢንተርፕራይዝ አለህ?**	1= Yes 2= No (>>skip to c3 and its sub-questions)	|______|
c2_1	Economic **sector** that the enterprise is engaged in ኢንተርፕራይዙ በየትኛው **ሴክተር** ውስጥ ይመደባል/ምን አይነት ስራ ነው?	1= ICT 2= Trade 3= Other Services 4=manufacturing 5= Construction 6= Agriculture/urban agriculture 7= Others(specify)	|______|
c2_2	When it started operation? **መቼ** ነው ስራ የጀመረው?	*For enumerators: if s/he has more than one business, please refer to the one where the entrepreneur works more hours in a week.*	
c2_2a	**Year** in E.C _it started operation ስራ የጀመረበት **ዓ/ም**		|______|

c2_2b	**Month** in E.C it started operation ስራ የጀመረበት **ወር**	*For programmer: Insert 13 months' codes; Start with September*	|______|
c2_3	Do you have a business **license**? **የንግድ ፈቃድ** አለው?	1= Yes 2= No (>>skip to c2_4)	|______|
c2_3a	**Year in E.C** when you have got the license? የንግድ ፈቃድ ያገኙበት **ዓ/ም**	For programmer: allow only 4 digits	|__ __ __ __|
c2_3b	**Months** in E.C when you have got the license? የንግድ ፈቃድ ያገኙበት **ወር**	*For programmer: Insert 13 months' codes;*	|______|
c2_3c	Where the business has been registered? (For enumerators: write govt's office name, kebele, woreda, city) የት ነው የተመዘገበው/የንግድ ፈቃዱን የወሰደበት ቢሮ ስም?		
c2_3c 1	What is your Tax Identification Number(TIN)?		
c2_3d	What is the **legal form** of the business? በየትኛው **አደረጃጀት** ነው የተደራጀው?	1= Sole proprietorship/በግል 2= Partnership/በሽርክና 3= Cooperative/በማህበራት 4= Private Limited company/ ኃላ የተ የግል ኩ 5= Joint venture 6= Corporation 7 = others	|______|
c2_4	What is the **name** of the enterprise, if any? **የድርጅቱ ስም?**		
c2_5	Total **numbeAr of owners** at this time? በአሁኑ ጊዜ የድርጅቱ **ባለቤቶች ብዛት** ስንት ነው?		|______|
c2_5a	How many of the owners are **females**? ምን ያህሉ **ሴት** ናቸው?		|______|
c2_6	**Total number of workers** including working owners, salaried, and apprentices, permanent or temporary/seasonal at this time በአሁኑ ጊዜ ያለው **የሰራተኛ ብዛት** /ድርጅቱ ውስጥ የሚሰሩ ባለቤቶች፣ ጊዚያዊና ቋሚ ጨምሮ/		|______|
c2_6a	How many of the current workers are **permanent** (including working owners)? ድርጅቱ ላይ የሚሰሩ **ባለቤቶቹን ጨምሮ** ምን ያህሉ **ቋሚ ሰራተኞች** ናቸው?		|______|
c2_6b	How many of the current workers **are salaried** (excluding owners)? (**ባለቤቶቹን ሳይጨምር** ምን ያህሉ ሰራተኞች **ደሞዝ** ይከፈላቸዋል?)		|______|
c2_6c	Number of working owners (በዚሁ ድርጅት ውስጥ የሚሰሩ ባለቤቶች ብዛት)		|______|

c2_7	Have you made any **business expenses (investment) so far**? ድርጅቱ ላይ እስካሁን **ያወጣኸው ወጭ አለ**?	1= Yes 2= No	|______|
c2_7a	Please **describe** the **most recent expenses** you had more in detail. (በቅርቡ **ለምን ጉዳይ** ወጭ እንዳደረክ ንገረኝ እስኪ?)		
c2_8a	What were the monthly **sales** of the business **last month** in Birr? (For enumerator: write 0 if there was no any sale) **ባለፈው ወር** የነበረው ወርሃዊ ጠቅላላ **ሽያጭ** ስንት ብር ነበር?		|______|
c2_8b	What were the **average monthly sales** of the business over the last **6 months** in Birr? (For enumerator: write 0 if there was no any sale) (ባላፉት **6 ወራት** ውስጥ የነበረው **አማካኝ ወርሃዊ ሽያጭ** ስንት ብር ይሆናል?)		|______|
c2_8c	Please **describe** the most **recent sale** you had more in detail. እስኪ በቅርቡ **ምን አይነት ምርት** ወይንም አገልግሎት **እንደሸጣችሁ** ንገረኝ?		
c2_9a	After paying all expenses (but not including any income you paid yourself and other owners), what was the **net income** of the business (the **profit** of the business) during the ***last* month**? ባለፈው ወር ድርጅቱ ምን ያህል ብር አተረፈ/ **ለባለቤቶቹ** የሚከፈል ድርሻ ወይም ደሞዝ ካለ የትርፉ አካል አድርገን እናስብ/		|______|
c2_9b	After paying all expenses (but not including any income you paid yourself and other owners), what was the average monthly net income of the business (the **average profit** of the business) over the last **6 months?** (ባላፉት **6 ወራት** ውስጥ የነበረው **አማካኝ ወርሃዊ** ትርፍ ስንት ብር ይሆናል?)	For programmer: after this question, skip to c7_1 for this group.	|______|
c3.	Did you have any business which is **permanently closed** at this moment? ሲሰራ ኑሮ እስከናካቴው **የተዘጋ ቢዝነስ** ነበረህ?	1= Yes 2= No (>>> skip to c4)	|______|
c3_1	**How long** the business had **operated** (in months) before it was closed? ከመዘጋቱ በፊት **ለምን ያህል ጊዜ ሰራ?**		|______|
c3_2	**When** the business was closed?		

	መቸ ነበር የተዘጋው?		
c3_2a	Year in E.C/ ዓ/ም		|______|
c3_2b	Month / ወር		|______|
c3_3	What were the **main reasons** for closing the business? (Multiple answers are possible) በምን ምክንያት ነበር የተዘጋው?	1= I lost interest (ፍላጎት ስላልነበረኝ) 2= bankruptcy/business was not profitable (በኪሳራ/አዋጭ ስላልነበር) 3= The personal conditions like health, family issues, child care (በግል ጉዳይ ምሳሌ ጤና፣ ሞት፣ት/ት፣ጋብቻ፣ መውለድ፣ቦታ ለመቀየር መፈለግ 4= A better **job** opportunity came along (የተሻለ የስራ እድል ስላገኘሁ) 5= Legal conditions/ gov't inspection (በህግ ጉዳይ/በመንግስት ትእዛዝ) 6=Dispute among owners (በአባላቱ አለመስማማት) 7=Forced migration (በአስገዳጅ ስደት) 8=Lack of finance (የገንዘብ እጥረት) 9=To change sector (ስራ ለመቀየር) 10= Losing the business by accident/shock (በአደጋ) 11= Political instability (በጸጥታው ችግር/የፖለቲካው አለመረጋጋት) 12= COVID-19 13=Others (specify)	|______|
c4.	Is there **any progress** that you have been making **to start** a business? (Multiple answers are possible) አዲስ **ቢዝነስ ለመጀመር** እያደረከው ያለ **ጥረት**/ተጨባጭ እንቅስቃሴ አለ? (For programmer: ask this question for anyone whose c2==2 regardless of their answer on c3. Then, if answers of c4 includes 7, ask c4_1, then skip to c6_1 & c6_2; if answers of c4 is limited to any of 1 to 6 here, skip to c6_1 & c6_2)	1= Yes, I finished developing the product/service (አዎ፣ ምርቱን/አገልግሎቱን አበልጽገናል) 2= Yes, I have got trade license (አዎ፣ ንግድ ፈቃድ አውጥተናል) 3=Yes, I have recruited staff (የሰው ሃይል እያሟላን ነው) 4=Yes, I have set up the business premises (አዎ፣ መስሪያ ቦታ አዘጋጅተናል/ተከራይተናል 5=Yes, I raised finance (አዎ፣ ከተለያዩ ምንጮች ገንዘብ አሰባስበናል) 6=Yes, I have established supply linkages (አዎ፣ አቅራቢና ገዥ/ደምበኛ አግኝተናል) 7= Yes, I have made business expenses (አዎ፣ ወጭዎችን ማውጣት ጀምረናል) 8=No (>>>> skip to c5) (የለም/ነበር ግን አሁን ትቸዋለሁ)	|______|
c4_1	Please **describe** the most recent **expenses** you had more in detail, if any. እስኪ በቅርቡ **ምን አይነት ወጭ** አወጣህ? (For programmer: skip this question if c4≠7)		

c5.	If not owning business at this time or did not make any progress to start a business, what are the **reasons** for this? (Multiple answers are possible) በአሁኑ ሰዓት በግልህ ወይም ከሌሎች ጋር በማህበር ባለቤት የሆንክበት አነስተኛ ወይም ጥቃቅን ኢንተርፕራይዝ **የሌለህ** ወይም እንዲኖርህ እየተንቀሳቀስክ ያልሆንከው **ለምንድን ነው?** (For programmer: ask this iff c4=8)	1= I was **not interested** on business /ፍላጎት ስለሌለኝ 2= A better **job** opportunity came along /የተሻለ ስራ ስላገኘሁ 3=The **personal conditions** did not allow me to start business /በግል ሁኔታዎች ወይም ጉዳይ ምክንያት 4= The legal environment and the **bureaucracy** are not easy for me to start business /የመንግስት ቢሮክራሲው ከባድ ስለሆነብኝ 5= Lack of **Finance** /የፋይናንስ ችግር 6=Lack of know-how or **skills** / የእውቀትና ክህሎት እጥረት 7= No feasible business idea (አዋጭ የሆነ የቢዝነስ ሃሳብ ስለለሌኝ) 8=Fear of failure (ከሳራን መፍራት) 9=Political **instability** /በፖለቲካ አለመረጋጋቱ ምክንያት 10= COVID-19 pandemic 11=Other (specify) /ሌላ ካለ ይጠቀሱ	
c5_1	Are **you interested** in starting a new business in **the next** 3 years? በሚቀጥሉት 3 አመታትስ ቢዝነስ **የመጀመር ሃሳብ** አለህ?	1= Yes 2= No (>>> skip to **c7_1)**	|______|
c6_1	**Describe** the nature of the business you intend to start **ምን አይነት** ቢዝነስ ለመጀመር ነው ያሰብከው?		
c6_2	**When** do you expect the business to **start operation**? በምን ያህል **ጊዜ** ውስጥ ስራ **የሚጀምር** ይመስልሃል?	1= less than 3 months 2=3-6 months 3=6-12 months 4= after 1 or more years.	|______|
c7_1	In the next six months, will there be **good opportunities** for starting a business in the area where you live? በሚቀጥሉት 6 ወራት በምትኖርበት አካባቢ/ከተማ ቢዝነስ ለመጀመር የሚያስችል **እድሎች/መልካም አጋጣሚዎች** ይኖራል? ((For programmer: ask c7_1, c7_2 & c7_3 for everyone)	1= Yes 2= No	|______|
c7_2	Do you have the **knowledge, skill and experience** required to start a new business? አዲስ ቢዝነስ ለመጀመር የሚያስችል **እውቀት፣ክህሎትና ልምድ አለህ?**	1= Yes 2= No	|______|
c7_3	Would **fear of failure** prevent you from starting a business? **ውድቀትን የመፍራት** ዝንባሌ ቢዝነስ ለመጀመር ያግድሃል?	1= Yes 2= No	|______|
c8	In your opinion, what are the three most important business obstacles for startups in Ethiopia in general?	Please refer the code book attached at the end of this questionnaire.	1.|______| 2.|______| 3.|______|

	በአንተ አስተያየት በአጠቃላይ በኢትዮጵያ ውስጥ ለጀማሪ ወይም አዳዲስ ኢንተርፕራይዞች እንቅፋት ከሆኑት ጉዳዮች ውስጥ እስኪ 3 ዋና ዋና የምትላቸውን ንገረኝ?		
	Section D: Access to finance, entrepreneurship training, and other services		
d1.	Over the past one year, have you **applied for loan** for your business from banks, MFIs, and other **formal sources**? **ባለፈው 1 አመት ወስጥ ከመደበኛ** የፋይናንስ ተቋማት(ባንክ፣ ማይክሮ ፋይናንስ) ለቢዝነስ የሚሆን **ብድር ጠይቀህ** ታውቃለህ?	1= Yes 2= No	|______|
d2.	Over the past one year, in which of the following external sources have you raised finance/ borrow from for your business? (Multiple answers are possible) **ባለፈው 1 አመት ወስጥ** ከየትኞቹ የፋይናንስ ምንጮች **ገንዘብ ማግኘት ቻልክ?**	1= Banks 2= Microfinance institutions 3=Saving and credit cooperatives 4= angel investors (equity investment) 5= friends and family 6=local money lenders 7= suppliers and customers 8= NGOs 9=Prize money/grant from winning competitions 10= others (specify)	|______|
d3.	Over the last one year, **how much** external finance have you raised in Birr from all sources? ባለፈው **አንድ አመት** ጊዜ ውሰጥ ከላይ ከተዘረዘሩት ምንጮች **በድምሩ ስንት ብር ማሰባሰብ ቻልክ?**		|_ _ _ _ _ _ _ |
d4.	Over the last one year, have you or any of your associates, if any, ever received any **entrepreneurship** related **training** relevant to start new business or expanding existing businesses? አንተ ወይም ሌላ የድርጅታችሁ አባል ከባለፈው **አንድ አመት ውዲህ** ከቢዝነስ ጋር የተያያዙ የኢንተርፕርነርሽፕና መሰል **ስልጠናዎችን** ሰልጥናችሁ ታውቃላችሁ?	1= Yes 2= No	|______|
d5	If you have ever had any types of entrepreneurship training, from **which program (s) or agent** you have received the training? (Multiple answers are possible) በየትኛውም ጊዜ የኢንተርፕርነርሺፕ ስልጠና አግኝታችሁ የምታውቁ ከሆነ **ስልጠናውን ያዘጋጀው አካል** ማን ነበር?	1=Government offices (SME office, cooperative agency, FeSMMIPA etc) 2=NGOs and International organizations (ILO, DOT, British council…) 3= TVET institutions 4= Universities 5= Specialized institutes (Kaizen institute, leather institute, etc) 6=EDC 7= Jobs Creation commission(JCC) (Bruh)/Ministry of labor and skills 8= Business plan competition programs (SolveIT, Chigign, Ethio-Talent power series…) 9 = Private Incubators and accelerators (Bluemoon, ICEADDIS, X-hub..) 10= Financial Institutions (Banks and MFIs) 11= others (specify)	|______|

d6	Which **contents of the training** you have ever covered? (Multiple answers are possible) ስልጠናው **ምን ምን ጉዳዮችን** ዳሷል?	1= **ETW** (Entrepreneurship training week) 2=entrepreneurship competency 3= business idea development and business plan preparation 4= Pitching skills 5=Marketing 6= visual prototype and product development 7=Book keeping (accounting) 8= legal business setup 9= Kaizen 10= Management trainings 11=Technical training 12= Others (specify)	|______|
d7	What were **the actual benefits** you have got by participating in this competition (in these entrepreneurship development programs)? (Multiple answers are possible)? በነዚህ ፕሮግራሞች በመሳተፋችሁ **ምን ጥቅሞችን አገኛችሁበት?**	1=Nothing (ምንም) 2= Skills improvements, coaching, and counselling 3= Expanding business networks 4=Easy access to other government services required for the business 5=Built self-confidence 6=Ignite my interest towards entrepreneurship activity 7= Prize money/grant/seed money 8= Loan 9= visibility/promotion 10= others (specify)	
d8	Over the last one year, besides the entrepreneurship training, **which other supports** did you get from government, NGOs, and private sector, and others? (Multiple answers are possible) **ከባለፈው አመት ወዲህ** ከሌሎች መንግስታዊም ሆነ መንግስታዊ ካልሆኑ ድርጅቶች **ምን አይነት ድጋፎችን** አግኝታችሁ ታውቃላችሁ?	1= Nothing 2=Grant/financial or material support 3= Loan facilitation 4= credit/ loan provision 5=Working premises 6= Other Skill dev't trainings 7= Market linkages 8=Reduced bureaucracy to get government services 9=Job offerings 10= other(specify)	|______|
	Section E: Wage employment conditions/ተቀጥሮ ስለመስራት እናውራ		
e1.	In the **last month**, did you work at **all for pay** as a **wage or salary** earner, casual worker, agricultural worker, commission worker, or other job? (**ባለፈው ወር** ውስጥ ደሞዝ ፣አበል ወይም ሌላ አይነት ክፍያ የሚያስገኝ (ጊዜያዊ፣ቋሚ፣ የኮሚሽን ወዘተ) ስራ **ተቀጥረህ ሰርተሃል**?)	1= Yes 2= No (>>> skip to e2)	
e1_1	How much did you earn from all wage employments last month (gross salary and other benefits in Birr) ከእንዲህ አይነት ስራዎች በወሩ ውስጥ **በድምሩ ስንት ብር** አገኝህ?		
e1_2	How many hours did you work in a typical week last month in this job? (በጠቅላላው በውሩ ውስጥ **ስንት ሰአት** ሰራህ?)		
e2.	Over the last one year, did you work at **all for pay** as a wage or salary earner, casual worker, agricultural worker, commission worker, or other job? **ባለፈው አንድ አመት ውስጥስ?** (ደሞዝ ፣አበል ወይም ሌላ አይነት ክፍያ የሚያስገኝ (ጊዜያዊ፣ቋሚ፣ የኮሚሽን ወዘተ) ስራ ሰርተሃል?)	1= Yes 2= No	

e3.	What is your **work experience** as a salaried worker? በጠቅላላው ክፍያ በሚያስገኝ ስራ የምን ያህል ጊዜ የስራ ልምድ አለህ?		
e3_a	Years		
e3_b	Months		
	Miscellaneous		
a4	Physical Address of the target entrepreneur/ሌላ ጊዜ በአካል ተገናኝተን ካወራን እስኪ ተጨማሪ አድራሻ ንገረኝ		
a4_1	Telephone (Mobile 1)		|______|
a4_2	Telephone (Mobile 2)		|______|
a4_3	Telephone (landline)		|______|
a4_4	Region of Current residence (የምትኖርበት ክልል) (	**For programmer: Use region codes**	|______|
a4_5	City/ sub-city (for Addis Ababa only)	**For programmer: Code sub-city if region==Addis**	

Thank you so much for your cooperation! I will recharge a 50 Birr worth airtime top-up to your mobile.
If possible, please send me photo of your trade license to my email or telegram account.

Annex: Codebook for question # c8

1.	Business licensing and registration
2.	Electricity (power supply)
3.	Market linkage problem (Finding customers)
4.	Lack of information
5.	Penalties (or deals to enforcing government officials) for informal activities
6.	Political instability
7.	Limited skills and know-how
8.	Lack of business support services
9.	Lack of access to Credit
10.	High interest rates to borrow
11.	Higher collateral requirement
12.	Lengthy process to get loan/lease from financial institutions
13.	Lack of adequate working premise
14.	Corruption
15.	Lack of raw material
16.	Poor quality of raw materials
17.	Higher price of raw materials
18.	Macroeconomic instability (inflation, shortage of foreign currency)
19.	High taxes
20.	Tax administration
21.	Stiff competition
22.	Unable to meet quality standard
23.	Telecommunications service
24.	Trust among businesses
25.	COVID-19 pandemic
26.	Other
27.	Government policies and regulations
28.	Water Shortage
-77	Not applicable
-88	Refusal
-99	Don't know

Appendix 2.4. Some descriptive results about training

Table 3.A.2. Cross tabulation of training status of the applicants (our sample) within the program of interest Versus any program.

	Any training over the last one year(self-reported) (%)					
Training status in our Program	Bruh sub-sample		EDC sub-sample		Full sample	
	Not trained	Trained	Not trained	Trained	Not trained	Trained
Not Trained	50.25	49.75	41.18	58.82	46.31	53.69
Trained	5.66	94.34	31.46	68.54	21.83	78.17
Total	40.87	59.13	37.6	62.4	39.27	60.73

Notes: this report shows more than half of the non-trainees (our potential controls) got trained by other programs in the same period. About 78% of our treated sample reported training access in the self-reported data too implying that there are some misreporting errors in the self-reported data. Trained status in our program is obtained from administrative data.

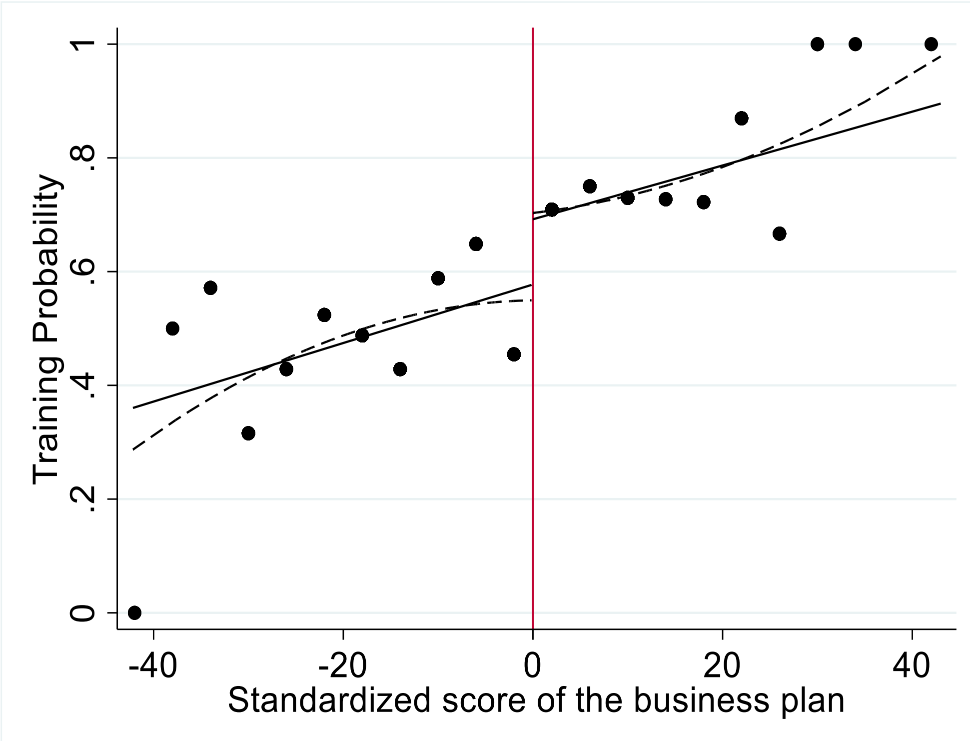
Table 2.A.3. Distribution of respondents who have ever had entrepreneurship training offered by various institutions or entrepreneurship development programs.

Training Providers	**percent**
Government offices (SME office, cooperative agency, FeSMMIPA etc)	22.9%
NGOs and International organizations (ILO, DOT, British council…)	28.1%
TVET institutions	5.5%
Universities	16.6%
Specialized institutes (Kaizen institute, leather institute, etc)	1.0%
EDC	15.6%
Jobs Creation commission (JCC) (Bruh)/Ministry of labor and skills	18.4%
Business plan competition programs (SolveIT, Chigign, Ethio-Talent power series	3.8%
Private Incubators and accelerators (Bluemoon, ICEADDIS, X-hub..)	12.3%
Financial Institutions (Banks and MFIs)	0.2%

Appendices 3.

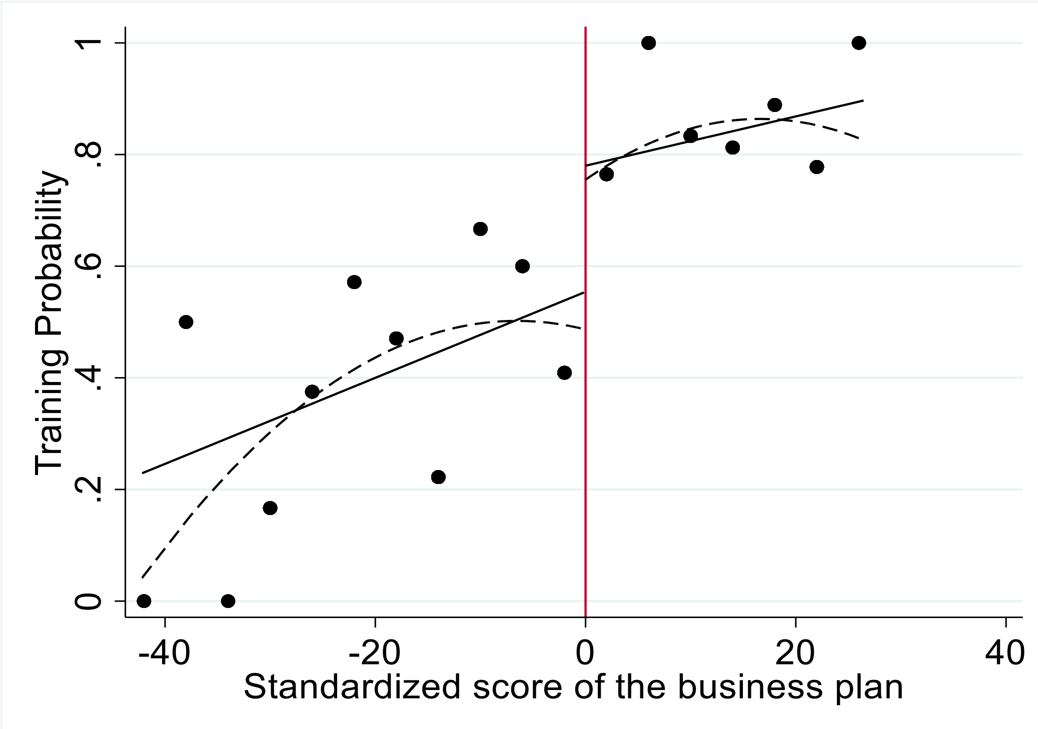
Appendix 3.1. Results for additional robustness checks for the effective first-stage

Figure 3.A.1. Effective first-stage for training by any program after the start of the business plan competition (Full sample)



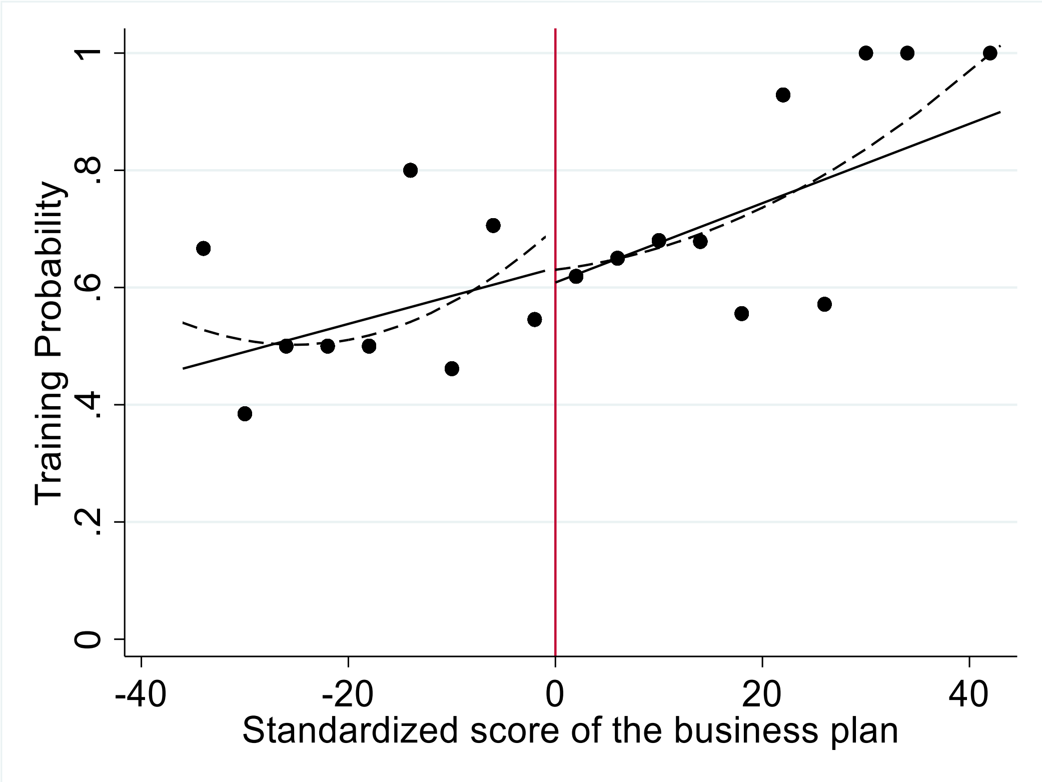
Notes: Dependent variable (treatment indicator) for this effective first-stage is dummy for taking any type of entrepreneurship training since participated in the Bruh/EDC business plan competitions. The timeline is synchronized (or limited) to be the same period with the implementation of the program of interest. Data for the treatment indicator (access to training) are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.2. Effective first-stage for training by any program after the start of the business plan competition (Bruh sub-sample)



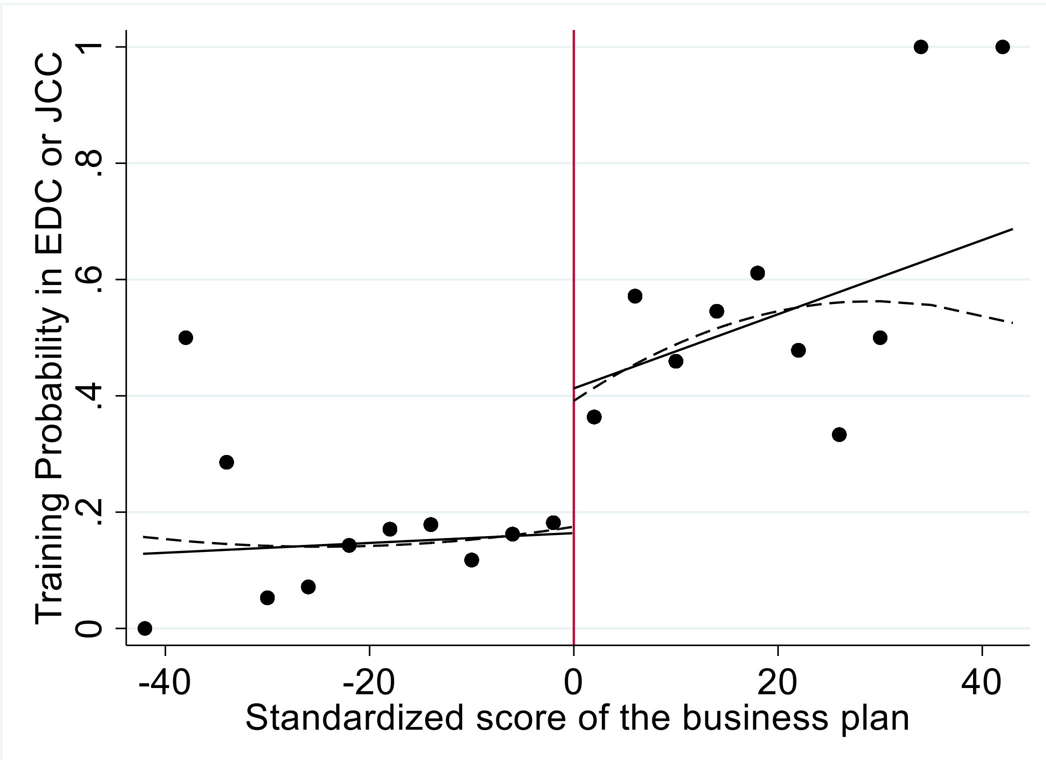
Notes: Dependent variable (treatment indicator) for this effective first-stage is dummy for taking any type of entrepreneurship training since participated in the Bruh/EDC business plan competitions. The timeline is synchronized (or limited) to be the same period with the implementation of the program of interest. Data for the treatment indicator (access to training) are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.3. Effective first-stage for training by any program after the start of the business plan competition (EDC sub-sample)



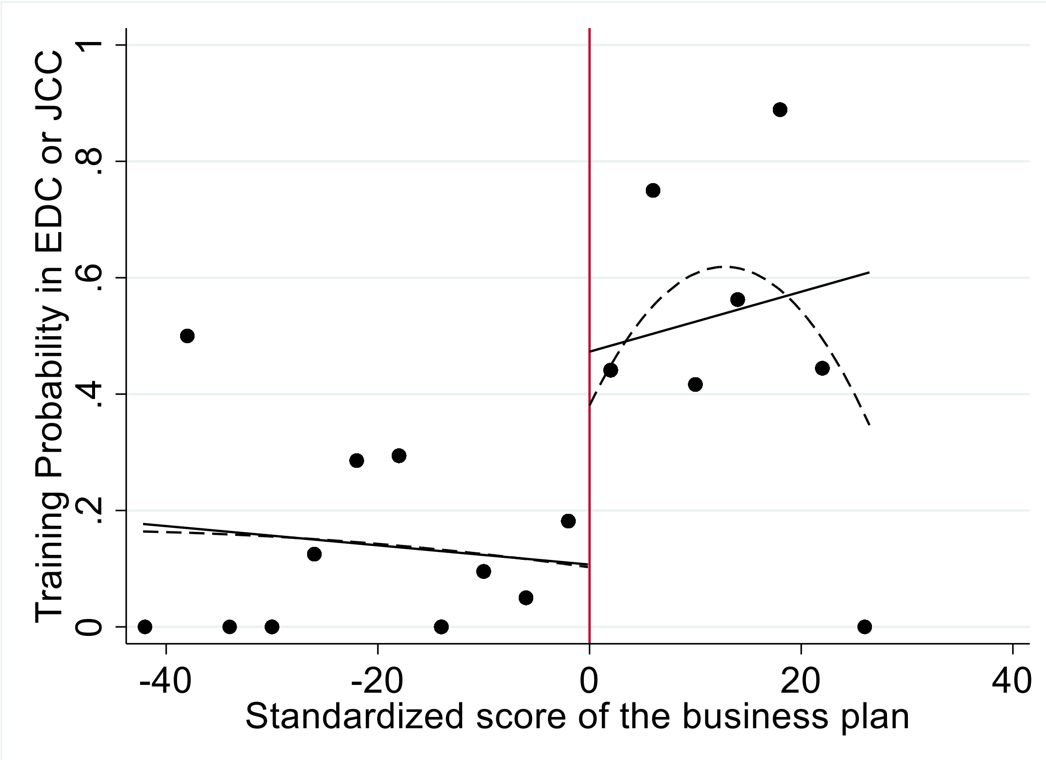
Notes: Dependent variable (treatment indicator) for this effective first-stage is dummy for taking any type of entrepreneurship training since participated in the Bruh/EDC business plan competitions. The timeline is synchronized (or limited) to be the same period with the implementation of the program of interest. Data for the treatment indicator (access to training) are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.4. Effective first-stage for training in EDC or JCC in anytime (Full sample)



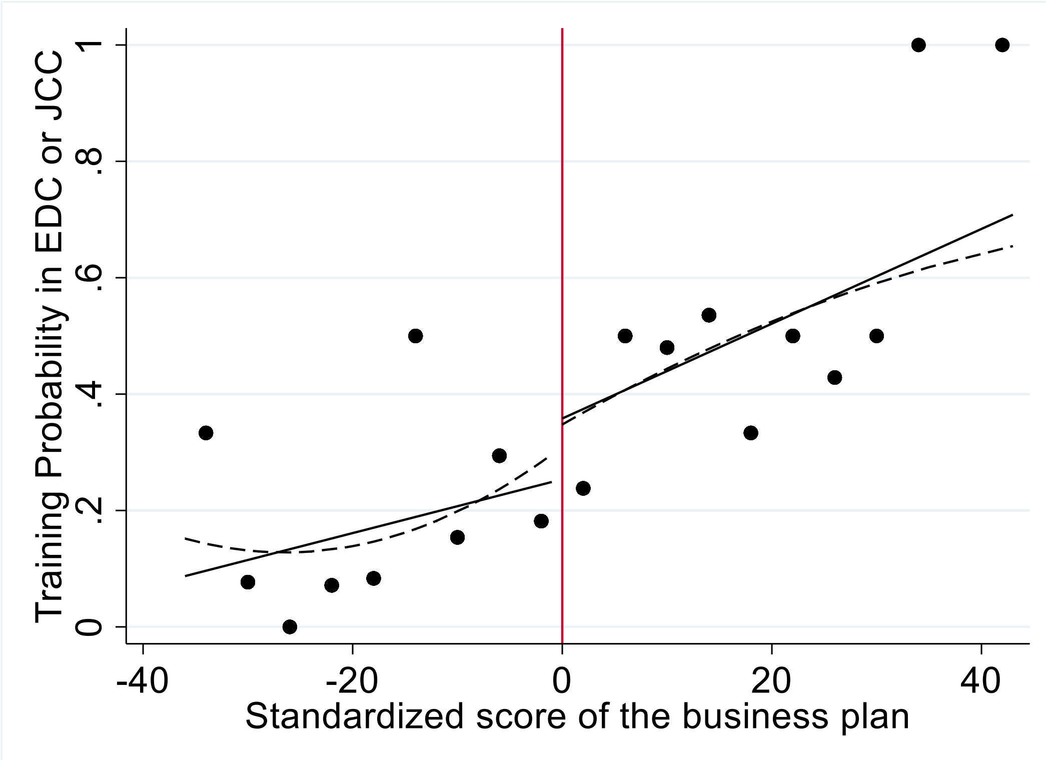
Notes: This effective first-stage graph is based on data from the full sample. Dependent variable (treatment indicator) is dummy for taking any type of entrepreneurship training in JCC or EDC (organizers of the business plan competitions) at any time. Data for the treatment indicator (access to training) are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.5. Effective first-stage for training in EDC or JCC in anytime (Bruh sub-sample)



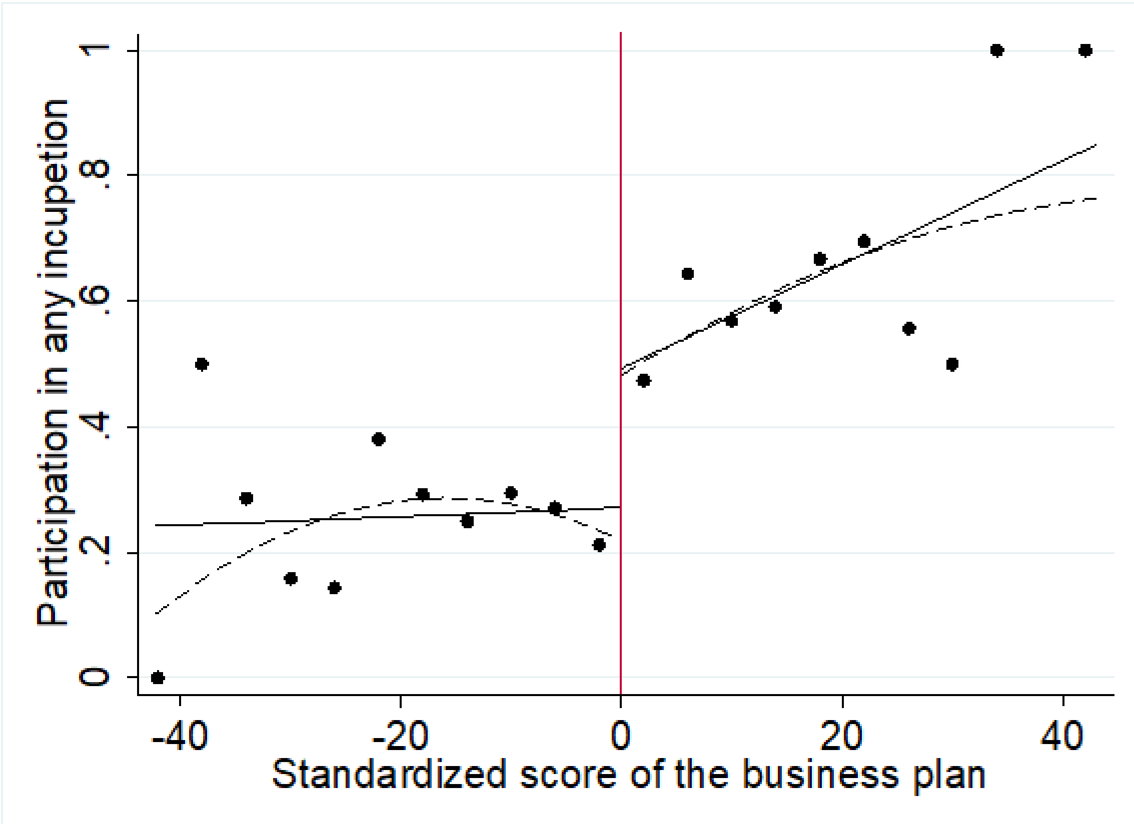
Notes: This effective first-stage graph is based on data from Bruh sub-sample. Dependent variable (treatment indicator) is dummy for taking any type of entrepreneurship training in JCC or EDC (organizers of the business plan competitions) at any time. Data for the treatment indicator (access to training) are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.6. Effective first-stage for training in EDC or JCC in anytime (EDC sub-sample)



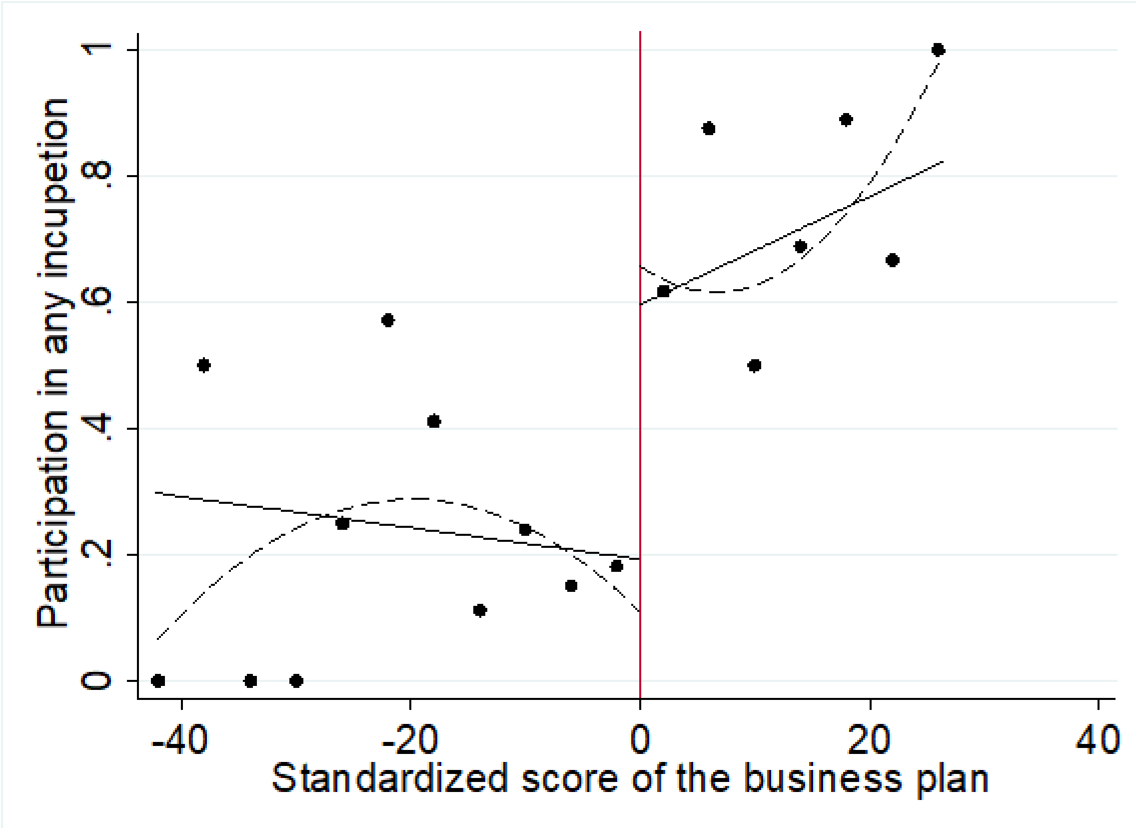
Notes: This effective first-stage graph is based on data from EDC sub-sample. Dependent variable (treatment indicator) is dummy for taking any type of entrepreneurship training in JCC or EDC (organizers of the business plan competitions) at any time. Data for the treatment indicator (access to training) are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.7. Effective first-stage result for access to training offered by any incupetion related program (full sample)



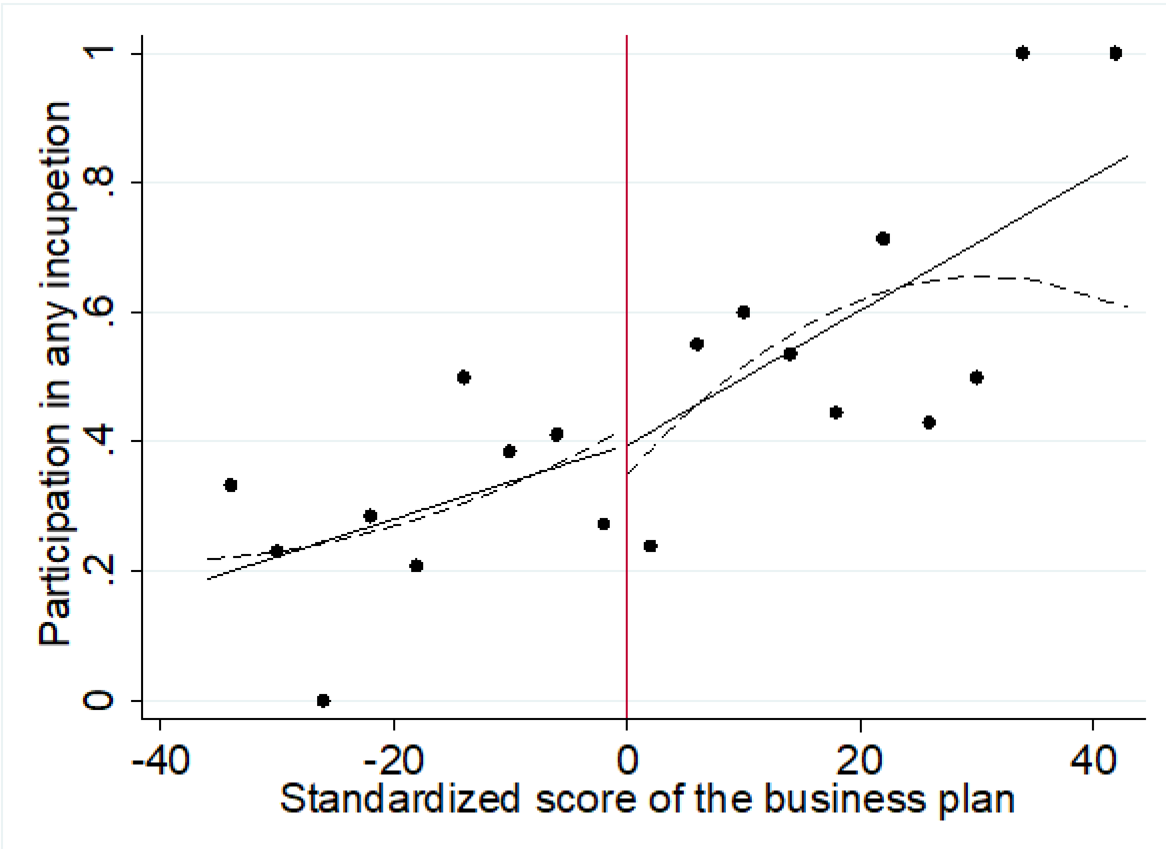
Notes: This effective first-stage graph is based on data from the full sample. Dependent variable (treatment indicator) is dummy which takes 1 if the entrepreneur has ever had entrepreneurship training in any business plan competitions, or JCC/ EDC (organizers of the business plan competitions under study) or business incubators or accelerators; and 0 otherwise. Data for the treatment indicator (access to training) are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.8. Effective first-stage result for access to training offered by any incupetion related program (Bruh sub-sample)



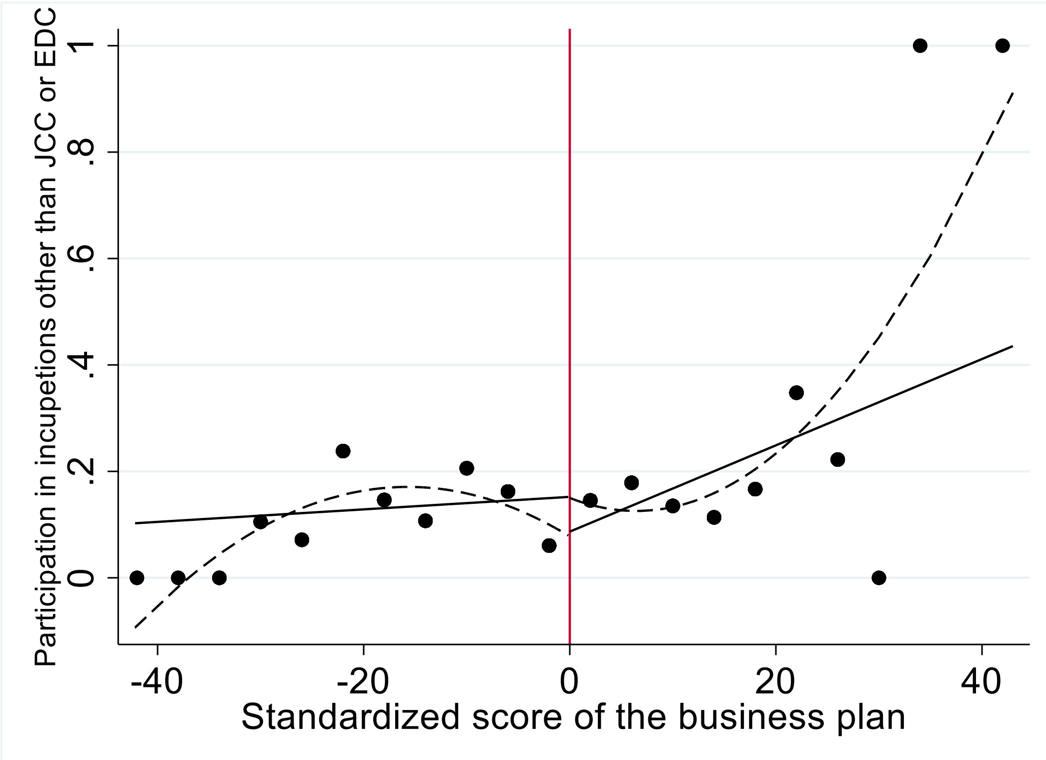
Notes: This effective first-stage graph is based on data from the Bruh sub-sample. Dependent variable (treatment indicator) is dummy which takes 1 if the entrepreneur has ever had entrepreneurship training in any business plan competitions, or JCC/ EDC (organizers of the business plan competitions under study) or business incubators or accelerators; and 0 otherwise. Data for the treatment indicator (access to training) are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.9. Effective first-stage result for access to training offered by any incupetion related program (EDC sub-sample)



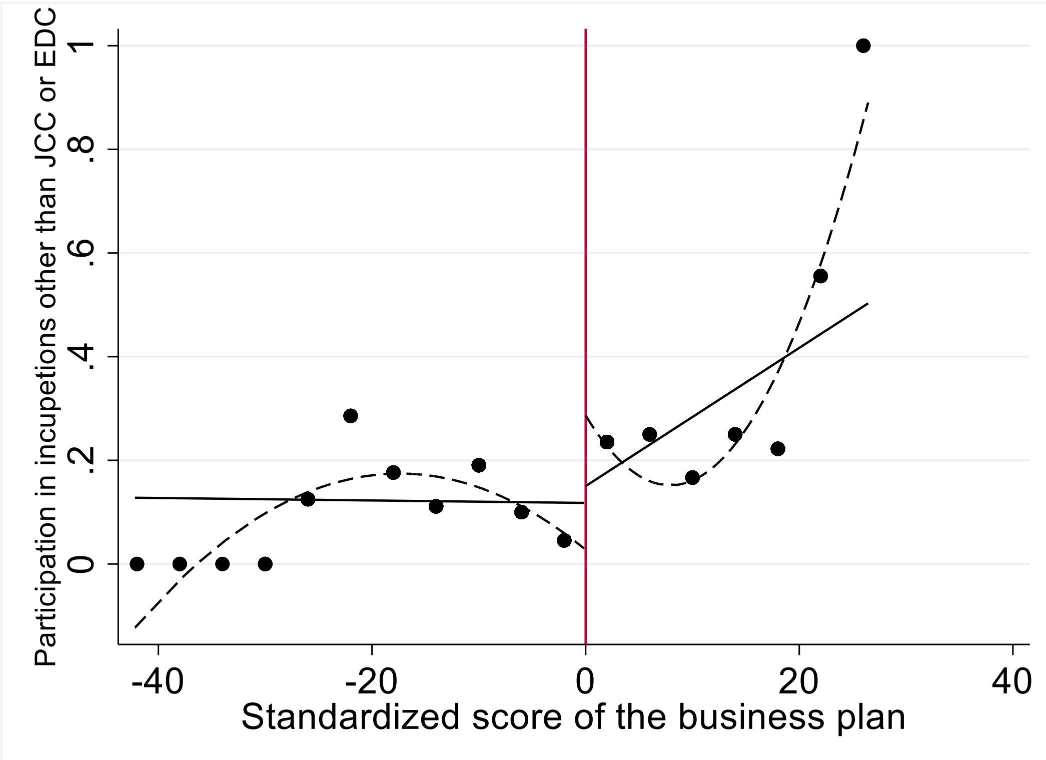
Notes: This effective first-stage graph is based on data from the EDC sub-sample. Dependent variable (treatment indicator) is dummy which takes 1 if the entrepreneur has ever had entrepreneurship training in any business plan competitions, or JCC/ EDC (organizers of the business plan competitions under study) or business incubators or accelerators; and 0 otherwise. Data for the treatment indicator (access to training) are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.10. Effective first-stage result for access to training offered by any incupetion related other than programs of JCC and EDC (full sample)



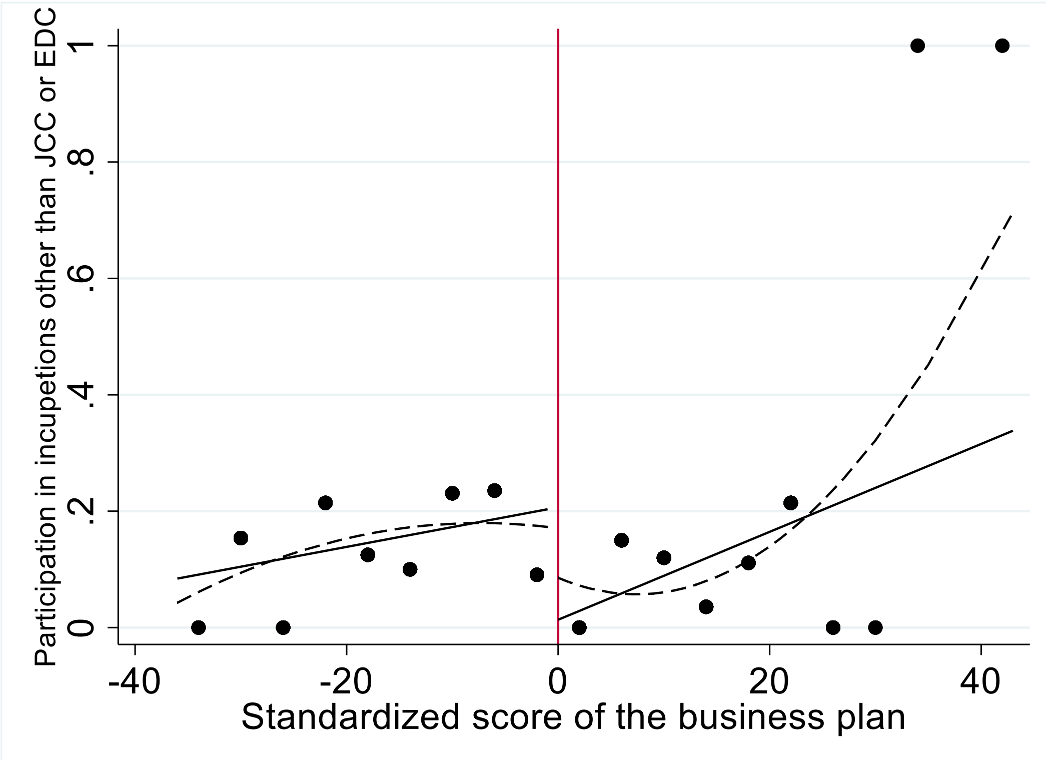
Notes: This effective first-stage graph is based on data from the full sample. Dependent variable (treatment indicator) is dummy which takes 1 if the entrepreneur has ever had entrepreneurship training in any business plan competitions excluding the business plan competition and incubation (or incupetion) under study or business incubators or accelerators; and 0 otherwise. Data for the treatment indicator (access to training) are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.11. Effective first-stage result for access to training offered by any incupetion related other than programs of JCC and EDC (Bruh sub-sample)



Notes: This effective first-stage graph is based on data from Bruh sub-sample. Dependent variable (treatment indicator) is dummy which takes 1 if the entrepreneur has ever had entrepreneurship training in any business plan competitions excluding the business plan competition and incubation (or incupetion) under study or business incubators or accelerators; and 0 otherwise. Data for the treatment indicator (access to training) are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.12. Effective first-stage result for access to training offered by any incupetion related other than programs of JCC and EDC (EDC sub-sample)

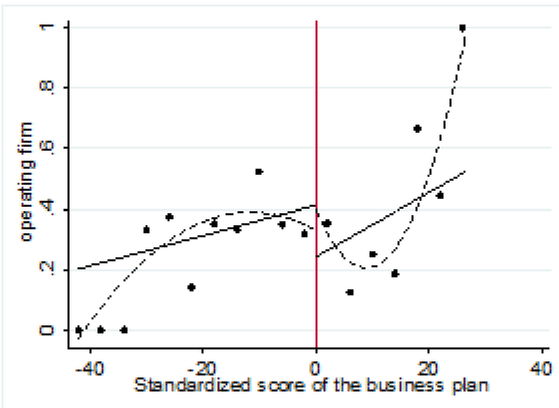


Notes: This effective first-stage graph is based on data from EDC sub-sample. Dependent variable (treatment indicator) is dummy which takes 1 if the entrepreneur has ever had entrepreneurship training in any business plan competitions excluding the business plan competition and incubation (or incupetion) under study or business incubators or accelerators; and 0 otherwise. Data for the treatment indicator (access to training) are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

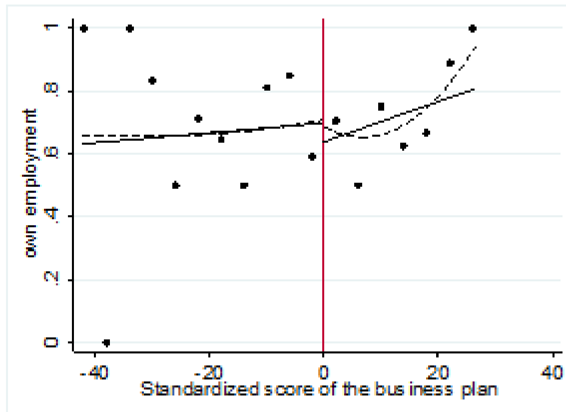
Appendix 3.2. Reduced form results for *Bruh* sub-sample

Figure 3.A.13. Reduced form results on measures of business entry (*Bruh* sub-sample)

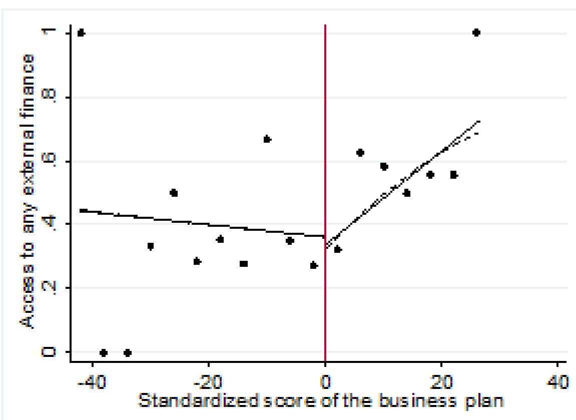
Panel A: Owning Operational business



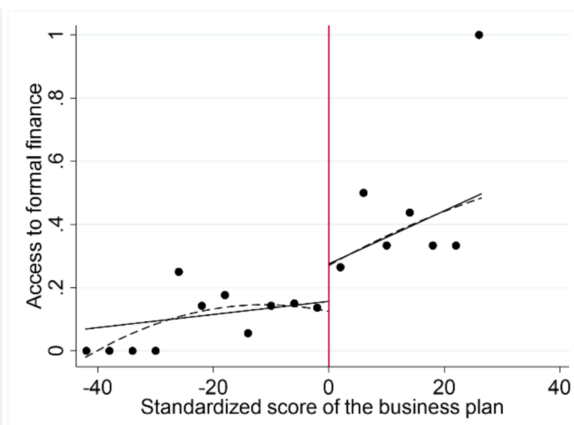
Panel B: Own employment



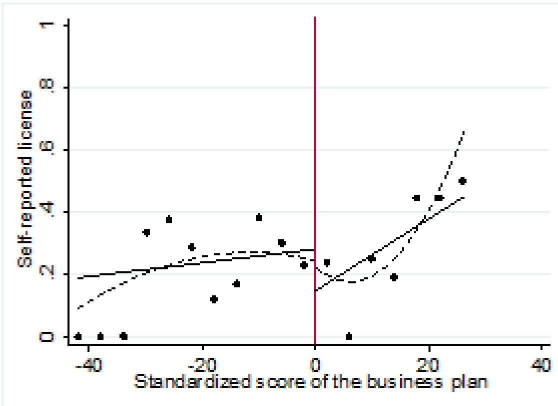
Panel C: Access to any external finance



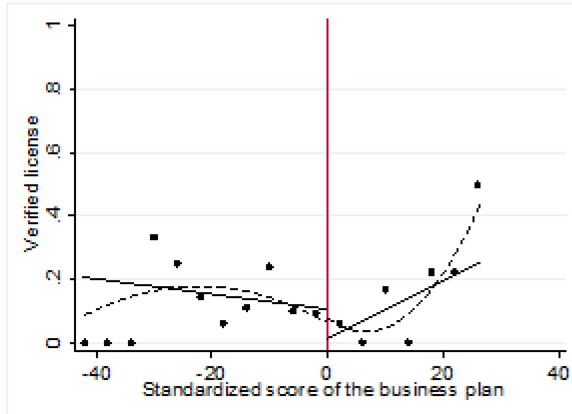
Panel D: Access to finance from formal sources



Pane E: licensed business (self-reported)



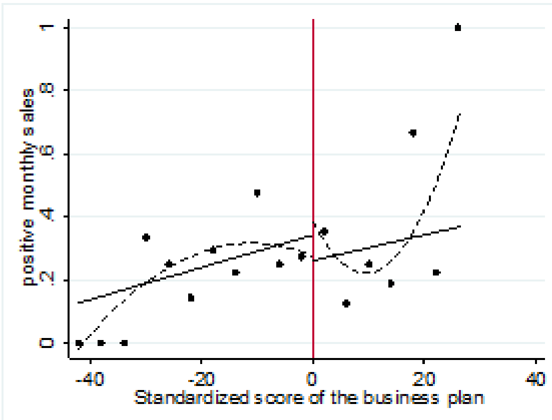
Panel F: licensed business (Verified)



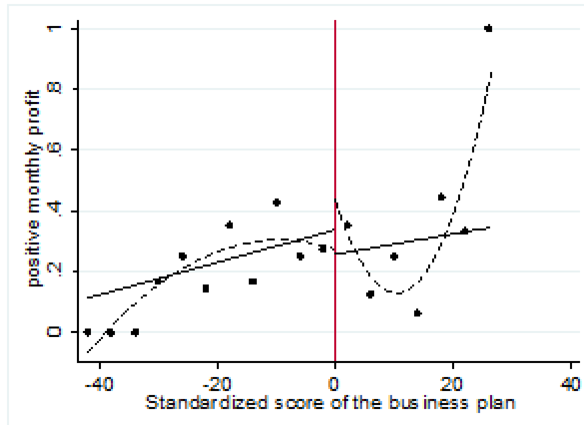
Notes: These graphs show results for Bruh sub-sample. Dependent variables in each panel are as defined for the full sample in the main body of this chapter. The results in panel E and F are not conditional on operating a business and outcomes of those who did not operate business were coded to zero. Data about outcomes from A to E self-reported in the follow-up survey while for panel F it is independently verified from local regulatory agencies. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.14. Reduced form results on dichotomous measures of firm performance (Bruh sub-sample)

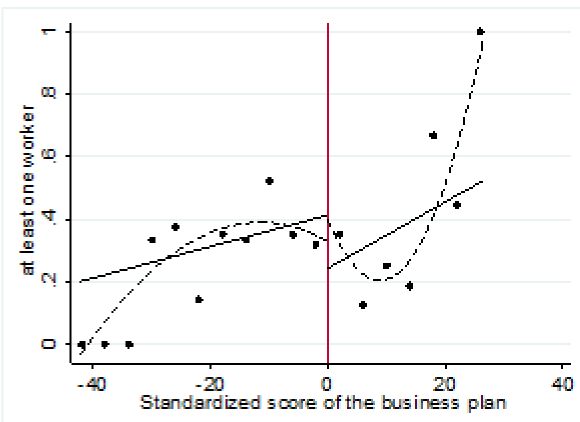
Panel A: Reporting Monthly sales



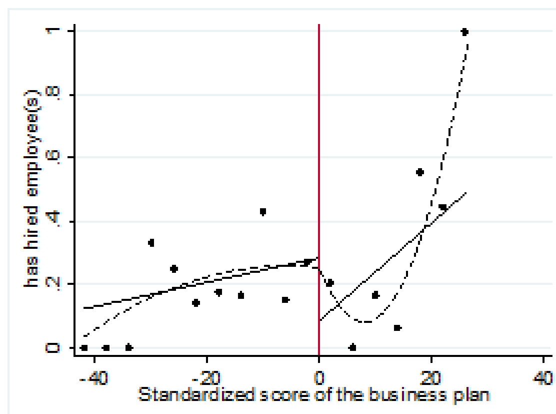
Panel B: Reporting Monthly profit



Panel C: Having at least one worker



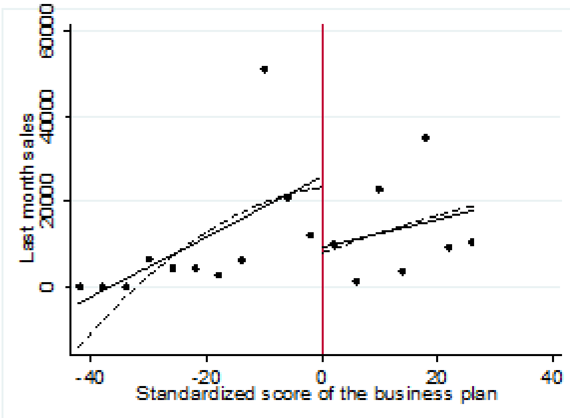
Panel D: Having hired employee



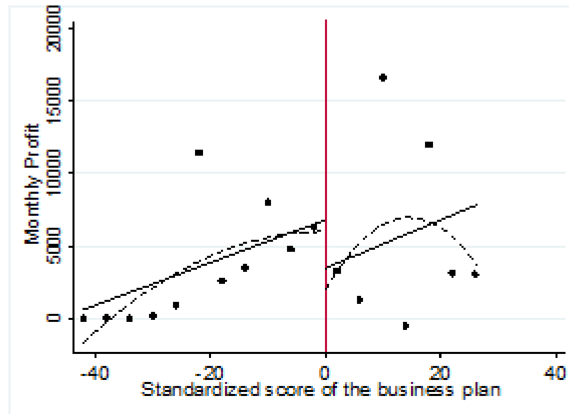
Notes: These graphs show results for Bruh sub-sample. Dependent variable for each panel is dummy for reporting any sales, profit, worker, and salaried worker (hired employee) for panel A, B, C, and D, respectively, one year after the application to the business plan competitions. The results are not conditional on operating a business and outcomes of those who did not operate business were coded to zero. Data about these outcomes are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.15. Reduced form results on continuous measures of firm performance (Bruh sub-sample)

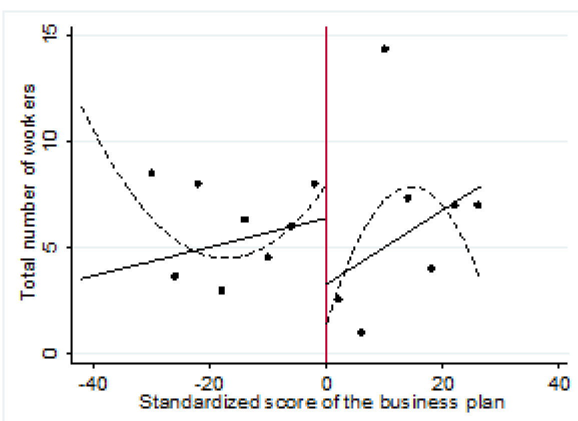
Panel A: Monthly sales



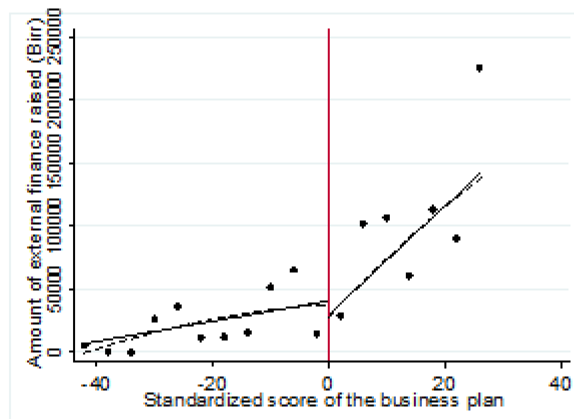
Panel B: Monthly profit



Panel C: Employment level



Panel D: Amount of external finance raised

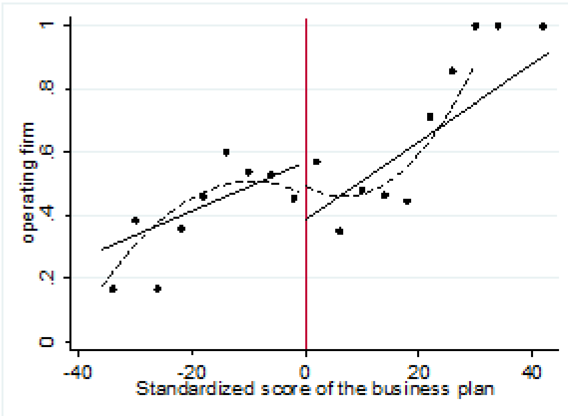


Notes: These graphs show results for Bruh sub-sample. Dependent variable for each panel the level of last month's sales (panel A) and profit (panel B) in Ethiopian Birr, total numbers of worker (panel C), and amount of external finance raised over a year in Ethiopian Birr (panel D), all as reported by the respondents a year after the application to program. The results are not conditional on operating a business and outcomes of those who did not operate business were coded to zero. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

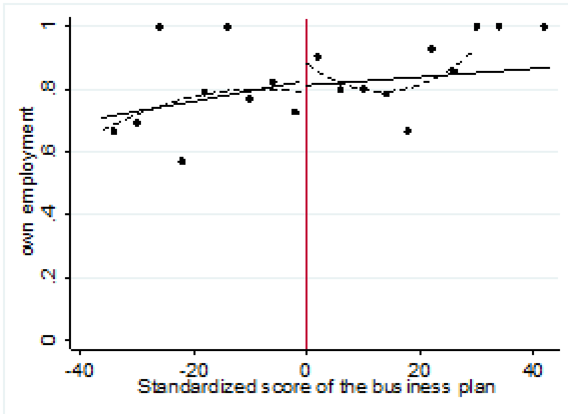
Appendix 3.3. Reduced form results for EDC sub-sample

Figure 3.A.16. Reduced form results on measures of business entry (EDC sub-sample)

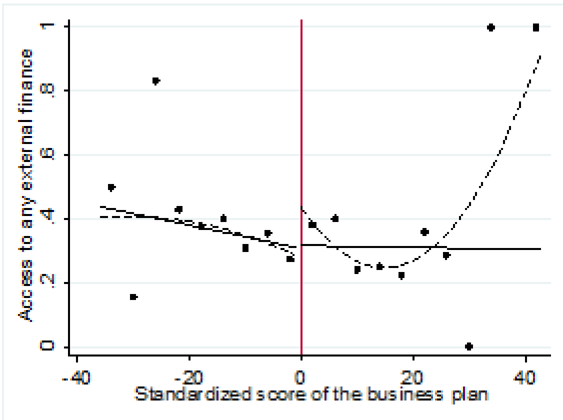
Panel A: Owning Operational business



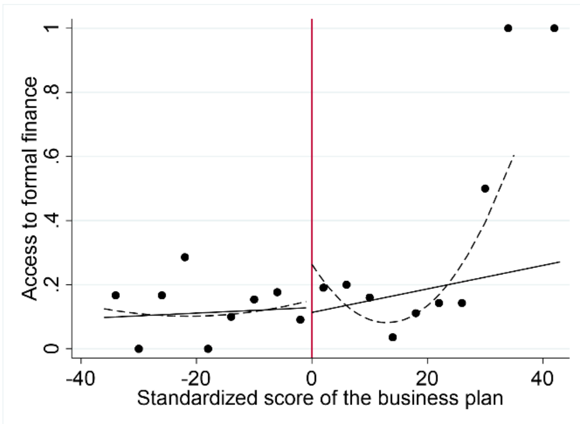
Panel B: Own employment



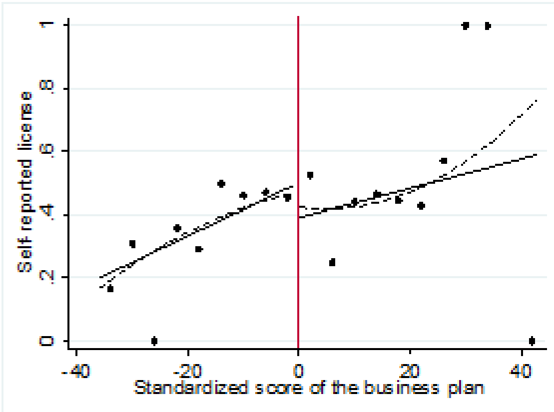
Panel C: Access to any external finance



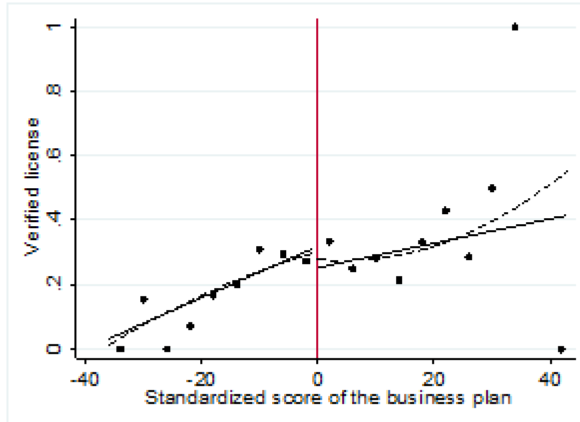
Panel D: Access to formal Loan



Panel E: licensed business (Self-reported)



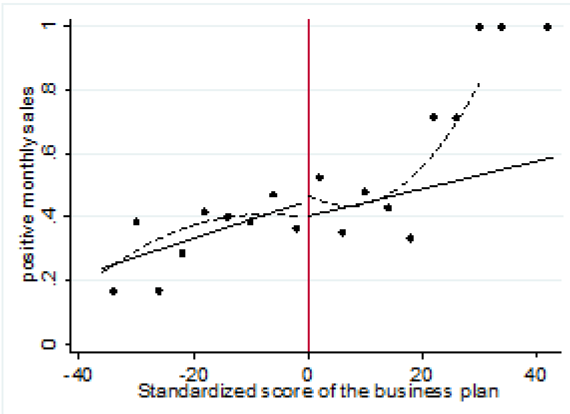
Panel F: Owning licensed business(Verified)



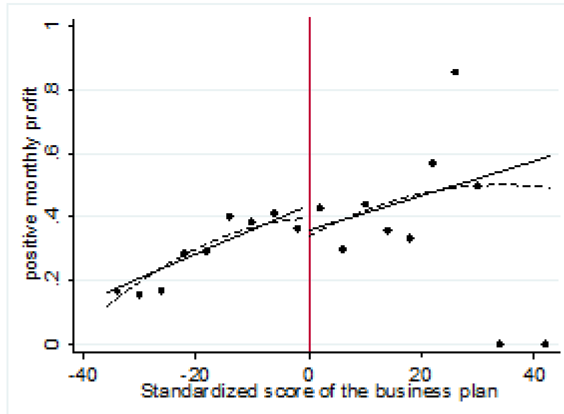
Notes: These graphs show results for EDC sub-sample. Dependent variables in each panel are as defined for the full sample in the main body of this chapter. The results in panel E and F are not conditional on operating a business and outcomes of those who did not operate business were coded to zero. Data about outcomes from A to E self-reported in the follow-up survey while for panel F it is independently verified from local regulatory agencies. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.17. Reduced form results on dichotomous measures of firm performance (EDC sub-sample)

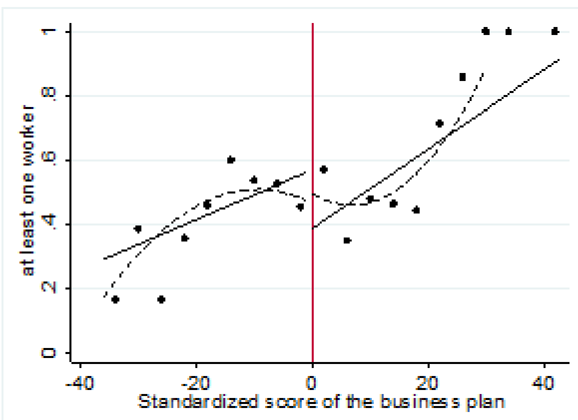
Panel A: Reporting Monthly sales



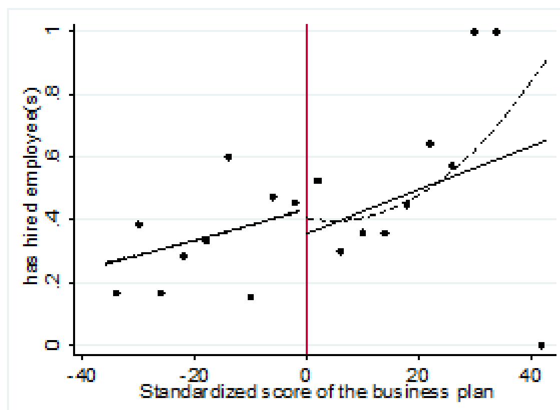
Panel B: Reporting Monthly profit



Panel C: Having at least one worker



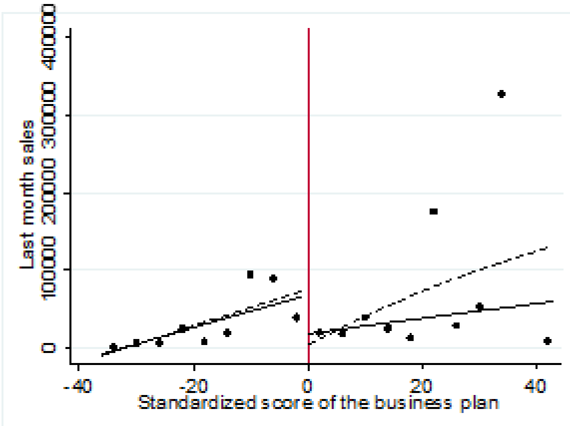
Panel D: Having hired employee



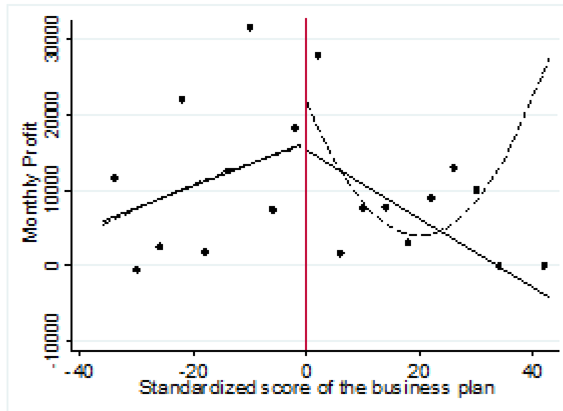
Notes: These graphs show results for EDC sub-sample. Dependent variable for each panel is dummy for reporting any sales, profit, worker, and salaried worker (hired employee) for panel A, B, C, and D, respectively, one year after the application to the business plan competitions. The results are not conditional on operating a business and outcomes of those who did not operate business were coded to zero. Data about these outcomes are self-reported in the follow-up survey. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Figure 3.A.18. Reduced form results on continuous measures of firm performance (EDC sub-sample)

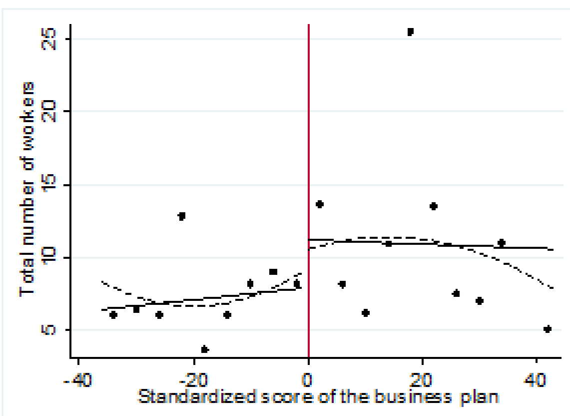
Panel A: Monthly sales



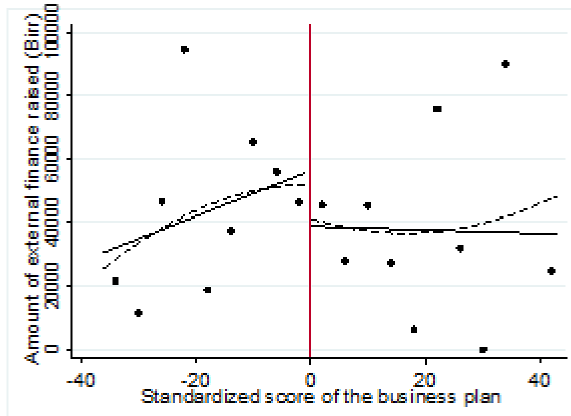
Panel B: Monthly profit



Panel C: Employment level



Panel D: Amount of external finance raised



Notes: These graphs show results for Bruh sub-sample. Dependent variable for each panel the level of last month's sales (panel A) and profit (panel B) in Ethiopian Birr, total numbers of worker (panel C), and amount of external finance raised over a year in Ethiopian Birr (panel D), all as reported by the respondents a year after the application to program. The results are not conditional on operating a business and outcomes of those who did not operate business were coded to zero. Zero is the cutoff for the running variable (the standardized score). The graphs were drawn for the entire support with a bin width of 4. The scattered black dots or circles are the bin means whereas the sold and dashed lines are linear and quadratic fits of the regressions.

Appendix 4: Additional results of the fourth chapter

Appendix 4.1. Results of linear Versus cubic specifications

Table 4.A.1. Point estimates of the effect of judges' score on probability of operating a business under with linear and cubic specifications

	Full sample		Trained	
VARIABLES	Linear	Cubic	linear	cubic
Judges' Score	0.00432***	-0.00147	0.00485***	-0.00024
	(0.00140)	(0.00249)	(0.00157)	(0.00276)
Judges' Score squared		0.00020**		0.00020**
		(0.00008)		(0.00009)
Judges' Score cube		0.00001***		0.00001***
		(0.00000)		(0.00000)
Constant	0.43265***	0.39556***	0.45104***	0.41340***
	(0.02337)	(0.03033)	(0.02615)	(0.03379)
Observations	456	456	358	358
R-squared	0.0189	0.0443	0.0239	0.0474
Adjusted R-squared	0.0168	0.0380	0.0212	0.0393
P-value for joint test of linearity		0.0024		0.0091

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is having a business in operation one year after the application. Estimation is based on a linear probability model and no other controls were included in this estimation.

Appendix 4.2: Marginal effects from the Probit and Logit regression models

Table 4.A.2A. Results of the marginal effects of the Probit model on prediction of probability of operating a business

	Full sample			Non-Trained			Trained		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Judges' Score	0.0044***	0.0034**	0.0032*	0.0001	0.0001	-0.0029	0.0050***	0.0038**	0.0043**
	(0.0015)	(0.0015)	(0.0018)	(0.0033)	(0.0033)	(0.0042)	(0.0017)	(0.0017)	(0.0020)
Existing firm		0.2273***	0.2013***		0.1519	0.1975		0.2371***	0.2044***
		(0.0510)	(0.0550)		(0.1192)	(0.1296)		(0.0565)	(0.0617)
Agriculture sector			0.0074			-0.1189			0.0323
			(0.1207)			(0.2672)			(0.1372)
IT sector			0.0231			-0.0596			0.0613
			(0.1121)			(0.2819)			(0.1278)
Manufacturing sector			0.0308			-0.2783			0.0927
			(0.1112)			(0.1925)			(0.1245)
Retail sector			-0.1123			-0.2020			-0.0654
			(0.1075)			(0.2319)			(0.1271)
Female			-0.0565			0.0745			-0.0933
			(0.0626)			(0.1533)			(0.0700)
TVET or some college			0.2402**			0.5425***			0.2126*
			(0.0936)			(0.1654)			(0.1086)
Undergraduate or graduate			0.0677			0.2642**			0.0395
			(0.0624)			(0.1123)			(0.0719)
Regional FE	No	No	Yes	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	456	456	456	98	98	97	358	358	358
Pseudo R-Squared	0.0140	0.0449	0.0784	1.29e-05	0.0137	0.130	0.0176	0.0516	0.0955

Notes: Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is having a business in operation one year after the application. Estimation is based on Probit model and the reported coefficients are marginal effects

Table 4.A.2B. Results of the marginal effects of the Logit model on prediction of probability of operating a business

	Full sample			Non-Trained			Trained		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Judges' Score	0.0044***	0.0034**	0.0032*	0.0001	0.0000	-0.0032	0.0050***	0.0039**	0.0042**
	(0.0015)	(0.0015)	(0.0018)	(0.0032)	(0.0033)	(0.0044)	(0.0017)	(0.0018)	(0.0021)
Existing firm		0.2278***	0.2031***		0.1519	0.2096		0.2377***	0.2085***
		(0.0510)	(0.0562)		(0.1193)	(0.1345)		(0.0564)	(0.0631)
Agriculture sector			0.0116			-0.1011			0.0386
			(0.1233)			(0.2804)			(0.1418)
IT sector			0.0259			-0.0514			0.0666
			(0.1145)			(0.3016)			(0.1325)
Manufacturing sector			0.0342			-0.2611			0.0971
			(0.1134)			(0.1989)			(0.1283)
Retail sector			-0.1111			-0.1890			-0.0649
			(0.1086)			(0.2394)			(0.1310)
Female			-0.0582			0.0846			-0.0962
			(0.0642)			(0.1610)			(0.0721)
TVET or some college			0.2412**			0.5474***			0.2164*
			(0.0951)			(0.1628)			(0.1118)
Undergraduate or graduate			0.0666			0.2509**			0.0401
			(0.0629)			(0.1157)			(0.0727)
Regional FE	No	No	Yes	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	456	456	456	98	98	97	358	358	358
Pseudo R-Squared	0.0140	0.0450	0.0781	1.27e-05	0.0137	0.131	0.0176	0.0517	0.0955

Notes: Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is having a business in operation one year after the application. Estimation is based on Logit model and the reported coefficients are marginal effects

Table 4.A.3. Prediction of success by additive model only using baseline covariates

VARIABLES	(1) Operating a firm	(2) Employment	(3) Sales	(4) Profit	(5) Aggregate growth
Existing firm	0.2067***	2.5063***	2.1753***	2.1827***	1.3082***
	(0.0541)	(0.9218)	(0.6074)	(0.6126)	(0.2873)
Agriculture sector	0.0431	-0.3643	0.8032	0.0740	0.0880
	(0.1132)	(2.0047)	(1.2273)	(1.1100)	(0.6284)
IT sector	0.0550	-1.9801	0.2234	-0.2119	-0.1134
	(0.1043)	(1.5727)	(1.0905)	(0.9765)	(0.5752)
Manufacturing sector	0.0594	-0.4954	0.4937	0.3762	0.1724
	(0.1030)	(1.6042)	(1.0746)	(0.9475)	(0.5726)
Retail sector	-0.0831	-1.4386	-0.5818	-0.6035	-0.4639
	(0.1043)	(1.6275)	(1.0921)	(0.9515)	(0.5924)
Female	-0.0479	-0.6978	-0.9846	-1.0921*	-0.4899
	(0.0588)	(0.8285)	(0.6179)	(0.6052)	(0.3294)
TVET or some college	0.2203**	3.8365*	1.6311	1.4890	1.0356**
	(0.0894)	(2.2181)	(1.0470)	(1.0727)	(0.5011)
Undergraduate or graduate	0.0687	0.7045	0.3436	0.3256	0.2279
	(0.0578)	(0.5514)	(0.6175)	(0.6219)	(0.3370)
Constant	0.2151*	1.5707	2.1821*	1.8539*	-0.8387
	(0.1095)	(1.5749)	(1.1222)	(1.0216)	(0.6288)
Regional FE	Yes	Yes	Yes	Yes	Yes
Panel FE	Yes	Yes	Yes	Yes	Yes
Observations	456	456	452	451	451
R-squared	0.0965	0.112	0.0871	0.0702	0.110
Adjusted R-squared	0.0657	0.0816	0.0557	0.0381	0.0794

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variables are as defined in previous tables

Appendix 4.3: Regression tables for Bruh and EDC sub-samples

Table 4.A.4. Prediction of firm ownership and survival by panel of experts by case

	Bruh			EDC		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Judges' Score	0.0025	0.0008	0.0004	0.0051***	0.0046**	0.0058**
	(0.0022)	(0.0022)	(0.0023)	(0.0018)	(0.0018)	(0.0023)
Existing firm		0.3095***	0.3384***		0.1484**	0.1002
		(0.0888)	(0.0890)		(0.0646)	(0.0653)
Agriculture sector			0.0283			-0.1312
			(0.1633)			(0.1613)
IT sector			0.2567*			-0.2509
			(0.1436)			(0.1538)
Manufacturing sector			0.1491			-0.0943
			(0.1405)			(0.1531)
Retail sector			0.1249			-0.3765**
			(0.1381)			(0.1567)
Female			-0.0310			-0.0691
			(0.0827)			(0.0800)
TVET or some college			0.3108**			0.1671
			(0.1312)			(0.1165)
Undergraduate or graduate			0.2125***			-0.0503
			(0.0748)			(0.0862)
Constant	0.3549***	0.2937***	-0.0458	0.4968***	0.4375***	0.7045***
	(0.0340)	(0.0367)	(0.1461)	(0.0319)	(0.0413)	(0.1693)
Regional FE	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes
Observations	214	214	214	242	242	242
R-squared	0.00558	0.0648	0.144	0.0292	0.0499	0.152
Adjusted R-squared	0.000886	0.0559	0.0933	0.0252	0.0420	0.0960

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is having a business in operation one year after the application. Estimation is based on a linear probability model

Table 4.A.5. Prediction of firm level of employment by panel of experts by case

	Bruh			EDC		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Judges' Score	0.0238	0.0106	0.0051	0.0939***	0.0836***	0.0930**
	(0.0186)	(0.0165)	(0.0143)	(0.0317)	(0.0305)	(0.0363)
Existing firm		2.3957***	2.7723***		2.8670**	1.6315
		(0.9049)	(0.9033)		(1.2387)	(1.2942)
Agriculture sector			-0.8453			-2.3407
			(1.2324)			(3.6536)
IT sector			1.0514			-5.9021*
			(1.2851)			(3.2985)
Manufacturing sector			0.4039			-2.0957
			(1.2036)			(3.3662)
Retail sector			1.0529			-4.6385
			(1.3136)			(3.4750)
Female			0.9437			-2.2593*
			(1.0963)			(1.2203)
TVET or some college			1.9379**			5.1782
			(0.9154)			(3.8058)
Undergraduate or graduate			1.6408***			-0.2193
			(0.4608)			(0.9942)
Constant	1.9710***	1.4967***	-0.7108	4.7600***	3.6161***	7.6825**
	(0.2988)	(0.2689)	(1.2575)	(0.6219)	(0.7353)	(3.4831)
Regional FE	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes
Observations	214	214	214	242	242	242
R-squared	0.00729	0.0575	0.117	0.0281	0.0499	0.143
Adjusted R-squared	0.00261	0.0486	0.0640	0.0241	0.0420	0.0859

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is total numbers of workers one year after the application.

Table 4.A.6. Prediction of firm sales by panel of experts by case

	Bruh			EDC		
	(1)	(2)	(3)	(4)	(5)	(6)
Judges' Score	0.0281	0.0099	0.0054	0.0471**	0.0407*	0.0308
	(0.0215)	(0.0209)	(0.0214)	(0.0210)	(0.0214)	(0.0273)
Existing firm		3.1595***	3.5236***		1.7046**	1.2172
		(0.9787)	(0.9550)		(0.7540)	(0.7506)
Agriculture sector			1.8106			-1.9813
			(1.3628)			(1.9678)
IT sector			3.0343***			-3.6726*
			(1.1098)			(1.8754)
Manufacturing sector			2.6309**			-2.1935
			(1.0858)			(1.8437)
Retail sector			2.7548***			-5.3288***
			(1.0422)			(1.8767)
Female			-0.0784			-1.5324*
			(0.8499)			(0.8104)
TVET or some college			3.2228**			0.9251
			(1.4193)			(1.4501)
Undergraduate or graduate			1.8095**			-0.6670
			(0.7764)			(0.9225)
Constant	3.1822***	2.5674***	-1.6022	4.7091***	4.0347***	7.6171***
	(0.3536)	(0.3726)	(1.1466)	(0.3673)	(0.4584)	(2.0237)
Regional FE	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes
Observations	213	213	213	239	239	239
R-squared	0.00673	0.0635	0.147	0.0199	0.0413	0.155
Adjusted R-squared	0.00203	0.0546	0.0960	0.0158	0.0332	0.0976

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is the Inverse hyperbolic transformation of monthly sales in Birr.

Table 4.A.7. Prediction of firm profit by panel of experts by case

	Bruh			EDC		
	(1)	(2)	(3)	(4)	(5)	(6)
Judges' Score	0.0209	0.0104	0.0114	0.0470**	0.0384*	0.0415
	(0.0216)	(0.0215)	(0.0232)	(0.0212)	(0.0213)	(0.0304)
Existing firm		1.8282	2.0226*		2.3987***	1.9240***
		(1.1186)	(1.1668)		(0.7252)	(0.7383)
Agriculture sector			2.0405*			-3.0256*
			(1.2101)			(1.7836)
IT sector			2.2333**			-3.8442**
			(1.0132)			(1.6307)
Manufacturing sector			2.2476**			-2.3918
			(0.9546)			(1.5967)
Retail sector			2.0588**			-4.5749***
			(0.9173)			(1.6018)
Female			-0.5804			-1.3983*
			(0.9046)			(0.8261)
TVET or some college			2.5971*			0.9321
			(1.5078)			(1.4877)
Undergraduate or graduate			1.4460*			-0.4703
			(0.7935)			(0.9706)
Constant	2.5332***	2.1775***	-1.0863	3.4763***	2.5249***	6.3011***
	(0.3554)	(0.3668)	(1.0677)	(0.3648)	(0.4661)	(1.8003)
Regional FE	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes
Observations	213	213	213	238	238	238
R-squared	0.00376	0.0229	0.0880	0.0199	0.0627	0.125
Adjusted R-squared	-0.000961	0.0136	0.0332	0.0158	0.0548	0.0656

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is the Inverse hyperbolic transformation of monthly profit in Birr.

Table 4.A.8. Prediction of firm aggregated growth by panel of experts by case

	Bruh			EDC		
	(1)	(2)	(3)	(4)	(5)	(6)
Judges' Score	0.0150	0.0063	0.0044	0.0331***	0.0287***	0.0283**
	(0.0111)	(0.0107)	(0.0110)	(0.0104)	(0.0105)	(0.0135)
Existing firm		1.5103***	1.7018***		1.2223***	0.8997**
		(0.5053)	(0.4932)		(0.3852)	(0.3844)
Agriculture sector			0.6257			-1.3821
			(0.6981)			(1.0584)
IT sector			1.3290**			-2.1862**
			(0.5938)			(0.9937)
Manufacturing sector			1.1163*			-1.1784
			(0.5799)			(0.9827)
Retail sector			1.1220**			-2.7481***
			(0.5646)			(1.0025)
Female			-0.0851			-0.7888*
			(0.4434)			(0.4115)
TVET or some college			1.6385**			0.7342
			(0.7303)			(0.7506)
Undergraduate or graduate			1.0139***			-0.3789
			(0.3879)			(0.4568)
Constant	-0.3774**	-0.6713***	-2.6477***	0.4571**	-0.0276	2.0993*
	(0.1800)	(0.1882)	(0.6087)	(0.1882)	(0.2260)	(1.0699)
Regional FE	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes
Observations	213	213	213	238	238	238
R-squared	0.00743	0.0578	0.136	0.0367	0.0780	0.183
Adjusted R-squared	0.00272	0.0488	0.0843	0.0326	0.0702	0.128

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is aggregate growth which is a sum of standardized values of the Inverse hyperbolic transformations of sales, profit and employment.

Table 4.A.9. Prediction of training access by score from panel of experts

VARIABLES	(1) Full Sample	(2) Bruh	(3) EDC
Judges' Score	0.0046***	0.0074***	0.0020
	(0.0014)	(0.0021)	(0.0020)
Existing firm	0.0367	-0.0346	0.0542
	(0.0430)	(0.0746)	(0.0548)
Agriculture sector	-0.1105	-0.0121	-0.1693
	(0.0878)	(0.1371)	(0.1059)
IT sector	-0.1884**	-0.1293	-0.2194**
	(0.0808)	(0.1091)	(0.1044)
Manufacturing sector	-0.0323	0.0266	-0.1031
	(0.0776)	(0.1039)	(0.1020)
Retail sector	-0.1158	-0.0846	-0.1239
	(0.0836)	(0.1071)	(0.1178)
Female	0.0146	-0.0084	0.0270
	(0.0462)	(0.0738)	(0.0627)
TVET or some college	-0.0275	-0.0156	-0.0462
	(0.0785)	(0.1380)	(0.0892)
Undergraduate or graduate	-0.0187	0.0720	-0.0775
	(0.0499)	(0.0752)	(0.0659)
Constant	0.9137***	0.8442***	0.9578***
	(0.0843)	(0.1171)	(0.1083)
Regional FE	No	No	No
Panel FE	No	No	No
Observations	456	214	242
R-squared	0.0706	0.112	0.0806
Adjusted R-squared	0.0367	0.0595	0.0195

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is having entrepreneurship training ever. Estimation is based on a linear probability model

Table A.4.11. Prediction of operating a firm and employment by panel of experts for EDC (excluding grant winners)

	Operating a firm			Employment		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Judges' Score	0.0047***	0.0042**	0.0051**	0.0972***	0.0866***	0.0961***
	(0.0018)	(0.0019)	(0.0024)	(0.0323)	(0.0311)	(0.0370)
Existing firm		0.1487**	0.1008		2.9130**	1.6452
		(0.0655)	(0.0661)		(1.2600)	(1.3123)
Agriculture sector			-0.1226			-2.3942
			(0.1632)			(3.6638)
IT sector			-0.2583*			-5.9091*
			(0.1563)			(3.3138)
Manufacturing sector			-0.1020			-2.0495
			(0.1554)			(3.3822)
Retail sector			-0.3761**			-4.6413
			(0.1585)			(3.4798)
Female			-0.0732			-2.3675*
			(0.0809)			(1.2373)
TVET or some college			0.1684			5.1627
			(0.1164)			(3.8051)
Undergraduate or graduate			-0.0598			-0.1744
			(0.0865)			(0.9986)
Constant	0.4886***	0.4294***	0.7014***	4.8078***	3.6478***	7.7339**
	(0.0323)	(0.0417)	(0.1715)	(0.6358)	(0.7489)	(3.4947)
Regional FE	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes
Observations	238	238	238	238	238	238
R-squared	0.0253	0.0460	0.150	0.0297	0.0518	0.145
Adjusted R-squared	0.0211	0.0379	0.0924	0.0256	0.0437	0.0871

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is having a business in operation one year after the application for the first three columns and total numbers of workers for the last three.

Table B3. Prediction of firm sales, profit, and aggregate growth by panel of experts for EDC (Excluding the grant winners)

	Sales			Profit			Aggregate Growth		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Judges' Score	0.0423**	0.0357*	0.0209	0.0416*	0.0325	0.0311	0.0313***	0.0267**	0.0241*
	(0.0212)	(0.0215)	(0.0274)	(0.0213)	(0.0212)	(0.0304)	(0.0106)	(0.0106)	(0.0135)
Existing firm		1.7426**	1.2704*		2.4526***	1.9922***		1.2410***	0.9214**
		(0.7626)	(0.7573)		(0.7311)	(0.7408)		(0.3905)	(0.3887)
Agriculture sector			-1.8261			-2.8451			-1.3207
			(1.9965)			(1.8069)			(1.0701)
IT sector			-3.7063*			-3.8813**			-2.2066**
			(1.9125)			(1.6665)			(1.0090)
Manufacturing sector			-2.3069			-2.4887			-1.2266
			(1.8739)			(1.6254)			(0.9959)
Retail sector			-5.3051***			-4.5460***			-2.7385***
			(1.9042)			(1.6254)			(1.0134)
Female			-1.6319**			-1.4913*			-0.8489**
			(0.8034)			(0.8218)			(0.4093)
TVET or some college			0.9449			0.9465			0.7423
			(1.4482)			(1.4806)			(0.7495)
Undergraduate or graduate			-0.8117			-0.6264			-0.4381
			(0.9236)			(0.9719)			(0.4573)
Constant	4.6051***	3.9171***	7.5092***	3.3621***	2.3914***	6.1855***	0.4122**	-0.0790	2.0597*
	(0.3713)	(0.4594)	(2.0566)	(0.3676)	(0.4647)	(1.8313)	(0.1907)	(0.2270)	(1.0831)
Regional FE	No	No	Yes	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	235	235	235	234	234	234	234	234	234
R-squared	0.0162	0.0386	0.158	0.0157	0.0608	0.126	0.0328	0.0753	0.185
Adjusted R-squared	0.0120	0.0303	0.0999	0.0115	0.0526	0.0658	0.0286	0.0673	0.129

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variables are as defined before.

Table A.4.10. Prediction of success by score from panel of experts disaggregated by firm type (with controls)

	Firm own		Employment		Sales		Profit		Aggregate growth	
VARIABLES	Existing	New	Existing	New	Existing	New	Existing	New	Existing	New
Judges' Score	0.0043	0.0021	0.0727*	0.0357	0.0366	0.0157	0.0498	0.0193	0.0299*	0.0118
	(0.0030)	(0.0019)	(0.0430)	(0.0236)	(0.0350)	(0.0197)	(0.0340)	(0.0221)	(0.0176)	(0.0096)
Agriculture sector	0.0909	0.0145	-1.1156	-0.4024	-0.3974	1.5556	-1.0490	0.9497	-0.3421	0.3183
	(0.2255)	(0.1362)	(2.1524)	(2.6572)	(2.6743)	(1.4027)	(2.4191)	(1.2766)	(1.3398)	(0.7366)
IT sector	-0.1528	0.0629	-0.9462	-3.4195*	-2.5773	0.9434	-2.5574	0.5471	-1.3779	0.0801
	(0.2153)	(0.1233)	(2.0853)	(2.0444)	(2.5361)	(1.1918)	(2.1900)	(1.0433)	(1.2781)	(0.6352)
Manufacturing sector	0.1767	-0.0106	3.1699	-3.0122	0.8088	0.3167	0.6584	0.2961	0.5694	-0.1112
	(0.2138)	(0.1210)	(2.0423)	(2.0231)	(2.5222)	(1.1645)	(2.2310)	(1.0048)	(1.2662)	(0.6213)
Retail sector	-0.0050	-0.1311	1.3035	-2.8836	-1.0858	-0.2294	-2.8517	0.2746	-0.8347	-0.3690
	(0.2270)	(0.1217)	(3.1695)	(1.9939)	(2.6808)	(1.1530)	(2.2946)	(0.9777)	(1.3529)	(0.6224)
Female	-0.1075	-0.0244	-0.5406	-0.6997	-2.1922*	-0.3686	-2.3230*	-0.3874	-0.9641	-0.2447
	(0.1171)	(0.0692)	(2.1606)	(0.6838)	(1.2871)	(0.7041)	(1.2767)	(0.6972)	(0.6807)	(0.3522)
TVET or some college	-0.0655	0.3350***	1.4965	4.6284	-0.1646	2.2597*	0.2956	1.7824	-0.0478	1.3920**
	(0.1379)	(0.1123)	(3.1770)	(2.8928)	(1.7352)	(1.3130)	(1.7223)	(1.4239)	(0.8982)	(0.6642)
Undergraduate or graduate	-0.2256*	0.1427**	0.6302	0.8485*	-1.9381	0.9944	-1.9422	0.8994	-1.0872	0.5791*
	(0.1160)	(0.0639)	(1.7236)	(0.5023)	(1.3460)	(0.6699)	(1.5298)	(0.6613)	(0.6939)	(0.3241)
Constant	0.8310***	0.1609	2.8107	3.1215	8.6509***	1.1237	7.1843***	0.9678	2.4054*	-1.1797*
	(0.2317)	(0.1268)	(2.4291)	(1.9320)	(2.6900)	(1.2124)	(2.5302)	(1.0982)	(1.3469)	(0.6399)
Regional FE	No	No	No	No	No	No	No	No	No	No
Panel FE	No	No	No	No	No	No	No	No	No	No
Observations	133	323	133	323	130	322	130	321	130	321
R-squared	0.118	0.0992	0.158	0.110	0.124	0.0642	0.134	0.0322	0.147	0.0667
Adjusted R-squared	0.00514	0.0552	0.0497	0.0662	0.00847	0.0183	0.0206	-0.0154	0.0353	0.0208

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variables are as defined in previous tables

Table A.4.11. Prediction of operating a firm and employment by panel of experts for EDC (excluding grant winners)

	Operating a firm			Employment		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Judges’ Score	0.0047***	0.0042**	0.0051**	0.0972***	0.0866***	0.0961***
	(0.0018)	(0.0019)	(0.0024)	(0.0323)	(0.0311)	(0.0370)
Existing firm		0.1487**	0.1008		2.9130**	1.6452
		(0.0655)	(0.0661)		(1.2600)	(1.3123)
Agriculture sector			-0.1226			-2.3942
			(0.1632)			(3.6638)
IT sector			-0.2583*			-5.9091*
			(0.1563)			(3.3138)
Manufacturing sector			-0.1020			-2.0495
			(0.1554)			(3.3822)
Retail sector			-0.3761**			-4.6413
			(0.1585)			(3.4798)
Female			-0.0732			-2.3675*
			(0.0809)			(1.2373)
TVET or some college			0.1684			5.1627
			(0.1164)			(3.8051)
Undergraduate or graduate			-0.0598			-0.1744
			(0.0865)			(0.9986)
Constant	0.4886***	0.4294***	0.7014***	4.8078***	3.6478***	7.7339**
	(0.0323)	(0.0417)	(0.1715)	(0.6358)	(0.7489)	(3.4947)
Regional FE	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes
Observations	238	238	238	238	238	238
R-squared	0.0253	0.0460	0.150	0.0297	0.0518	0.145
Adjusted R-squared	0.0211	0.0379	0.0924	0.0256	0.0437	0.0871

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variable is having a business in operation one year after the application for the first three columns and total numbers of workers for the last three.

Table A.4.12. Prediction of firm sales, profit, and aggregate growth by panel of experts for EDC (Excluding the grant winners)

	Sales			Profit			Aggregate Growth		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Judges' Score	0.0423**	0.0357*	0.0209	0.0416*	0.0325	0.0311	0.0313***	0.0267**	0.0241*
	(0.0212)	(0.0215)	(0.0274)	(0.0213)	(0.0212)	(0.0304)	(0.0106)	(0.0106)	(0.0135)
Existing firm		1.7426**	1.2704*		2.4526***	1.9922***		1.2410***	0.9214**
		(0.7626)	(0.7573)		(0.7311)	(0.7408)		(0.3905)	(0.3887)
Agriculture sector			-1.8261			-2.8451			-1.3207
			(1.9965)			(1.8069)			(1.0701)
IT sector			-3.7063*			-3.8813**			-2.2066**
			(1.9125)			(1.6665)			(1.0090)
Manufacturing sector			-2.3069			-2.4887			-1.2266
			(1.8739)			(1.6254)			(0.9959)
Retail sector			-5.3051***			-4.5460***			-2.7385***
			(1.9042)			(1.6254)			(1.0134)
Female			-1.6319**			-1.4913*			-0.8489**
			(0.8034)			(0.8218)			(0.4093)
TVET or some college			0.9449			0.9465			0.7423
			(1.4482)			(1.4806)			(0.7495)
Undergraduate or graduate			-0.8117			-0.6264			-0.4381
			(0.9236)			(0.9719)			(0.4573)
Constant	4.6051***	3.9171***	7.5092***	3.3621***	2.3914***	6.1855***	0.4122**	-0.0790	2.0597*
	(0.3713)	(0.4594)	(2.0566)	(0.3676)	(0.4647)	(1.8313)	(0.1907)	(0.2270)	(1.0831)
Regional FE	No	No	Yes	No	No	Yes	No	No	Yes
Panel FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	235	235	235	234	234	234	234	234	234
R-squared	0.0162	0.0386	0.158	0.0157	0.0608	0.126	0.0328	0.0753	0.185
Adjusted R-squared	0.0120	0.0303	0.0999	0.0115	0.0526	0.0658	0.0286	0.0673	0.129

Robust Standard error in parenthesis. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Dependent variables are as defined before.

Appendix 4.4. Graphical representations of the Quintile analyses

Figure 4.A.1. Proportion of applicants who operate businesses by score quintile (**full sample**)

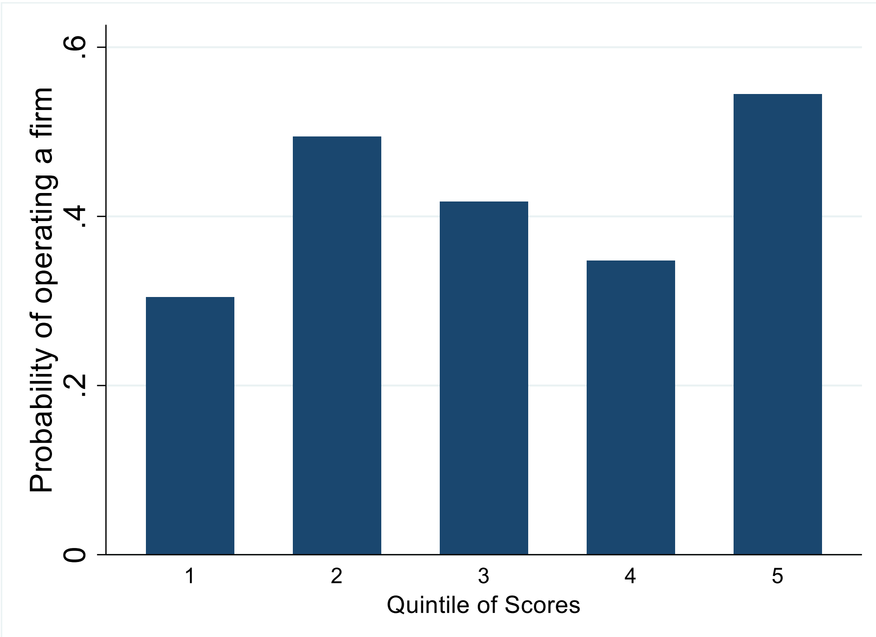


Figure 4.A.2. Level of employment by score quintile (**full sample**)

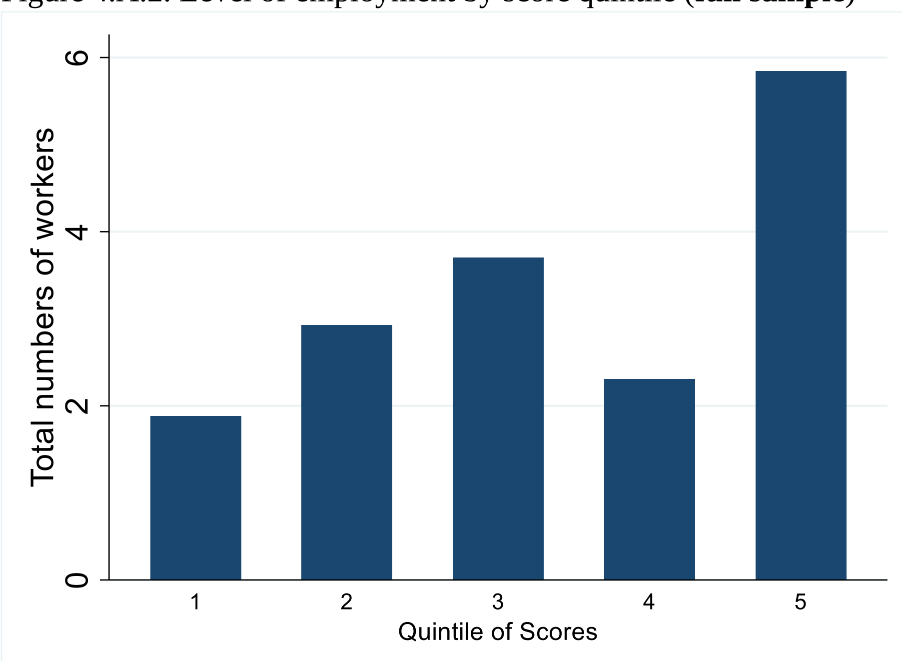


Figure 4.A.3. IHS transformation of average monthly sales in Ethiopian Birr by score quintile (**full sample**)

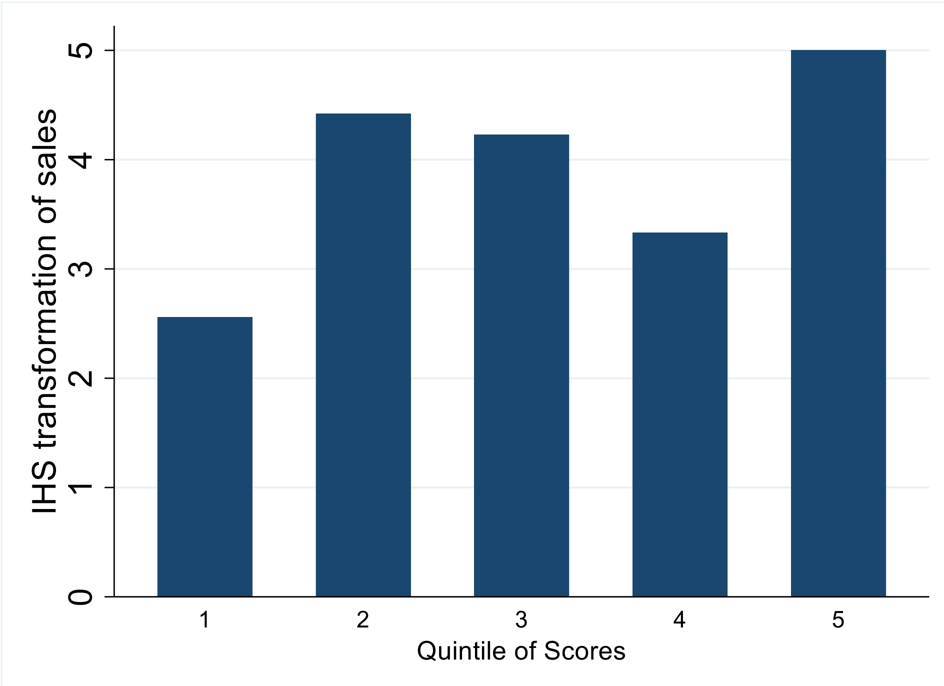


Figure 4.A.4. IHS transformation of average monthly profit in Ethiopian Birr by score quintile (**full sample**)

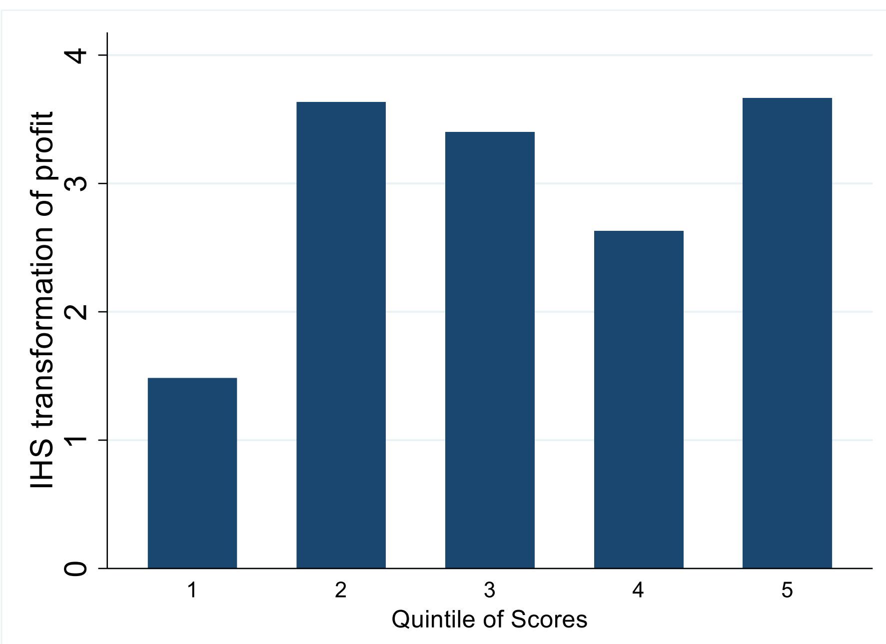


Figure 4.A.5. Aggregate growth index by score quintile (**full sample**)

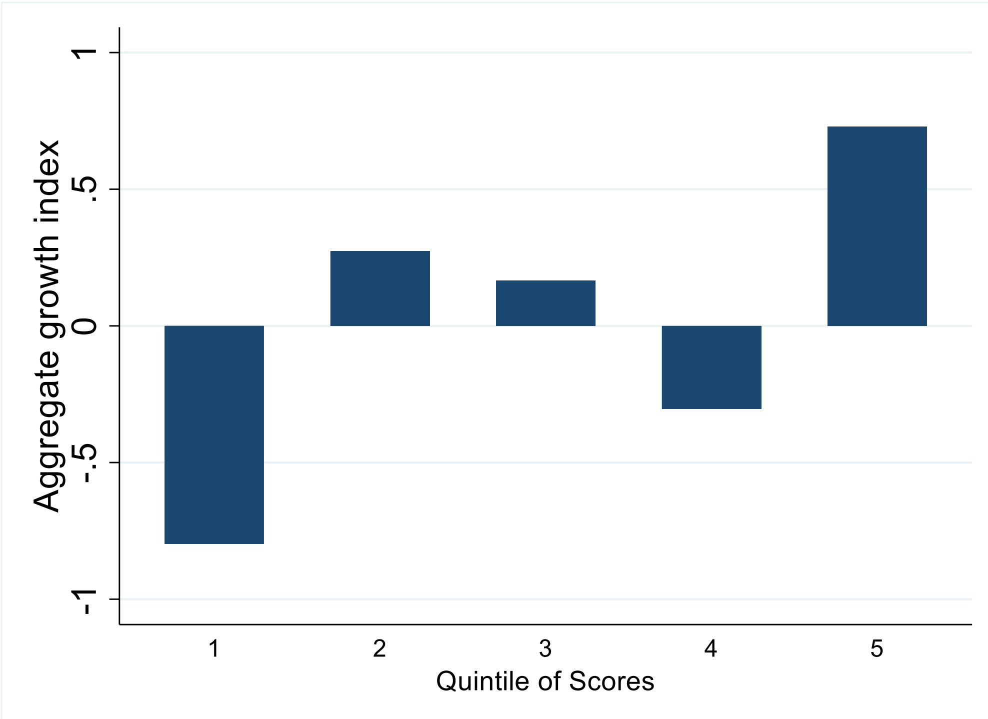


Figure 4.A.6. Proportion of applicants who operate businesses by score quintile (**Bruh**)

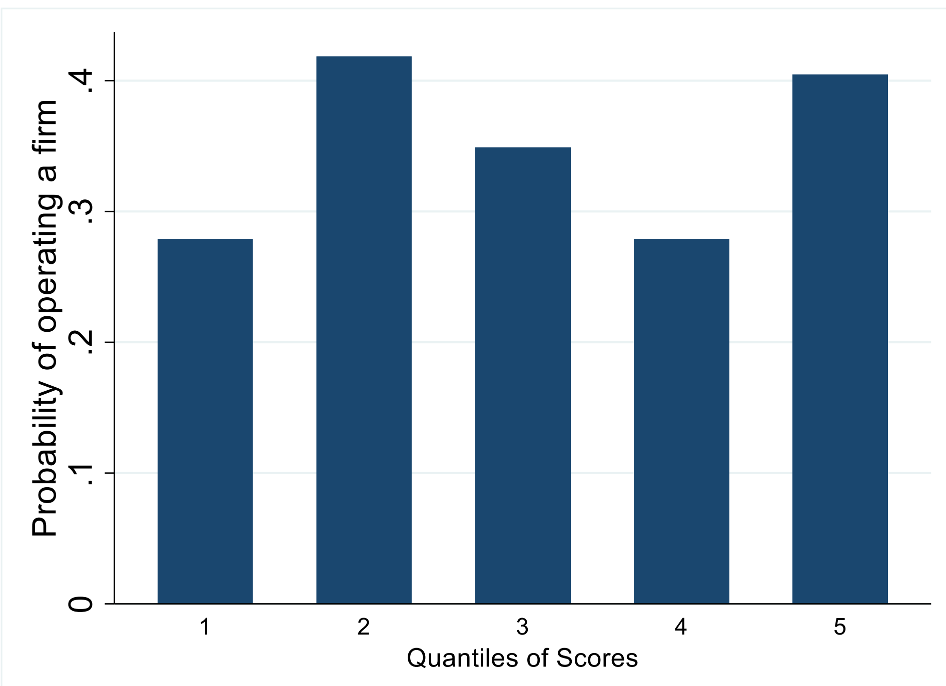


Figure 4.A.7. Proportion of applicants who operate businesses by score quintile (**EDC**)

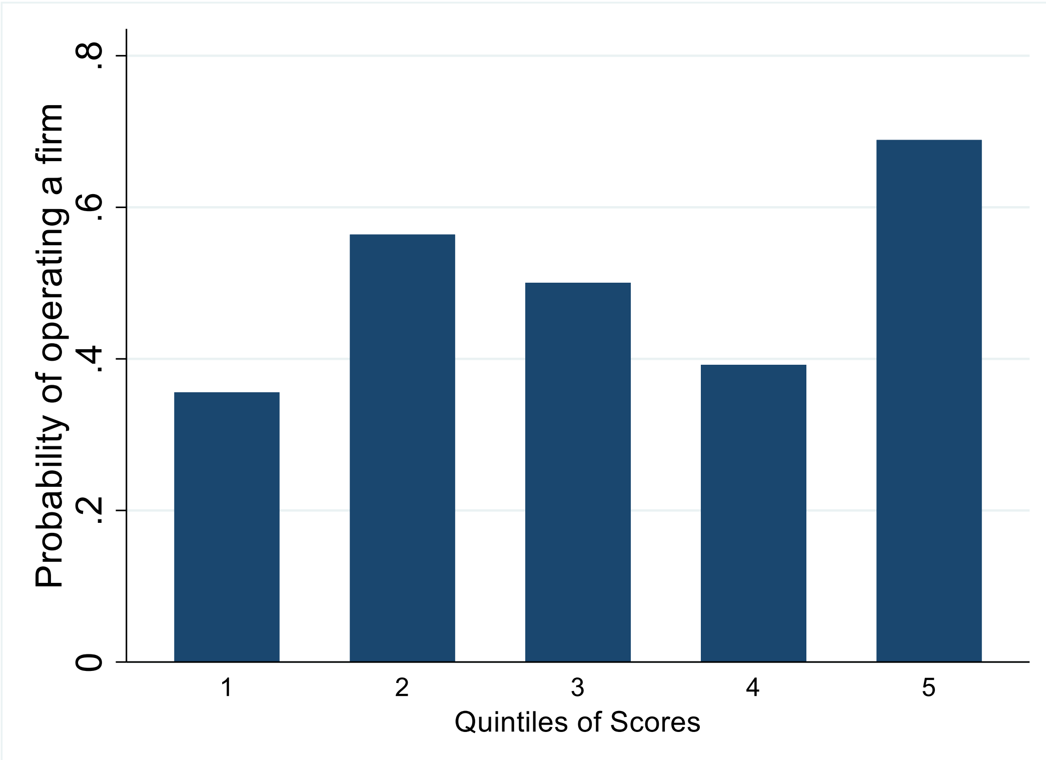


Figure 4.A.8. Aggregate growth index by score quintile (**Bruh**)

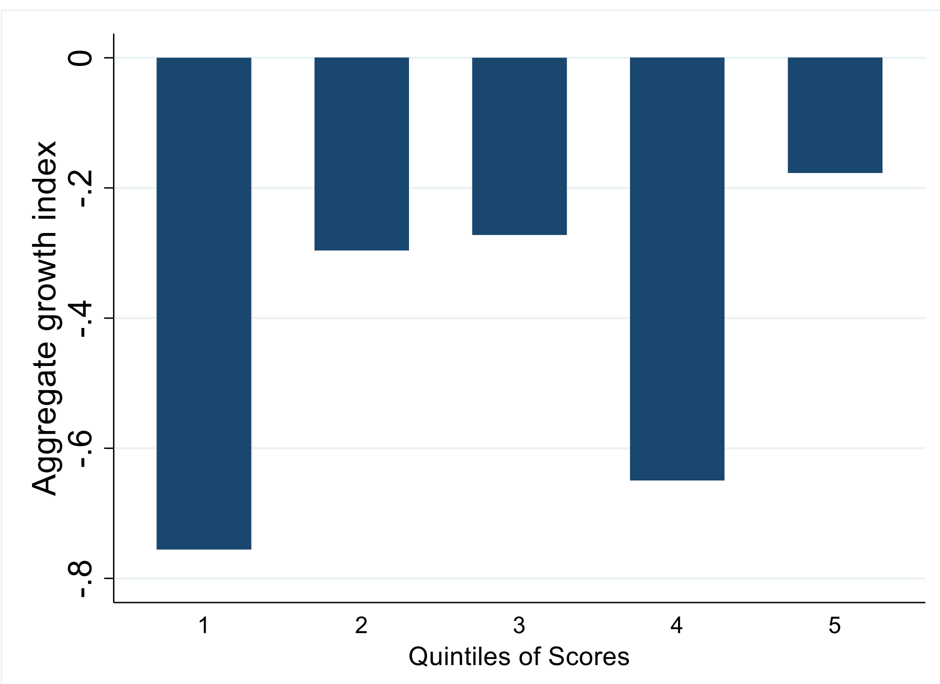


Figure 4.A.8. Aggregate growth index by score quintile (**EDC**)