NBER WORKING PAPER SERIES

LEARNING TO QUIT?
A MULTI-YEAR, MULTI-SITE FIELD EXPERIMENT WITH INNOVATION-DRIVEN ENTREPRENEURS

Esther Bailey
Daniel Fehder
Eric Floyd
Yael Hochberg
Daniel J. Lee

Working Paper 34755
http://www.nber.org/papers/w34755

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2026

We thank conference and seminar participants at Bocconi University, Duke University, Georgia Tech, Harvard Business School, McGill University, Michigan State University, Purdue University, University of Chicago, University of Colorado, University of Mannheim, University of Southern California, University of Utah, University of Washington, the Advances in Field Experiments Conference, and the Southern California Strategy Conference for insightful feedback on both experiment design and/or findings. The Ewing Marion Kauffman Foundation provided generous financial support for this project under Foundation Grant G-201508-20151274. All errors and omissions are our own. Bailey: University of Houston; Fehder: Georgia Institute of Technology; Floyd: University of California at San Diego; Hochberg: Rice University, ECGI & NBER; Lee: University of Delaware. A randomized controlled trials registry entry is available at AEARCTR-0001566 (Fehder et al., 2016). This study was conducted under the oversight of the Rice University Institutional Review Board (IRB). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Learning to Quit? A Multi-Year, Multi-Site Field Experiment with Innovation-Driven Entrepreneurs
Esther Bailey, Daniel Fehder, Eric Floyd, Yael Hochberg, and Daniel J. Lee
NBER Working Paper No. 34755
January 2026
JEL No. C93, D22, M13, M53, O32

ABSTRACT

We use a randomized experiment with 553 science- and technology-based startups in 12 co-working spaces across the US to evaluate the effects of intensive, short-term entrepreneurial training programs on survival and performance for innovation-driven startups. Treated startups are more likely to shut down their businesses and do so sooner than control startups. Conditional on survival, however, treated startups are more likely to raise external funding for their ventures, raise funding faster, and raise more funding than the control group; they also exhibit higher employment and revenue. Treated founders are less likely to found a new startup after shutdown. Our findings are consistent with practitioner arguments that early entrepreneurship training interventions can help entrepreneurs with less viable ventures "rationally quit" ("fail fast"). We use machine learning techniques (causal random forest) to provide exploratory insights on the most impacted subgroups.

Esther Bailey
esther.e.bailey@gmail.com

Daniel Fehder
Georgia Institute of Technology
daniel.fehder@scheller.gatech.edu

Eric Floyd
University of California, San Diego
ejfloyd@ucsd.edu

Yael Hochberg
Rice University
Jones Graduate School of Business
Department of Finance
and Massachusetts Institute of Technology
and also NBER
hochberg@rice.edu

Daniel J. Lee
University of Delaware
djlee@udel.edu

A online appendix is available at http://www.nber.org/data-appendix/w34755
A randomized controlled trials registry entry is available at AEARCTR-0001566 (Fehder et al., 2016).

LEARNING TO QUIT? A MULTI-YEAR, MULTI-SITE FIELD EXPERIMENT WITH INNOVATION-DRIVEN ENTREPRENEURS*

ESTHER E. BAILEY
DANIEL C. FEHDER
ERIC J. FLOYD
YAEL V. HOCHBERG
DANIEL J. LEE

January 8, 2026

We use a randomized experiment with 553 science- and technology-based startups in 12 co-working spaces across the US to evaluate the effects of intensive, short-term entrepreneurial training programs on survival and performance for innovation-driven startups. Treated startups are more likely to shut down their businesses and do so sooner than control startups. Conditional on survival, however, treated startups are more likely to raise external funding for their ventures, raise funding faster, and raise more funding than the control group; they also exhibit higher employment and revenue. Treated founders are less likely to found a new startup after shutdown. Our findings are consistent with practitioner arguments that early entrepreneurship training interventions can help entrepreneurs with less viable ventures "rationally quit" ("fail fast"). We use machine learning techniques (causal random forest) to provide exploratory insights on the most impacted subgroups.

Keywords: Entrepreneurship, Fast Failure, Rational Quitting, Field Experiments, Education

JEL Classification: C93, D22, M113, M53, O32

---

* We thank conference and seminar participants at Bocconi University, Duke University, Georgia Tech, Harvard Business School, McGill University, Michigan State University, Purdue University, University of Chicago, University of Colorado, University of Mannheim, University of Southern California, University of Utah, University of Washington, the Advances in Field Experiments Conference, and the Southern California Strategy Conference for insightful feedback on both experiment design and/or findings. The Ewing Marion Kauffman Foundation provided generous financial support for this project under Foundation Grant G-201508-20151274. All errors and omissions are our own. Bailey: University of Houston; Fehder: Georgia Institute of Technology; Floyd: University of California at San Diego; Hochberg: Rice University, ECGI & NBER; Lee: University of Delaware. A randomized controlled trials registry entry is available at AEARCTR-0001566 (Fehder et al., 2016). This study was conducted under the oversight of the Rice University Institutional Review Board (IRB).

1. INTRODUCTION

Innovation-driven entrepreneurial activity is a major driver of economic growth. Venture capital-backed startups, which typically fall into this category, account for 41% of total US stock market capitalization and 62% of publicly traded companies' research and development spending (Gornall and Strebulaev, 2023). Policies that enable the entry of new innovation-driven entrepreneurs in a region can spur future growth in both venture capital allocation and employment, potentially yielding a virtuous cycle of growth (Hausman 2022; Fehder, Hausman, and Hochberg 2025). As a result, policy makers often undertake significant efforts to encourage and support entrepreneurial activity. At the same time, most new innovation-driven ventures fail. Many innovation-driven entrepreneurs may not have sufficient knowledge to fully understand the entrepreneurial process, the challenges they may face, or how to evaluate the likelihood of success. While post-entry learning and subsequent exit are central to many economic models of entrepreneurship (Jovanovic 1982; Vereshchagina and Hopenhayn 2009), the efficiency of the learning process and its connection to startup performance and exit are less explored (Manso 2016). Increasingly, the stated goals of entrepreneurship training and acceleration programs are twofold: not only to improve and accelerate success for viable ventures, but also to encourage "fast failure" for non-viable ventures or entrepreneurs (Cohen et al. 2019).[1] Whether such programs achieve either of these goals, however, remains an open question. In this paper, we assess the potential for such short-term entrepreneurship training to affect the performance and survival of innovation-driven startups.

From a theoretical perspective, the impacts of entrepreneurship training are ambiguous. A natural null hypothesis is that training should make entrepreneurs "better," which is often interpreted as prolonging startup survival. On the other hand, treated startups could shut down faster if training helps entrepreneurs realize their ventures do not warrant additional investment of time or resources, either because of weaknesses in the venture, product, team, or market, or because the founding team did not wish to continue pursuing an entrepreneurial venture once they had more knowledge of the process and challenges in hand—a form of "rational quitting." From a practical standpoint, identifying a causal impact of training interventions is also challenging. To address causality, random variation is needed, yet there are few settings in which entrepreneurship training

[1] An example of a startup accelerator program is Techstars, whose graduates have raised $25.7B in total funding and created $98.6B in market capitalization. These programs offer short, intensive educational programs to inform founders about the entrepreneurial process, help evaluate ideas and markets, and aid in raising financing. See Hochberg (2016) for a detailed definition and description of such programs.

is randomly or exogenously provided. For example, MBA programs, which often house entrepreneurship curricula, are not filled with randomly selected students, nor are entrepreneurship-specific certificate or executive education programs. Quasi-experiments resulting from random shocks to training are similarly uncommon.

To explore the effects of training on venture success, here, instead, we implement a multi-year field experiment with a large set of operating innovation-driven startups, who were screened ex-ante to have high growth potential. As part of the study, we randomized a training intervention that provides initial coverage of the frameworks contained in an MBA curriculum in innovation-driven entrepreneurship condensed into six four-hour overview sessions conducted over consecutive weeks. The total course time is comparable to the instruction time in a nine-week (i.e., quarter-long) class at business schools. The short term, intensive nature of the curriculum was not chosen ad hoc. Rather, it was identified ex-ante as the most intensive curriculum that could be offered to active entrepreneurs operating tech- and science-based going concerns that would conceivably allow them to find time to participate, and which could be scaled post-experiment, if successful. The training program mimics the overview nature of entrepreneurial training sessions offered as part of many startup accelerator programs and executive education short courses. The goal of the condensed curriculum was to introduce participants to key entrepreneurship-specific knowledge.

Our randomized control trial (RCT) spanned metro areas across the US, partnering with 12 startup incubators and co-working facilities that provide workspace to technology- and science-based startups, and was conducted over a period of four years. We marketed our efforts as the "Entrepreneurial Success Initiative" (ESI). Participating startups were randomized within each participating site. Treated startups received the training program and a $1,000 participation incentive paid to their corporate account. Control group startups received only the participation incentive. We collected extensive information on the startups and their management teams at baseline and conducted periodic surveys at six-month intervals thereafter.[2] We then examine the effect of the entrepreneurship training program across six categories of interim outcomes: survival; pivoting to a new customer segment, business model, or product offering; external fundraising; employment; revenue; and individual founder outcomes post-startup shutdown (employment vs entrepreneurship).

Our experiment reveals a number of intriguing findings. First, treated entrepreneurs are *more* likely to shut down their ventures and do so *sooner* than the control group startups. This increased

[2] Individual founders and management team members participated in the data collection on a voluntary basis under IRB supervision and were free to withdraw from the experiment at any time.

rate of firm closure is not due to positive exits (e.g., acquisitions), which are extremely infrequent. Of course, a startup may, rather than choosing to shut down, instead choose to change either their market, customer, business model, or product offerings, a practice often referred to as "pivoting" (Ries 2011). When we estimate panel models for pivoting, conditional on survival, we find that treated startups are less likely to pivot than those in the control group, suggesting that treated startups may be choosing to shut down rather than experiment with alternatives when their original approach does not prove fruitful.

We next examine a variety of interim measures of performance. First, the ability to raise funding from outside investors, typically venture capitalists or angels, is often used in the financial economics literature as a signal of interim success and quality for innovation-driven ventures. Both unconditionally, and accounting for the competing outcome of startup shutdown via Fine-Gray competing risk models and Cragg double-hurdle estimators, we find that treated startups are more likely to raise external capital, raise capital sooner, and raise much higher amounts of funding than do control group startups by the end of the study period. The magnitudes of these effects are large and economically meaningful. Second, we examine firm employment and revenues. We estimate panel models over the RCT followup periods. Treated firms employ more people and have higher revenue than control firms.

On their own, the survival results could be interpreted in a number of ways. One is as a negative effect: training could harm the startup or entrepreneur and lead to faster shutdown by either discouraging the entrepreneur, wasting their time, or confusing them in some manner. Alternatively, the training may help entrepreneurs better recognize their venture's prospects and the skills and behavioral characteristics required for success, and lead those with poorer prospects or poor suitability for entrepreneurship to shut down sooner. Overall, however, the fact that treated firms raise more funding and have higher employment and revenue suggests that treated firms that do not shut down are, on average, of *higher* quality than surviving control group firms; even accounting for the early shutdown effect, treatment improves the treated startups' ability to raise funds, and treatment leads to higher employment and revenue. Furthermore, the empirical patterns do not appear to be consistent with training merely discouraging entrepreneurs overall; if the mechanism for shutdown is discouragement, it appears to be of startups with poorer prospects, on average.

An additional possibility we consider, given the increased incidence of fundraising in the treatment group, is the possibility that our survival effects are mediated by the increased fundraising in the treatment group, with startups that obtain VC funding encouraged to pursue higher risk paths that lead to either fast failure or large success. While we cannot rule this possibility out entirely, the overall empirical evidence does not appear to be consistent with this alternative mediation channel: the incidence of fundraising among firms that shut down is relatively small (14%), suggesting that fundraising is unlikely to drive the full treatment effect we observe. We additionally find no clear positive association between fundraising and company shutdown—those startups that do not raise funding are considerably more likely to shutdown—and the magnitude of the treatment effect does not diminish substantially when we control for fundraising status. While we cannot completely rule out a modest indirect effect through fundraising, the evidence is most consistent with education having a direct effect on closure, rather than operating primarily through fundraising.

We then turn to post-shutdown labor market choices. If training helps entrepreneurs better understand how to assess the viability of a startup or better evaluate their suitability for entrepreneurial activity, we might expect treated entrepreneurs to change their subsequent labor market choices: they may raise their bar for entry into entrepreneurship in the future or change their evaluation of their expected value from a startup relative to their outside option. Utilizing information on the post-shutdown labor market activities of founders, we show evidence consistent with this hypothesis: conditional on having left their current startup, founders of treated startups are less likely to immediately found another startup.

Since attendance at the training sessions was not mandatory (and could not be made mandatory), our estimates represent "intent to treat" (ITT) rather than "treatment on the treated" (ToT). We utilize attendance data to estimate the TOT. In the first stage, we instrument two measures of attendance (attended at least one session and number of sessions attended) using the randomly assigned treatment. We then utilize those estimates to calculate the ToT effect for survival, fundraising, maximum employment, and post-shutdown labor market choices. As expected, the magnitude of the TOT (attended at least one session) is somewhat larger than the ITT magnitude.

Finally, the overall treatment effect may mask important heterogeneity. Using theory to predict the effects of moderating variables on the treatment in our setting is subject to (at least) three limitations. First, there is an absence of theory that explicitly identifies which variables we would expect to moderate the effects. Second, heterogeneity is not randomly assigned in the same way as treatment status. As a result, it is difficult to use heterogeneous effects to make causal inferences. Third, if the researcher only considers variables to include in this type of analysis ex-ante, they may unintentionally ignore important variation in the data ex post.

To this end, we implement a causal random forest approach to detect heterogeneous treatment effects (Wager and Athey 2018; Athey, Tibshirani, and Wager 2019). The algorithm allows us to examine a large set of potential subgroups and identify patterns in the data that are potentially hidden by the local average treatment effect (LATE). Our analysis suggests some heterogeneity for the treatment effect on startup survival and fundraising. For survival, marital status and optimism both moderate treatment. For fundraising, incorporation status, prior startup experience, and prior P&L experience all appear to be promising moderators. These patterns provide some initial insights for policy makers seeking to target training interventions to appropriate populations. Importantly, this analysis is descriptive, not causal. Future research will be necessary to address potential causal implications.

Our paper contributes to several literatures. First, we contribute to the interdisciplinary topic of "rational quitting." Although many economic models suggest that rational economic agents should quit when their outside options exceed the benefits of their project, few studies have been able to empirically show interventions that can increase the propensity of agents to behave in line with economic theory. "Rational quitting" serves as an important counterpoint to many mainstream dialogues about the unequivocal benefits of perseverance (List 2022). A growing literature has documented how entrepreneurs face substantial information frictions when trying to assess the quality of their business ideas (Lerner and Malmendier 2013; Scott, Shu, and Lubynsky 2019; Bennett and Chatterji 2023). Our study suggests the possible presence of broader information frictions that may impact an entrepreneur's understanding of the growth process, resulting in reduced ability to identify when it is rational for them to quit.

Second, our paper contributes to the literature on the efficacy and effects of targeted interventions for entrepreneurs. The economics literature in this area has primarily focused on interventions in the developing world (e.g. McKenzie 2017; De Mel, McKenzie and Woodruff 2019; Davies et al. 2024), often involving either SMEs (e.g. Bruhn, Karlin, and Schoar 2018) or subsistence entrepreneurs (e.g. Karlin and Valdivia 2011). Recent experimental studies in the management literature span both the developing and developed world, and include the effects of posting non-pecuniary versus pecuniary messaging when recruiting participants to an entrepreneurship program (Sen, Guzman, and Oh, 2018), the effect of communications and networking assistance for small business performance (Kotha et al. 2023), the impact of training on methods of scientific experimentation on entrepreneurial entry and revenues conditional on entry for would-be entrepreneurs across a wide variety of business types (Camuffo et al, 2020), and the effects of availability of peer networking for startup survival and growth (Chatterji et al, 2019).

Our study places specific focus on innovation-driven entrepreneurs running active technology-related businesses with significant growth aspirations—a group of significant interest to policy makers and private funders (Botelho, Fehder and Hochberg 2025). Substantial resources are spent on education for entrepreneurs each year. Our study explores the effect of broad coverage innovation-driven entrepreneurship-specific training on outcomes over a substantial time period. Our findings provide unique insights on the viability of general entrepreneurship training programs in developed innovation ecosystems characterized by high human and financial capital. If entrepreneurship training can reduce inefficient continuation and boost the performance of higher-potential startups, policymakers can potentially harness training interventions to encourage growth and business development. On the other hand, if educating entrepreneurs is mostly ineffective, knowing this could save considerable resources by deterring policymakers from focusing on an intervention that serves little purpose.[3] Our findings suggest that education for particular sets of entrepreneurs may help them better assess the potential for their ventures and prevent inefficient continuation.

Finally, an additional contribution of our study pertains to the setting and population studied. A sizeable proportion of field experiments in economics and business-related fields have been implemented in developing economies. Even for innovation-driven entrepreneurship specific studies, it is unclear whether results obtained from studies conducted in developing countries would generalize from studies to the developed world. Notably, one common challenge to conducting entrepreneurship field experiments in developed countries is finding settings that allow for appropriate population sample sizes with suitable power to detect effects. Here, we achieve this through partnership with practitioner organizations. Our RCT demonstrates that partnerships and ongoing relationships with practitioners can be an important solution for achieving scale in these settings. Addressing our research question of interest would not have been possible within a single tech-oriented co-working facility alone and required the participation and partnership of eleven different organizations (twelve different co-working spaces).

[3] McKenzie (2021) estimates that more than $1 billion is spent on entrepreneurship training each year. For example, the US government spends roughly $40 million on iCorps training yearly for innovation-driven entrepreneurs.

2. BACKGROUND AND EXPERIMENTAL DESIGN

Our analyses represent the first set of results from the RCT registered under AEARCTR-0001566 (Fehder et al., 2016). We recruited innovation-driven startups from 12 startup-focused co-working entrepreneurial spaces in ten metro areas (eleven cities) across the US. Importantly, participating locations did not include cohort-based accelerator programs or other programs that offer financing or structured programming for startups.[4] Rather, we used startup-oriented co-working spaces that at most offered ad hoc workshops.

2.1. Participant Population

Interested startups were screened to confirm that: (i) the startup was innovation-driven (this excludes traditional-type businesses such as food trucks, marketing agencies, development shops, consultancies, retail stores, and other non-tech or non-science-based businesses); (ii) they had not yet raised any institutional funding (firms that had previously raised venture capital or other institutional financing were restricted from participating); and (iii) their founders and top management team members (when not founders) were willing to participate in the field study data collection effort, which included the randomized entrepreneurial training program offered as a participation incentive under the terms of the IRB-approved study.[5]

At the corporate entity level, participating startups signed contracts which paid $1,000 up front to the company legal entity (not the founders) in exchange for an obligation on behalf of the corporate entity to fill out a series of experimental surveys administered by the research team over a multi-year period. The baseline survey for the startup firms includes background information on the company and its activities and performance to date. At the individual level, members of the management teams of the contracted startup companies participated in the research experiment under Institutional Review Board (IRB) supervision, with the option for the individuals to withdraw from the study at any time. Baseline information collected includes demographic, education, and other information such as behavioral characteristics.

Follow-up surveys sent to the participants every six months included questions about the startup and its ongoing performance. The company survey questions tracked standard measures of interim startup success such as funding, employment, team composition, and various corporate practices. Both the baseline and follow-up surveys were designed to be as comprehensive as

[4] For a discussion of the differences between co-working spaces, incubators and accelerators, see Cohen et al. (2019).

[5] Willingness to participate did not require individual members of the team to show up for training if their startup was randomized into treatment but did require them to fill out the baseline individual survey.

possible while also deferring to the realities of participant time constraints and burden of completion.

Figure 1 presents a map of treatment sites, and Table 1 presents further details on participating incubators and co-working spaces. The cities covered in the RCT were deliberately chosen to be a mix of more established startup ecosystems (Boston, Austin, New York City) and newer startup ecosystems (e.g. Chicago, St. Louis, Houston), while co-working spaces were chosen with an eye towards significant tech- and science-based startup membership, but deliberately avoided locations that themselves offered formal training sequences for their lessees. We did not run the experiment at sites in the San Francisco Bay Area due to the extreme prevalence of local entrepreneurship related programming, support organizations, and tech and science founder social networks.

2.2. Intervention

We randomized a condensed version of a standard innovation-driven entrepreneurship (IDE) curriculum to the startups in our sample. This consisted of six sessions of four hours each, delivered at each site over six consecutive weeks. An ex-ante concern is the large set of potential interventions; if we did not find a meaningful treatment effect with a particular intervention, it would not preclude the possibility that other potential treatments could produce a treatment effect. Furthermore, even if we did find a treatment effect, other interventions could produce even more pronounced effects. Given these considerations, our intervention was designed based on feedback from a pilot study. The curriculum was chosen to be long enough to convey a useful amount of knowledge (particularly critical information the inexperienced might not be aware of) but condensed enough that startup teams would participate given extensive demands on their time. Importantly, the modules covered were a condensed overview of typical entrepreneurship content taught in top business school executive education programs and other training programs (such as accelerator programs). A more detailed description of the curriculum is provided in Appendix E.[6]

2.3. Randomization

We randomized startups into treatment and control *within* each site rather than placing some co-working spaces into treatment and some into control. Consequently, each site had startups that were part of both the treatment and control. The disadvantage to this approach is the possibility of spillover from the treatment to control (i.e., the treated group could, in theory, teach the control

[6] The study was not powered or designed to detect the differential impact of the different subject matter addressed in the curriculum; treated entrepreneurs were encouraged to attend all sessions and overwhelmingly did so.

group what they had learned), which would bias against us finding differences.[7] The advantages include increased statistical power and better control for the specific characteristics of any given city or co-working space, which outweighed this disadvantage. Due to capacity constraints on classrooms for training, in most locations, fewer startups were randomized into treatment than into control.

Randomization occurred after startups enrolled in the program, in order to avoid treatment specific selection bias (Floyd and List, 2016). Startups in the treatment group, however, were not contractually obligated to send their management team members to the training sessions; individual participation in the experiment was voluntary. As a result, our average effects represent intent-to-treat. We supplement our analysis with attendance data to better identify the potential effect of the treatment on the treated. A summary schematic of our field experiment design is presented in Figure 2.

2.4. Post-Intervention Tracking and Attrition Considerations

A between-survey period of six months was chosen so that we could track startups as they evolved while simultaneously allowing enough to time to pass for reasonable changes to occur. Survey response rates over time are presented in Table 2. Participation rates for the surveys were high, notably so compared to most prior survey work. For example, at the first follow-up survey, our lowest percentage of respondents was 82.1% (Los Angeles) while our highest was 100% (Saint Louis and Houston). To achieve this, we made use of the claw-back provision in the spirit of Fryer et al. (2022), which allowed us to demand reimbursement for up to the full $1,000 if startups did not fulfill their contractual survey obligations.

Survey participation rates in the full sample fall over time, as startups attrite from the sample due to failure. Conditional on survival, response rates at the fourth follow-up (24 months) still reached approximately 80%. When looking at attrition in our data, what is particularly relevant for our empirical analysis is not whether a firm dropped out of our analysis because the firm did not survive (which is a study outcome variable we specifically model and analyze), but rather whether the firm continued to operate but chose not to respond to our survey—in other words, whether there was a differential ability to collect data across the two groups. Attrition concerns apply only to variables collected from surveys. They do not apply to our outcome variables that are determined from outside sources such as funding or survival.

[7] In untabulated analysis, we examine survey data on which companies our participants interact with at their co-working facilities, and find no evidence that spillovers from treatment to control are cause for concern.

Online Appendix A4 provides a detailed description of our differential attrition analysis. At the individual survey period level, we find a statistically significant difference between treatment and control response rates only for the first follow-up period. When we test for overall difference in response rates, the analysis suggests that treated startups were approximately 5.3% less likely to respond to the followup surveys than control group startups. Given this differential attrition, we follow McKenzie (2017) for our survey-level variables (pivoting, employment, and revenue) and assess the robustness of our results to differences across attritted and non-attritted startups, following the procedures describe in Online Appendix A4, and provide Lee (2009) bounds and Imbens and Manski (2004) bounds. Overall, our conclusions remain similar to those in the main analysis below.

The quality of responses to the survey questions, conditional on filling out the survey, exhibited more variability, which was expected. Certain types of written answers required cleaning to ensure that the answers recorded are logical and consistent.

2.5. Timing and Sample

A detailed description of the structure and timing of the RCT across sites is presented in Online Appendix A1, Online Appendix Figure A1, and Online Appendix Table A1. We discuss the procedure for processing surveys and external data in Online Appendix A2, Online Appendix Tables A2.1 and A2.2, and Online Appendix Figure A2. While the total length of the experiment conduct was four years, the actual intervention was rolled-out sequentially over a two-year period, site by site. The resulting dataset is a panel of 553 firms over four calendar years, which in total represents the time from the intervention start at the first sample site until the final survey was collected at the last sample site. For each individual site, the time from the end of the administration of the intervention until the final survey was collected is two years.

3. DATA

The experiment resulted in an extensive set of data. Given the scope of our analyses in this paper, we focus on only a subset. Online Appendix A3 provides a more detailed description of the construction of our major variables.

3.1. Outcome variables

Because our startups were only contracted to a two-year follow-up reporting period, and due to the often-extended time period it takes to exit an innovation-driven startup, we necessarily cannot follow all firms through to their end outcome. Instead, we focus on interim performance measures indicative of startup progress and performance. For some variables, we can supplement data collection beyond the experimental survey instruments, as described below, allowing us to extend our analysis period outside of the two-year follow-up period, while also alleviating concerns about non-reporting or attrition. We utilize six categories of outcome variables, computed either from survey data or from commercial databases that allow us to track participants without concerns about attrition.

3.1.1. Survival/Shutdown

The first of these categories is firm survival. A startup is coded as shut down either based on survey response or given an external signal of closure such as the first point in time that every (former) employee of the firm on LinkedIn has a job somewhere other than the startup. A full description of the tracking and coding process is provided in Online Appendix A2. The use of outside indicators for shutdown allows us to track survival for an additional year beyond the two year contractual RCT period.

For startups that indicated that they shut down in a followup survey, we ask for the primary reason for closure. The survey provides 7 possible reasons: (1) no market need for the product; (2) personal reasons; (3) better outside opportunity; (4) bankruptcy; (5) acquisition; (6) Other; and (7) prefer not to answer. 174 startups indicated shutdown on their follow-up surveys; of these, 123 provided a specific reason from the list we had provided, 49 answered “Other,” and two did not respond. For the “Other” category, we ask for a free text description of the reason. We then use the free text responses to further categorize the closure reasons (note that a startup could provide multiple reasons in their free text response). Ultimately, 18 gave an answer that cited lack of financing, 17 indicated that they were not reaching growth milestones (either market or technical), eight had founders who had elected to move on to another startup, six cited founder disagreements, five cited limited demand, and three cited new entrant competition. We separately code startups that experience a “positive” exit in the form of an acquisition. Only 7 of the startups in our sample report experiencing a merger or acquisition during the tracking period. We found no further indications of acquisition exit (of either company or assets) and no other responses in the “Other” category that could plausibly be considered positive exit.

In addition to startups that shutdown and explicitly reported this to us when we sent the next period’s tracking survey, an additional 109 startups shutdown during the study period as determined by our shut down tracking procedure described in Online Appendix A2. When we manually searched the entries for these startups on Pitchbook, we found no evidence of additional acquisition exits. All in all, a total of 283 startups were determined to be shutdown by the end of the study period. Startups that shutdown due to acquisition are treated as a positive exit and treated as survivors in our survival analysis. We construct an indicator for whether a startup has shutdown for each followup period. We also construct a measure of time to closure for a startup, defined as the total number of days from the date in which the startup enrolled in our study to the date in which the startup was no longer operating. Finally, we construct a simple indicator of whether the startup shut down at any point during the timespan of the RCT.

3.1.2. Pivots

Our second category of outcome variable is pivots. Pivots, as coined by Ries (2011), are major changes to either the customer, product, or revenue/distribution channel. In each followup period, startups provided an elevator pitch and their major goals for the next six months. We use these qualitative answers and follow the definition of pivot used in Camuffo et al (2020) to code whether a startup had a major change to either the customer, product, or revenue/distribution channel for each followup period, as described in Online Appendix A2. For each followup period, we construct an indicator as to whether the startup survey response indicates a pivot took place in their business.

3.1.3. Fundraising

The third category of outcome variable is funding. In order to ensure a standardized measure of fundraising for all participants without attrition concerns, we utilize data on fundraising from Pitchbook. We track the startups for an additional year beyond the contractual two-year RCT period for this analysis, for a total of three years of followup fundraising data. We define five variables: an indicator for whether the company raised capital in a given followup period, an indicator as to whether the startup raised capital by the end of three years, funding amount raised in each followup period, total funding raised by the end of three years, and time (days) to first fundraising post intervention. We utilize the natural logarithm of one plus the dollar amounts in our analysis (results are robust to the use of inverse hyperbolic sine instead as shown in Online Appendix B; we report the results using the log 1+ transform herein as they make interpretation of coefficients straightforward).

3.1.4. Employment

Our fourth category is startup employment. We use survey answers to determine the total number of startup non-founder employees in each followup period. Participating startups were asked to report the number of founders still present as well as the number of full-time and part-time non-founder employees of the firm.

3.1.5. Revenue

The next category of outcome variable is startup revenue, obtained from survey responses. Our followup survey asks the startup for quarterly revenue for the last quarter. Participants were not forced by the survey instrument to report revenue; some chose not to and skipped the answer. As a result, our unbalanced panel does not include all firms in all periods they were still operating. We use the natural logarithm of one plus quarterly revenue reported in each followup period (results are robust to the use of inverse hyperbolic sine instead as shown in Online Appendix B; we report the results using the log 1+ transform herein as they make interpretation of coefficients straightforward).

3.1.6. Post-shutdown Founder Labor Choices

Our last category explores founder employment choices after they leave their startup. We utilize Linkedin to determine whether founders launched another startup immediately after leaving the ESI-enrolled startup, as described in Online Appendix A2, Section A2.4. We construct an indicator variable for whether each founder of a shutdown startup identifies as a founder or co-founder of a new startup in their next job role.

3.2. Behavioral, Demographic, and Company Variables

We include a number of additional variables. These include basic behavioral and demographic characteristics as well as corporate characteristics. The data for these additional variables is taken directly from our surveys. Our set of covariates is motivated by ex ante theoretical considerations of what might be expected to potentially affect startup performance, as well as the volume of past research that suggests the possibility of correlation between certain variables and startup outcomes, decision-making, or performance.

3.2.1. Individual Behavioral Parameters

A large literature has connected founder and manager behavioral characteristics to firm performance and managerial decision making (see e.g. Kerr, Kerr, and Dalton, 2019). We control for three individual founder behavioral parameters including optimism (measured using the LOT-R scale), time preference (measured as in Loewenstein, Read, and Baumeister (2003)), and risk preference (measured using Barsky et al. (1997) as implemented by Graham et al (2013)). Because we measure behavioral parameters at the founder/individual team member level, we must first determine how to aggregate parameters to arrive at a single value for the startup. For simplicity, we use the maximum of each behavioral characteristic across the startup management team. Risk preferences are measured on a scale of 1 to 4, following Barsky et al. (1997), with a 4 being most risk loving. Our firms have a mean of 3.60 compared to a mean of 1.72 in Barsky et al. (1997), indicating that our sample of entrepreneurs is relatively more risk loving than average, consistent with prior research. Similarly, the max optimism score of almost 20 indicates that our sample exhibits high optimism relative to the mean of 14.33 reported for the standard laboratory population used to develop our measure of optimism (LOT-R, Scheier, Carver, and Bridges 1994).

3.2.2. Demographic Variables

We use answers to the individual founder surveys at baseline to construct a number of variables reflecting the demographic characteristics of the founder team. Our set of covariates are motivated by the substantial literature in finance, economics, and management that relates founder demographics to differential funding and performance (see e.g. Blanchflower, Levine, and Zimmerman (2003), Fairlie and Robb (2007), Ewens and Townsend (2020), Azoulay et al. (2020), Fairlie, Robb, and Robinson (2022)). We construct an indicator for whether the startup has at least one founder with a child, whether at least one founder is married, whether at least one founder is female, whether at least one founder comes from an underrepresented minority (defined as BIPOC—Black, Indigenous, People of Color). We also construct indicators for whether at least one founder has managerial experience with direct responsibility for P&L (profit and loss), whether at least one founder has prior experience working at a startup, whether at least one founder has a STEM graduate or undergraduate degree, whether at least one founder holds a business degree, whether at least one founder holds a degree from a top 10 (top 11-20) university based on the 2016 US News and World Report ranking, and the average age of the startup's founders at baseline.

3.2.3. Company Characteristics

A number of corporate characteristics are thought to be correlated with startup funding and outcomes. We utilize a number company characteristics as covariates, all measured at baseline. First, we construct an indicator for whether the company is registered as a C corporation (e.g. Guzman and Stern (2020)). Second, we construct an indicator for whether the startup was incorporated in Delaware (Guzman and Stern (2020)). We also categorize the startup industry using the primary industry they report in their baseline survey (education, finance, healthcare/biotechnology, e-commerce, IT and enterprise software, media, mobile and wireless, natural resources/energy, social enterprise, social media/networking, tourism, other), and the stage of development of the startup using their survey responses (concept, working prototype in development, functional product with users, scaling sales). Finally, we construct measures of our performance outcome variables as measured at baseline, pre-intervention: the startup's quarterly revenue in the last quarter at baseline, startup employment at baseline (full-time and part-time non-founder employees), and the amount of funding the startup had already raised at baseline.

3.3. Summary Statistics and Sample Balance

Table 3 presents the overall summary statistics for the treatment and control sample at baseline, and Table 4 presents summary statistics separated by treatment and control along with a test of balance. The first column of Table 4 presents the mean and standard deviation (in parentheses) of the baseline covariates of the startups in the control group. The second column shows the mean and standard deviation (in parentheses) of the baseline covariates of the startups in the treatment group. The third column shows the mean and standard deviation (in parentheses) of the baseline covariates of all the startups in the sample.

To test for sample balance, we estimate separate regressions for each covariate of the form $Treat = \beta_0 + \beta_1\ Covar + \epsilon$. In the column, we show the estimates of β_1 for each covariate. Randomization appears to have achieved balance across treatment and control samples. Of the full set of baseline variables presented, only C Corporation status, Business Education, and one industry category (finance) demonstrate statistically significant differences between the treatment and control groups. The existence of some differences is to be expected given the dimensionality.

We also compute an omnibus F-Test first following the procedure in McKenzie (2017) and second following Cattaneo et al. (2018). The joint orthogonality test cannot reject balance on baseline characteristics for the sample in either approach (F=1.17, p=0.24; F=1.26, p=0.16). Despite this, to account for any residual differences post-randomization, we control for both unbalanced baseline variables and the full set of baseline variables in our specifications.

4. MAIN RESULTS

4.1. Survival

A natural null hypothesis is that training should make entrepreneurs "better," thus prolonging their startups' survival. On the other hand, treated startups could shut down faster because training encourages a form of "rational quitting." Unsurprisingly, a large number of startups shut down over the study period. Table 5 reports the raw percentage of startups still in business at different time periods.[8] Overall survival at three years post intake is 49%. At the three-year mark, 56% of control firms are still operating, while only 41% of treated firms remain in business. This difference is economically large.

In the raw survival data, treatment startups shut down operations approximately 99 days sooner on average than control firms. Figure 3 Panel A shows Kaplan-Meier plots for survival. It is visually apparent that the difference in survival rates between treatment and control in our sample manifests relatively early—within approximately one year. A Wilcoxon rank-sum test confirms the treatment and control groups do not have the same survival functions ($p<0.0001$). Survival in our analysis is restricted to the narrow window (3 years) in which we observe startup outcomes.

Kaplan-Meier survival curve analysis does not allow for the analysis of survival differences conditional on control variables. In Table 6, Panel A, we present estimates from linear probability panel regression models of the type:

$$Y_{it} = \alpha + \gamma_s + \varphi_t + X_i'\beta + \epsilon_{it} \qquad (1)$$

where α is a constant, Y_{it} is an indicator for whether the company i shut down during followup period t, γ_s is a site-level fixed effect, φ_t is a followup period fixed effect, and X_i is a vector of covariates measured at baseline. We first estimate a specification that has only site and followup period fixed effects. We additionally estimate a specification that adds only unbalanced covariates and a specification that adds all baseline covariates. Standard errors are clustered at the startup level. In the panel models, treated firms exhibit an increased probability of closure of 9.4 to 10.5 percentage points.

We next estimate a series of Cox proportional hazards models for survival (time to shutdown), where the independent variable of interest is treatment. We estimate variations on the following model:

$$h(t|x) = h_0 \exp\left(\beta_1 treated + \gamma_s + X_i'\beta\right) \qquad (2)$$

where h_0, the baseline hazard rate, is estimated non-parametrically, γ_s is a site-level fixed effect, and X_i is a vector of covariates. The exponentiated coefficient, e^{β_1}, provides an estimate of the impact of treatment with an intuitive interpretation as a hazard rate ratio. To illustrate interpretation, a hazard ratio of 0.5 represents a 50% reduction in the hazard of shutdown, whereas a coefficient of 1.5 would correspond to a 50% increase. Throughout the paper, we show exponentiated coefficients.

[8] As noted in Section 3.1.1., startups that were acquired are categorized as survivors for this analysis, as acquisition is generally considered to be a "positive" exit rather than a (negative) shut down. Our results are robust to categorizing such startups as shutdowns instead.

Panel B of Table 6 presents the estimates. Treatment exhibits an encouraging effect on startup shutdown. The exponentiated coefficients (hazard ratio) on the treatment indicator are positive and significant across all models. The estimated magnitudes are fairly stable across specifications and suggest that treatment increases the hazard of shutdown by 51-58%.

4.2. Pivoting

Rather than quit, a startup that feels its current approach is not viable may instead choose to change their business, adjusting either their market, customer, business model, or product offerings to pursue a different path, a practice called pivoting (Ries 2011; Camuffo et al., 2020). To examine pivoting, we estimate linear probability panel models similar to those we estimate for survival. The outcome variable is an indicator for whether the startup conducted a pivot in that period. The panel is unbalanced, and observation inclusion is conditional on not having shutdown or exited via acquisition.

The estimates are presented in Table 7. Column (1) includes only site and followup-period fixed effects. Column (2) adds unbalanced covariates as controls, while column (3) adds all covariates as controls. Standard errors are clustered at the startup level. Treatment is associated with a statistically significant decrease in pivoting in each period of between 1.9 and 2.6 percentage points. The estimates are consistent with the notion that when faced with challenges, treated startups are more likely to decide that further pursuit of their startup is not worthwhile, and simply shut down, while control startups try to find other paths forward rather than admit defeat. The training program provides treated founders with clear guidance as to what a prospective startup growth path would need to look like for the startup to have a chance to succeed. Treated founders may view the opportunity costs as too high to justify attempting to find alternative paths, and shut down rather than pivot.

4.3. Fundraising

Our evidence thus far is consistent with our hypothesis that treatment encourages "rational quitting." To better elucidate the mechanism driving the observed treatment effect, we next turn to measures of startup performance. Fundraising from external investors is considered a key interim performance measure for innovation-driven startups because external financing is particularly important for these types of entrepreneurs. Because innovation-driven startups are attempting to bring novel products and services into the economy, they require investment of financial capital in advance of receiving significant revenue or even entering the market (Botelho, Fehder, and Hochberg 2025). Here, we focus on a variety of measures: whether the startup raised capital in any given survey followup period, how much they raised in any given followup period, time to first funding round, whether the startup managed to raise any funding by endline, and the logged total amount of funding raised over the study period.

To set the stage, Figure 3 Panel B shows the Kaplan-Meier plots of time to first financing for treatment and control startups. Treated startups acquire their first funding round more quickly and in higher proportion. A log-rank test of the difference in the survival curves between the two groups is statistically significant ($p < 0.01$).

We then begin our formal analysis by estimating linear regression panel models where the dependent variable is an indicator for whether or not the startup has raised external funding in a given followup period. The estimates from our panel models are presented in the first three columns of Panel A of Table 8. We estimate similar specifications to those for survival. Across all three specifications, we observe a positive and statistically significant coefficient on the treatment indicator, with magnitudes suggesting that treatment leads to a 2.3 to 2.7 percentage point increase in the probability of raising external funding in any given followup period. We then estimate similar panel models where the dependent variable is the natural logarithm of one plus the amount of external capital raised by the startup in a given followup period (results are robust to the use of the inverse hyperbolic sine transformation instead – See Appendix B, Table B1.1; we present the logarithmic transform estimates for ease of coefficient interpretation). The estimates are presented in columns (4)-(6). We find a positive and statistically significant coefficient on the treatment indicator across all models. The magnitude of the coefficients suggest that the treatment affect leads to a 34-41% increase in per follow-up period funding.

Next, we analyze the impact of treatment on the hazard of a fundraising using a Cox Proportional Hazard model. The estimates are presented in Panel B of Table 8. Columns (1)-(3) present estimates from the Cox Proportional Hazard models. We find that treatment is associated with a statistically significant 63 to 81% increase in the hazard of a startup receiving funding across the different specifications.

Of course, a startup's choice to shutdown (or its acquisition and absorption into a buyer) might preclude us from observing a fundraising event, creating competing risks. To address this issue, we estimate Fine-Gray models of competing risks. In these models, we consider the risk of our primary event, fundraising, while also considering the competing risk of firm shutdown or acquisition (which induces exit from the risk pool). The subdistribution hazards of both events are considered while estimating the cumulative incidence function of the primary event using the following model:

$$1 - CIF(t) = \left(1 - CIF_0(t)\right)\exp\left(\beta_1 treated + \gamma_s + X_i'\beta\right) \quad (3)$$

where CIF_0, the baseline cumulative incidence function, is estimated non-parametrically, γ_s is a site-level fixed effect, and X_i is a vector of covariates that varies by model. The exponentiated coefficient, e^{β_1}, provides an intuitive interpretation as a relative change in the subdistribution hazard function for the primary event—fundraising. Here, an exponentiated coefficient of 1.5 can be interpreted as a 50% increase in the risk of the primary event in subjects that remain event free or experienced a competing event (Austin and Fine 2017).

Columns (4)-(6) present estimates from the Fine-Gray Competing Risk models. Here, the estimates suggest a statistically significant 62 to 80% increase in the risk of raising funds after accounting for the competing risk of startup shutdown. Overall, the estimates indicate that treatment significantly impacted the speed at which treated startups were able to receive funding.

We next look at startup fundraising status at endline. First, we estimate simple unconditional models of the following form:

$$Y_i = \beta_0 + \beta_1 treated_i + \gamma_s + X_i'\beta + \epsilon_{is} \quad (4)$$

where Y_i in this regression is either an indicator variable set to one if the startup received any funding by endline and a zero otherwise ($I\{Raise\}$) or the natural logarithm of one plus the total dollars raised over the course of the study period. γ_s are site fixed effects and X_i' is a vector of additional covariates that vary by model. The standard errors in our model are clustered by site. For the indicator variables, we estimate a linear probability model (LPM) which provides a simple interpretation of β_1 as a percentage point change in the probability of receiving funding. We show robustness of our results to alternative logit specifications in Online Appendix Table B1.1 of Online Appendix B.

The estimates for the linear probability models are provided in Columns (1)-(3) of Panel C of Table 8. The estimates of the impact of treatment are a statistically significant 9.6 to 10.6 percentage point increase in the likelihood that a treated startup ever raises funding. As compared to the unconditional mean fundraising probability of 23.3%, this represents a roughly 41 to 45% increase in incidence of fundraising by endline.

Next, we analyze the amount of funding raised by endline. To do so, we use the natural logarithm of one plus the total dollars raised post enrollment as the dependent variable ($Logged\ Funding$). The results of our analysis are provided in Columns (4)-(6) of Panel C of Table 8. Across all the models, the estimates suggest a positive and statistically significant impact of treatment on the total amount of post-enrollment funding. Depending upon the specification, the estimates suggest that treated startups raise approximately 288 to 346% more funding than control startups. Our results are robust to an alternative transformation using the inverse hyperbolic sine (Table B1.1 of Online Appendix B). If training equally increases the chance that startups close earlier irrespective of quality, we would expect to see decreases in funding rates and total funding in treated startups. This is separate from the human capital channel for our educational program which might increase the quality of the surviving startups and thus both the likelihood of raising and the total amount of funding raised.

While the results above show large differences between treatment and control in amounts raised, they do not account fully for several features of the data generating process. First, they do not account for the extensive and intensive margins for raising funds. Second, they do not account for the relationship between survival and fundraising. To address these issues, we jointly estimate the extensive and intensive margin of treatment while incorporating survival time in a double hurdle model (Cragg 1971). The double hurdle model jointly estimates the probability of raising zero funding and the total amount of funding conditional on being a non-zero observation. We follow Cragg (1971) because the model provides a more flexible functional form for estimating the extensive margin relative to other truncated or censored regression models (Tobin 1958; Heckman 1976; Powell 1986).

The estimates are presented in Panel D of Table 8. Across all specifications, treatment has a significant positive effect on the likelihood of raising funds on the extensive margin after also accounting for survival time. As expected, survival time influences both the extensive and intensive margins; that is, the longer a firm survives, the more likely they are to raise funds, and the more funding they raise. Across most specifications, treatment exhibits a positive and statistically significant impact on both the extensive and intensive margins. Coefficient estimates for the effect of treatment on funding raised are somewhat larger than in the unconditional models.

Figure 4 presents a kernel density plot of the distribution of funding for treatment and control. A Kolmogorov-Smirnov test of the equality of the distributions rejects the null of the same distribution of funding (p=0.023). The figure qualitatively demonstrates that the control group has a longer left tail of smaller fundraising outcomes. Thus, it appears that startups that raise low levels of funding are "missing" in the treatment group.

4.4. Employment

To further reinforce our conclusions, we turn next to an additional performance metric, non-founder employment. Table 9 presents estimate from a set of Panel models similar to those run for survival, pivoting, and fundraising. Across all three specifications, the coefficients are positive and significant, indicating higher employment levels for the treated startups relative to the controls. The magnitude of the difference is approximately 1.6 additional employee per period for treated startups, relative to an unconditional mean of 3.91 employees, or a 41% increase relative to controls. These results lend support to the conclusion that treated startups perform better than controls in their ability to grow and pursue their objectives.

4.5. Revenue

Next, we examine startup quarterly revenue, another proxy for startup performance. Table 10 presents estimate from a set of panel regression models similar to those run for survival, pivoting, fundraising, and employment, where the outcome variable is the natural logarithm of one plus the startups reported revenue for the followup period's last reported quarter. Results are robust to the use of the inverse hyperbolic sine transform instead (see Online Appendix B). Because not all startups reported revenue numbers, the panel includes only firms reporting revenue numbers. Across all three specifications, the coefficients are significant and indicate higher quarterly revenue levels for the treated relative to the controls. Treated startups exhibit between 74 and 87% higher quarterly revenue than their control group counterparts. These results lend further support to the conclusion that treated startups perform better than controls in their ability to grow and pursue their objectives.

4.6. Direct Effects? Or Mediation from Fundraising?

An additional possibility we consider, given the increased incidence of fundraising in the treatment group, is the possibility that our survival effects are mediated by the increased fundraising in the treatment group, with startups that obtain VC funding encouraged to pursue higher risk paths that lead to either fast failure or large success. We undertake a number of analyses to evaluate this possibility. While it is the case that treated startups that shut down raised more funding than control startups that shut down (about $240K for treated shutdowns versus $125K for control shutdowns), a number of other pieces of evidence suggest that mediation of this type is unlikely to be driving the main effect we document.

First, we examine the incidence of fundraising among firms that shut down. We find that the incidence of fundraising among negative outcome firms is relatively small (14%), which makes it unlikely that fundraising is unlikely to drive the full treatment effect we observe. The low incidence of external financing among shutdown ventures suggests that the training/education direct effect operates largely independently of any fundraising-induced growth or death, and makes it unlikely that funding explains the full treatment effect we observe. Second, we do not observe a clear association between fundraising and closure when we define it as (non-acquisition) shutdown by the end of the sample period. In fact, if anything, it appears that those that do not raise funding are more likely to shut down by the end of our sample period in both treatment and control: 68% of treated firms that did not raise funding shutdown by endline, versus 40% of those that did raise; in the control group, 49% of those that did not raise capital shutdown by endline, versus only 20% of those control group startups that did raise capital. Third, in unreported results (available upon request), we estimate a linear probability model regressing an indicator for shutdown on treatment status, fundraising status, and amount raised. The inclusion of the indicator for fundraising or the dollar amount raised in the regressions does not substantially shrink the main coefficient of treatment on closure. If we also include interactions of the fundraising variables with treatment, the coefficients on the interaction term are close to zero and not statistically significant. This is consistent with a direct treatment effect that goes beyond fundraising status.[9] Finally, when we re-run our main survival analysis in Table 6, restricting the sample solely to companies that never raised financing during the sample period, the hazard of shutdown for treated firms becomes slightly larger, and remains highly statistically significant, further supporting the notion that our results stem primarily from an effect separate from any VC "swing for the fences" risk-taking push. Thus, while we cannot completely rule out a modest indirect effect through fundraising, the data suggest that much of the shutdown effect occurs directly from the educational treatment.

4.7. Post-Shutdown Entrepreneurial Choices

Thus far, our results have demonstrated that treatment causes startups to shut down earlier, but, conditional on survival, leaves startups of higher quality. The empirical patterns are consistent with the notion that this is a rational effect and is inconsistent with the notion that the training intervention made treated entrepreneurs "worse" or simply discouraged them. Rather, the results suggest that treated firms who do not shut down are of higher, rather than lower quality. Our hypothesis is that treated startups shutdown more quickly because founders have a better understanding of the quality level required to grow and succeed. A natural follow up question is whether there are longer term changes in how these startup founders choose to engage in entrepreneurship after the shutdown of their current startup. In general, founders frequently engage in serial entrepreneurship founding new businesses when one fails. The increased understanding of the hurdles associated with growing an innovation-driven startup that is provided by treatment

[9] Because fundraising is itself affected by treatment, conditioning on funding status does not permit a clean causal decomposition; these results should therefore be interpreted as descriptive evidence rather than formal mediation tests.

should spillover into their assessment of future ventures. Thus, training may not only lead founders to "rationally quit," it may also lead them to evaluate potential new startup opportunities more stringently. If this were so, we would expect fewer treated startup founders to jump into entrepreneurship again immediately after shutting down their business.

To investigate this, we collect data on founder employment post-shutdown from LinkedIn. Data is collected after the conclusion of the experiment across all sites, and the cessation of follow-on surveying. We code whether founders of shutdown startups in our experiment start a new startup following the termination of their original experiment startup. We then estimate a series of linear probability models similar in structure to equation (3). Our dependent variable measuring the immediate entrepreneurship choices of founders after shutdown of an ESI enrolled startup.

Table 11 presents the estimates from our analysis. Treated entrepreneurs have a statistically significant reduction of 8.5 to 9.5 percentage points in the likelihood of pursuing entrepreneurship again immediately after shutdown of their startup. Our results support the idea that the training not only impacts decisions about the current startup, but also how startup founders weigh the costs and benefits of entrepreneurship more broadly.

4.8. Treatment on Treated

Thus far, all of our analyses measure the effects of intention-to-treat. Attendance in the educational program is not mandatory, and not all teams chose to attend. Startups were randomized into treatment, but ultimately there was no mechanism to force treated individuals to take the class. First, as in all human subject studies, individual participants (founders and management team members) retained the right to cease participation at will. While the possibility existed of incentivizing them to attend sessions via a claw-back mechanism that conditioned on the education program, we decided against this approach. A primary consideration in this design choice is that, if our experiment is to inform policy, understanding voluntary participation rates in the program are likely as important as how effective the training is conditional on startups participating.

While the ITT estimates of our educational intervention are the most policy-relevant, we provide additional TOT estimates for several of our key dependent variables. To do this, we use the randomized assignment to treatment as an instrument to predict various measures of attendance in the ESI classes. We draw our attendance data from RSVPs and ex post measures of attendance. Our two measures of attendance are the total number of sessions attended by the startup and whether they attended at least one session. We then use the instrumented attendance variables to estimate TOT.

The results of our analysis are presented in Table 12. In Panel A, we provide our estimates of the first-stage relationship between treatment assignment and two measures of attendance levels. As expected, there is a large and statistically significant relationship between treatment and attendance. The average first-stage F statistics range between approximately 200 and 300, well above the thresholds to ensure minimal bias (Stock and Yogo 2005).

In Panel B, we provide the results of the instrumental variable analysis for four key dependent variables measured at endline: (1) time to shut down, (2) whether the startup raised any funding by experiment endline, (3) logged dollar funding amount raised by end of the study, and (4) whether a startup founder immediately started another company after shutdown. Our estimates of TOT for the program are consistent with the ITT findings. In Panel C, we provide the estimates from the second stage for our panel model outcome variables: (1) whether the startup pivoted in a given followup period; (2) number of employees in each followup period; and (3) logged quarterly revenue in each followup period. Our estimates of TOT for the program again are consistent with the ITT findings. Appendix B shows robustness to the use of inverse hyperbolic sine for the logged variables.

4.9. Treatment Heterogeneity

The exploration of heterogeneity for this study is critical. While our education program is similar to programs at major business schools and entrepreneurship training programs provided by accelerator programs, the program itself is reasonably expensive in terms of both founder time and cost of instructors. We do not expect that all entrepreneurs in our study will benefit equally from the training, but it is not clear which entrepreneurs, as identified by their observable characteristics, are most likely to improve because of treatment. Understanding which characteristics are most important can help policymakers better target the education program towards those that would most benefit.

Without predictions from theory, it is not clear which characteristics are likely to provide treatment heterogeneity. There are enough baseline covariates in our sample data that simply including linear interaction terms into our regressions may provide spurious findings of treatment heterogeneity. In the time that has passed since we initially pre-registered the exploration of treatment heterogeneity in our study, advances in post-processing of RCT data, especially through machine learning (ML) tools, have been considerable. They now provide options for assessing treatment heterogeneity across many covariates in a rigorous manner that does not require a theoretical perspective ex ante.

A number of methods have emerged building upon the tools of machine learning to provide robust estimation of treatment heterogeneity (Chernozhukov et al. 2018; Athey, Tibshirani, and Wager 2019). We focus on the causal random forest because it allows efficient estimation of treatment heterogeneity in settings with smaller sample sizes.[10] The causal random forest approach captures complex, non-linear relationships between covariates, including the treatment variable, and outcomes that might be missed by exploring heterogeneity using OLS regressions with linear interaction terms. A number of existing papers provide useful primers on the underlying theory behind causal random forests (Knittel and Stolper 2019; Davis and Heller 2020). Importantly, we consider the following analysis of heterogeneity to be a descriptive exercise, rather than a test of theory.

Assessment of treatment heterogeneity proceeds after first assessing the degree to which the ML prediction algorithms have adequately captured both the overall average treatment effect (ATE) and the systematic deviations from the average treatment effect for different subgroups, the conditional average treatment effect (CATE). For the sake of brevity, we document these "pre-steps" in Online Appendix C and focus on the outputs of the heterogeneity analysis. In particular, Chernozhukov et al. (2018) propose a test of the model's ability to fit ATE and heterogeneity. The test results provided in Online Appendix Table C1 show that our causal random forests provide good estimates of both the ATE and systematic deviations from it for survival and funding. In addition, Online Appendix Figure C1 indicates that there is significant heterogeneity in the treatment effects for these two outcome variables. The key output of these pre-steps for this method is a validated non-linear model of the CATE which produces a separate estimate of CATE for all unique combination of baseline covariates (see Athey et al., 2019 for further description).

To begin our exploration of the relationship between treatment heterogeneity and baseline covariates, we estimate the best linear projection (BLP) of our baseline covariates onto our estimated CATE distribution. Specifically, we estimate the following regression:

$$\hat{\tau}(X_i) = \beta_0 + X_i'\beta + \epsilon_i \tag{5}$$

where $\hat{\tau}(X_i)$ is the estimated CATE given a vector of covariates, X_i. Because the relationship between X_i and $\hat{\tau}(X_i)$ can be complicated and non-linear, β provides a simplified representation of these complicated relationships. Each parameter in the equation estimates how a unit increase in the covariate is associated on average with a change in our estimated CATE.

[10] Ensemble methods like the causal random forest can be particularly efficient because they can randomly subsample estimation and evaluation data sets individually for each tree (Wager and Athey 2018).

The results are presented in Table 13. The first column presents the estimates for the model of the relationship between treatment and time to shutdown (survival). The estimates suggest significant relationships between the CATE and the indicators for married founder and founder with children, optimism, time preferences, and the industry indicators for healthcare and biotechnology and social enterprise.

The estimates for the model of the relationship between treatment and funding are presented in column two. Four baseline variables demonstrate a statistically significant linear relationship with the CATE estimates: incorporation status (C corporation), prior entrepreneurial experience, prior P&L experience, and the indicator for the mobile industry.

We can now assess a related question with clear policy implications: would any given variable be useful for targeting if we had to prioritize treatment to different startups? Because the distribution of CATE is complex and non-linear in the distribution of baseline covariates and equation 4 is a simple linear description, the results of Table 13 do not immediately answer this question. To do so, we estimate the Targeted Operator Characteristic (TOC) curve for the variables that show a significant linear relationship (Yadlowsky et al. 2021). The TOC measures the difference in estimated CATE for the top q-th fraction of startups in the distribution of a covariate, X, from the overall estimated ATE. Specifically, the TOC measures:

$$TOC(q) = E[Y_i(1) - Y_i(0) \mid X_i \geq F^{-1}(1-q)] - E[Y_i(1) - Y_i(0)] \quad (6)$$

where $F(q)$ is the distribution of the covariate X in the sample. The TOC formula has a nice visual representation as the area under a curve representing the excess treatment effect for prioritized treatment units (i.e., higher, or lower values of the covariate). They provide information on the change to the average treatment effect if treatment was prioritized by the covariate. Observations are ranked by the covariate value (in decreasing order), and the horizontal axis reports the fraction of treated units included, starting from those with the highest predicted values of the covariate. The vertical axis shows the change in the dependent variable relative to the average treatment effect of the entire population. The solid line provides the estimates of the differential treatment effect if treatment was restricted to companies at higher or lower values of the covariate.

Figure 5 plots the TOC curves as well as 95% confidence intervals for the covariates that observe significant relationships between our CATE estimates for survival and baseline covariates. Six baseline characteristics emerged in the BLP analysis in Table 13. Unmarried founders are associated with earlier shutdown relative to the ATE; in other words, marital status appears to moderate how strongly training accelerates exit. Founder optimism also appears to be a promising moderator of shutdown effects. Training accelerates shutdown most strongly for less optimistic founders; highly optimistic founders appear less responsive to training in terms of exit timing. While time preference on the surface appears to contain useful heterogeneity, but with large confidence intervals that suggest caution is warranted in policy application. The industry variables do not appear useful for targeting, and neither does whether the founder have children.

Figure 6 presents the same analysis for funding. The TOC curves suggest that the strongest covariates for targeting purposes are prior startup experience status and prior P&L experience status. The TOC curves in this case suggest that targeting those *without* those two types of experience is most effective. They also suggest that those already incorporated as C Corps benefit more from the training. Industry does not look promising for targeting, in contrast.

While these methods provide an initial look at potential sources of treatment heterogeneity, we caution that they are descriptive, not causal. Nevertheless, many of the relationships are intuitive and support the motivations behind entrepreneurial training curriculums. For example, substantial excess treatment effects on startup founders with less startup experience or managerial experience suggests that the intervention is affecting the population that most policymakers would ex ante expect it to. Further work, likely experimental, will be needed to better identify and explore treatment heterogeneity.

5. CONCLUSION

Despite its potential importance for stimulating economic growth, little is known about the factors that contribute to innovation-driven startups' success or failure. In this paper, we use a field experiment in combination with a rich dataset collected from surveys of participating startups to investigate the importance of realistic-scope short-course entrepreneurship training on future performance for a sample of innovation-driven startups. Our results suggest that policymakers can use short education curriculums to improve outcomes for startups—a stated goal of many foundations and local development programs. Moreover, our analysis of treatment heterogeneity provides initial information on the characteristics of startups and entrepreneurs that policymakers might consider for prioritizing treatment in order to maximize the efficacy of these type of programs.

In addition, our experiment was designed specifically to easily scale towards broad implementation of our program. This design choice is in line with a recent push in the economics literature to consider the potential for scaling of offerings within the context of experimental interventions (List 2022; Al-Ubaydli, Lai, and List 2023). Often, considerable uncertainty remains as to whether experimental interventions can be scaled beyond the scope of the smaller scale randomized study. Our funding source was interested in the effects of education programs that could be potentially offered to a much larger audience.[11] To ensure that our chosen intervention was implementable at scale post-experiment at reasonable cost, we conducted extensive qualitative research at the pilot stage. The final educational program was selected based on its ability to realistically be offered to working entrepreneurs outside of experimental conditions and in the absence of participation incentives. Furthermore, we explicitly considered the ability of the intervention to be scaled through a train-the-trainer model. We utilized corporate trainers who had experience teaching short entrepreneurship training classes who were provided materials and content for the modules that were taught. In other words, we sought to ensure that none of the instructors was sufficiently unique that their main contributions couldn't be replicated by others.

Pinning down the overall welfare consequences of our intervention is difficult. Our results are consistent with the interpretation that we cause less viable startups to shut down earlier. Given the earning potential and high prior wages of individuals involved in innovation-driven ventures, it may be beneficial to society if training causes unproductive entrepreneurs to preserve their time and return to the labor force. If successful, certain innovation-driven startups by their nature may lead to outsized societal and economic impact. Such startups are inherently risky, and luck may play a large role in eventual success. We cannot rule out the possibility that some ventures that chose to shutdown sooner as a result of treatment would have ultimately been successful. From a societal perspective, it may instead be optimal to encourage all startups to continue, despite the possibility of entrepreneurs wasting time and resources, if policy makers' primary focus is to maximize the probability of increasing innovation. We consider this discussion to be outside the scope of our experiment, and worthy of future research.

REFERENCES

Al-Ubaydli, O., Lai, C.-Y., List, J.A., 2023. A simple rational expectations model of the voltage effect. NBER Working Paper No. 30850.

Athey, S., Tibshirani, J., Wager, S., 2019. Generalized random forests. Annals of Statistics 47 (2), 1148–1178.

Austin, P.C., Fine, J.P., 2017. Practical recommendations for reporting Fine–Gray model analyses for competing risk data. Statistics in Medicine 36 (27), 4391–4400.

Azoulay, P., Jones, B.F., Kim, J.D., Miranda, J., 2020. Age and high-growth entrepreneurship. American Economic Review: Insights 2 (1), 65–82.

[11] In this sense, we were not interested in exploring interventions that might produce a larger treatment effect but were outside the scope of what could reasonably be offered to a broad set of entrepreneurs.

Barsky, R.B., Juster, F.T., Kimball, M.S., Shapiro, M.D., 1997. Preference parameters and behavioral heterogeneity: An experimental approach in the Health and Retirement Study. Quarterly Journal of Economics 112 (2), 537–579.

Bennett, V.M., Chatterji, A.K., 2023. The entrepreneurial process: Evidence from a nationally representative survey. Strategic Management Journal 44 (1), 86–116.

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

Botelho, T.L., Fehder, D., Hochberg, Y.V., 2025. Innovation-driven entrepreneurship. Journal of Economic Literature, forthcoming.

Bruhn, M., Karlan, D., Schoar, A., 2018. The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in Mexico. *Journal of Political Economy* 126 (2), 635–687.

Camuffo, A., Cordova, A., Gambardella, A., Spina, C., 2020. A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial. Management Science 66 (2), 564–586.

Cattaneo, M.D., Jansson, M., Ma, X., 2018. Manipulation testing based on density discontinuity. Stata Journal 18 (1), 234–261.

Chatterji, A., Delecourt, S., Hasan, S., Koning, R.M., 2019. When does advice impact startup performance? Strategic Management Journal 40 (3), 331–356.

Chernozhukov, V., Demirer, M., Duflo, E., Fernandez-Val, I., 2018. Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India. NBER Working Paper No. 24678.

Cohen, S., Fehder, D.C., Hochberg, Y.V., Murray, F., 2019. The design of startup accelerators. Research Policy 48 (7), 1781–1797.

Cragg, J.G., 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica 39 (5), 829–844.

Davies, E., Deffebach, P., Iacovone, L., McKenzie, D., 2024. Training microentrepreneurs over Zoom: Experimental evidence from Mexico. Journal of Development Economics 167, 103244.

Davis, J.M.V., Heller, S.B., 2020. Rethinking the benefits of youth employment programs: The heterogeneous effects of summer jobs. Review of Economics and Statistics 102 (4), 664–677.

de Mel, S., McKenzie, D., Woodruff, C., 2019. Labor drops: Experimental evidence on the return to additional labor in microenterprises. American Economic Journal: Applied Economics 11 (1), 202–235.

Ewens, M., Townsend, R.R., 2020. Are early-stage investors biased against women? Journal of Financial Economics 135 (3), 653–677.

Fairlie, R.W., Robb, A.M., 2007. Why are Black-owned businesses less successful than White-owned businesses? The role of families, inheritances, and business human capital. Journal of Labor Economics 25 (2), 289–323.

Fairlie, R.W., Robb, A.M., Robinson, D.T., 2022. Black and White: Access to capital among minority-owned start-ups. Management Science 68 (4), 2377–2400.

Fehder, Daniel et al. 2019. "Assessing the Effectiveness of Entrepreneurship Education: A Randomized Control Trial." AEA RCT Registry. June 24. https://www.socialscienceregistry.org/trials/1566/history/48634

Fehder, D.C., Hausman, N., Hochberg, Y.V., 2025. Innovation and capital. Journal of Financial Economics 169, 104029.

Floyd, E., List, J.A., 2016. Using field experiments in accounting and finance. Journal of Accounting Research 54 (2), 437–475.

Fryer, R.G., Levitt, S.D., List, J.A., Sadoff, S., 2022. Enhancing the efficacy of teacher incentives through framing: A field experiment. American Economic Journal: Economic Policy 14 (4), 269–299.

Gornall, W., Strebulaev, I.A., 2023. The economic impact of venture capital: Evidence from public companies. Journal of Finance 78 (5), 2467–2522.

Guzman, J., Stern, S., 2020. The state of American entrepreneurship: New estimates of the quantity and quality of entrepreneurship for 32 U.S. states, 1988–2014. American Economic Journal: Economic Policy 12 (4), 212–243.

Graham, J.R., Harvey, C.R., Puri, M., 2013. Managerial attitudes and corporate actions. *Journal of Financial Economics* 109 (1), 103–121.

Hausman, N., 2022. University innovation and local economic growth. Review of Economics and Statistics 104 (4), 718–735.

Heckman, J.J., 1976. A life-cycle model of earnings, learning, and consumption. Journal of Political Economy 84 (4, Part 2), S9–S44.

Hochberg, Y.V., 2016. Accelerating entrepreneurs and ecosystems: The seed accelerator model. Innovation Policy and the Economy 16 (1), 25–51.

Imbens, G.W., Manski, C.F., 2004. Confidence intervals for partially identified parameters. *Econometrica* 72 (6), 1845–1857.

Jovanovic, B., 1982. Selection and the evolution of industry. Econometrica 50 (3), 649–670.

Karlan, D., Valdivia, M., 2011. Teaching entrepreneurship: Impact of business training on microfinance clients and institutions. Review of Economics and Statistics 93 (2), 510–527.

Kerr, S.P., Kerr, W.R., Dalton, M., 2019. Risk attitudes and personality traits of entrepreneurs and venture team members. Proceedings of the National Academy of Sciences 116, 17712–17716

Christopher R Knittel, Samuel Stolper, Using Machine Learning to Target Treatment: The Case of Household Energy Use, *The Economic Journal*, Volume 135, Issue 672, November 2025, Pages 2377–2401.

Kotha, R., Vissa, B., Lin, Y., Corboz, A.-V., 2023. Do ambitious entrepreneurs benefit more from training? Strategic Management Journal 44 (2), 549–575.

Lee, D.S., 2009. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies* 76 (3), 1071–1102.

Lerner, J., Malmendier, U., 2013. With a little help from my (random) friends: Success and failure in post-business school entrepreneurship. Review of Financial Studies 26 (10), 2411–2452.

List, J.A., 2022. The voltage effect: How to make good ideas great and great ideas scale. Currency, New York.

Loewenstein, G., Read, D., Baumeister, R.F., 2003. Time and decision: Economic and psychological perspectives on intertemporal choice. Russell Sage Foundation, New York.

Manso, G., 2016. Experimentation and the returns to entrepreneurship. Review of Financial Studies 29 (9), 2319–2340.

McKenzie, D., 2017. Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition. American Economic Review 107 (8), 2278–2307.

McKenzie, D. (2021). Small business training to improve management practices in developing countries: re-assessing the evidence for ‘training doesn’t work’. *Oxford Review of Economic Policy*, *37*(2), 276-301.

Powell, J.L., 1986. Censored regression quantiles. Journal of Econometrics 32 (1), 143–155.

Ries, E., 2011. The lean startup: How today’s entrepreneurs use continuous innovation to create radically successful businesses. Crown Business, New York.

Scheier, M.F., Carver, C.S., Bridges, M.W., 1994. Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. Journal of Personality and Social Psychology 67 (6), 1063–1078.

Scott, E.L., Shu, P., Lubynsky, R.M., 2019. Entrepreneurial uncertainty and expert evaluation: An empirical analysis. Management Science 65 (11), 5039–5056.

Sen, A., Guzman, J., Oh, J.J., 2018. Recruiting entrepreneurs: Evidence from a field experiment. *Management Science* 64 (11), 4992–5008.

Stock, J.H., Yogo, M., 2005. Testing for weak instruments in linear IV regression. In: Andrews, D.W.K., Stock, J.H. (Eds.), Identification and inference for econometric models: Essays in honor of Thomas Rothenberg. Cambridge University Press, Cambridge, pp. 80–108.

Tobin, J., 1958. Estimation of relationships for limited dependent variables. Econometrica 26 (1), 24–36.

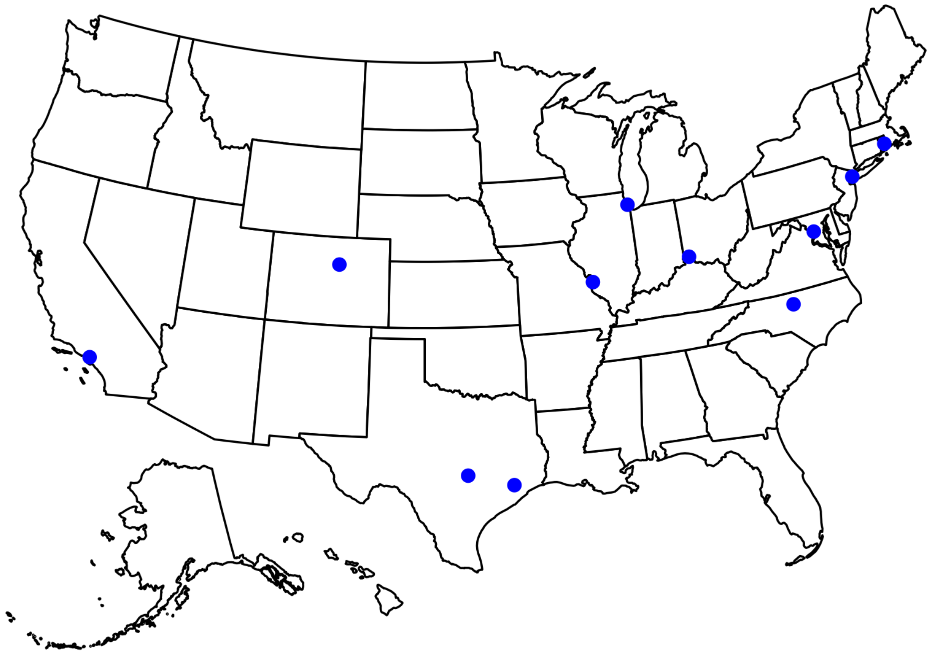
Vereshchagina, G., Hopenhayn, H.A., 2009. Risk taking by entrepreneurs. American Economic Review 99 (5), 1808–1830.

Wager, S., Athey, S., 2018. Estimation and inference of heterogeneous treatment effects using random forests. Journal of the American Statistical Association 113 (523), 1228–1242.

Yadlowsky, S., Fleming, S., Shah, N., Brunskill, E., Wager, S., 2025. Evaluating treatment prioritization rules via rank-weighted average treatment effects. Journal of the American Statistical Association, 120(549).

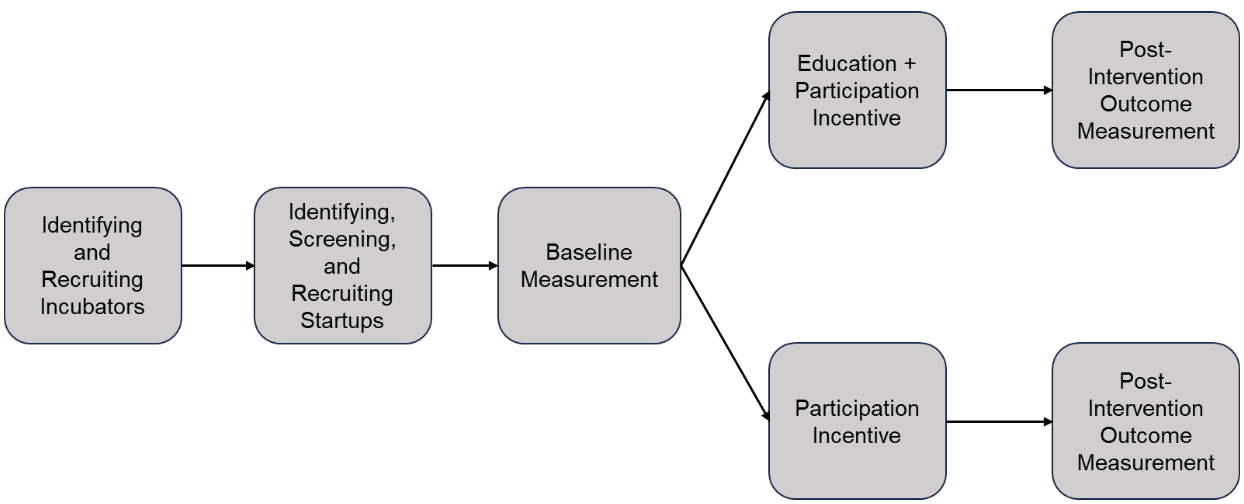
FIGURES

Figure 1: Map of Locations of Incubator and Startup Locations



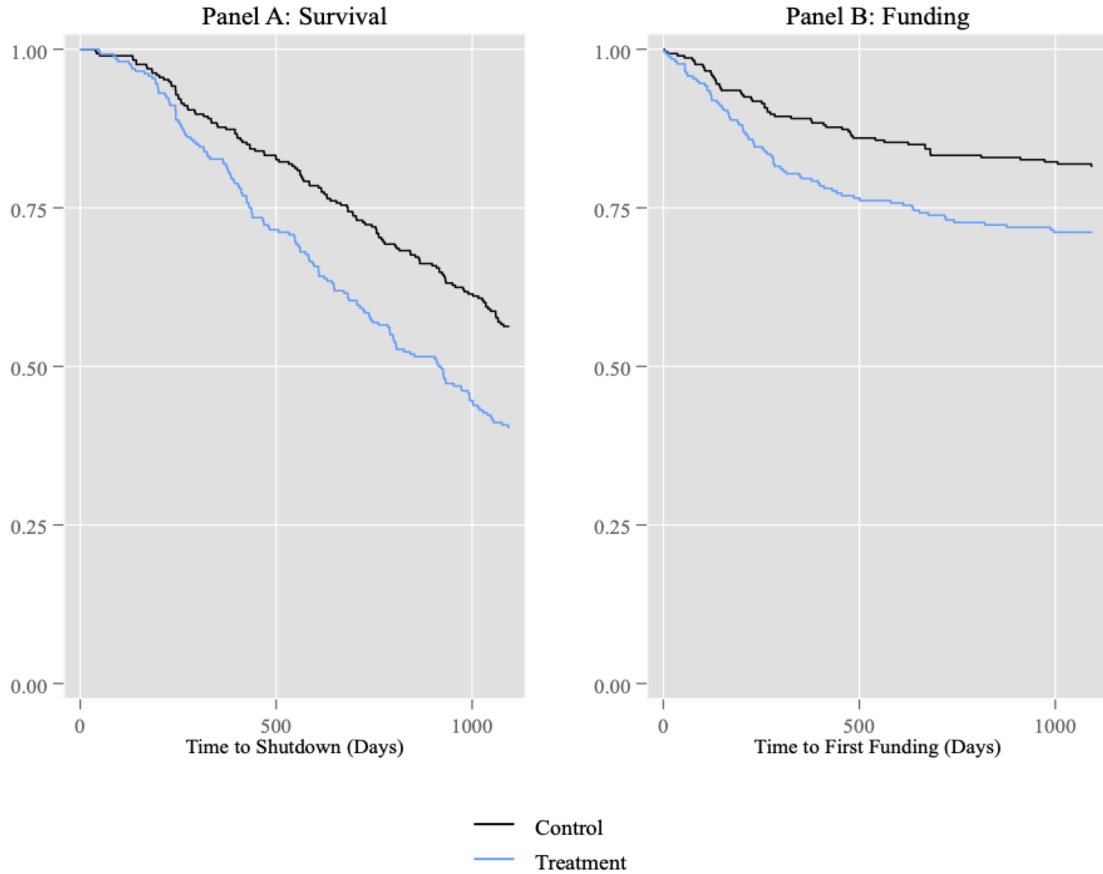
Notes. This map shows the location of the incubators and startups enrolled in our field experiment. The twelve incubator locations inclusive of our pilot site (Matter in Chicago) are indicated by the blue dots.

Figure 2: Schematic of Experimental Design



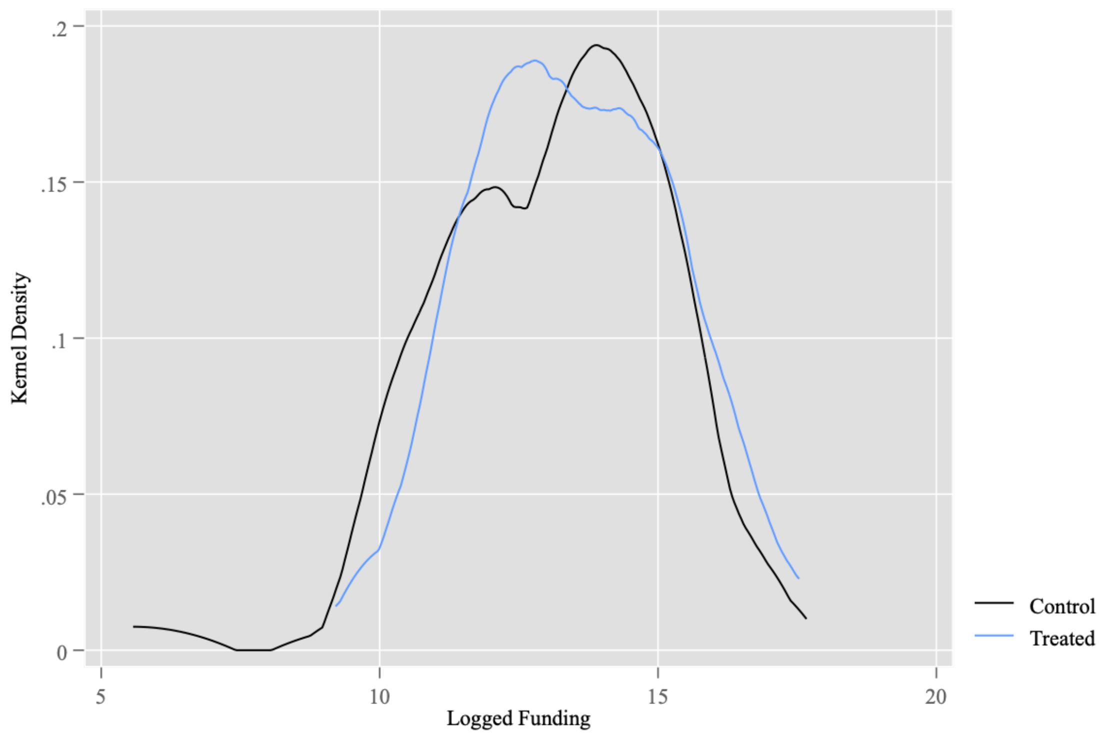
Notes. This figure presents a schematic of our RCT design.

Figure 3: Kaplan-Meier Plots by Treatment Status



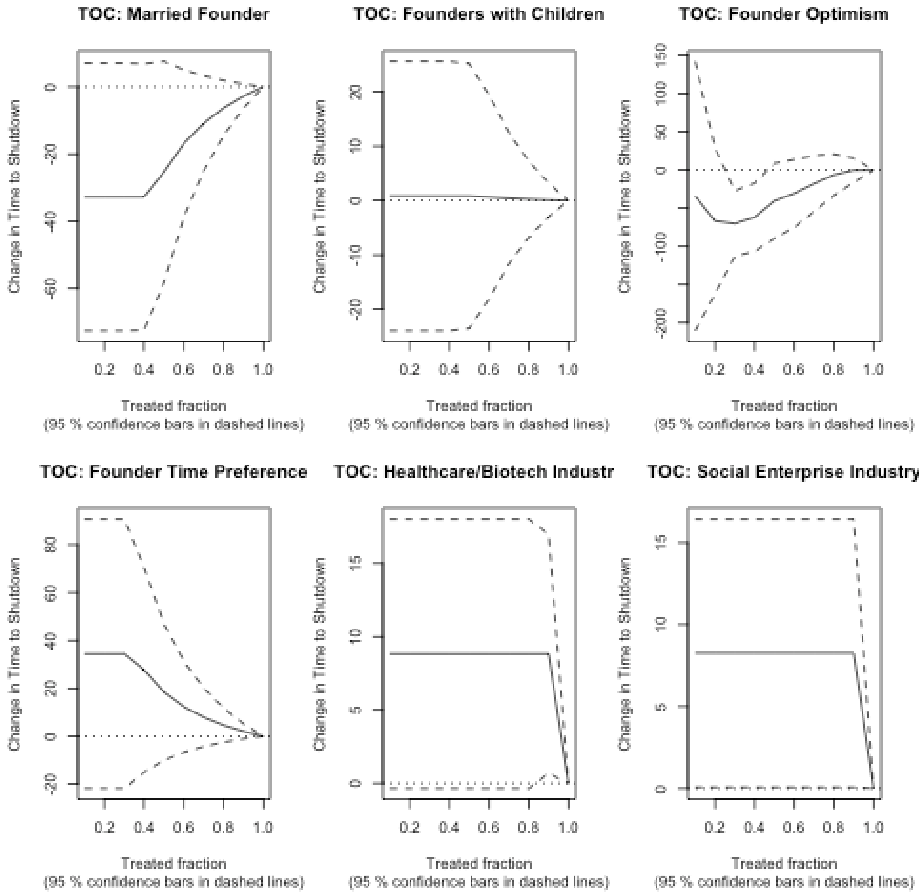
Notes. This figure contains a series of graphs that each contain plots of the Kaplan-Meier survivor function by treatment condition for different time-based dependent variables. In Panel A, we provide the plot of the Kaplan-Meier survivor function for time to shutdown for startups in days in the treatment and control conditions of our experiment. Log-rank test of the difference in time to shutdown between treated and control startups is statistically significant (X^2= 15.79, $p < 0.0001$). In Panel B, we show the plot of the Kaplan-Meier survivor function for time to first fundraising in days for startups in the treatment and control conditions of our experiment. Log-rank test of the difference in time to first VC round between treated and control startups is statistically significant (X^2= 8.76, $p < 0.001$). Log-rank test of the difference in time to first revenue between treated and control startups is not statistically significant (X^2= 0.22, $p < 0.639$).

Figure 4: Kernel Density Plot of Funding by Treatment Status



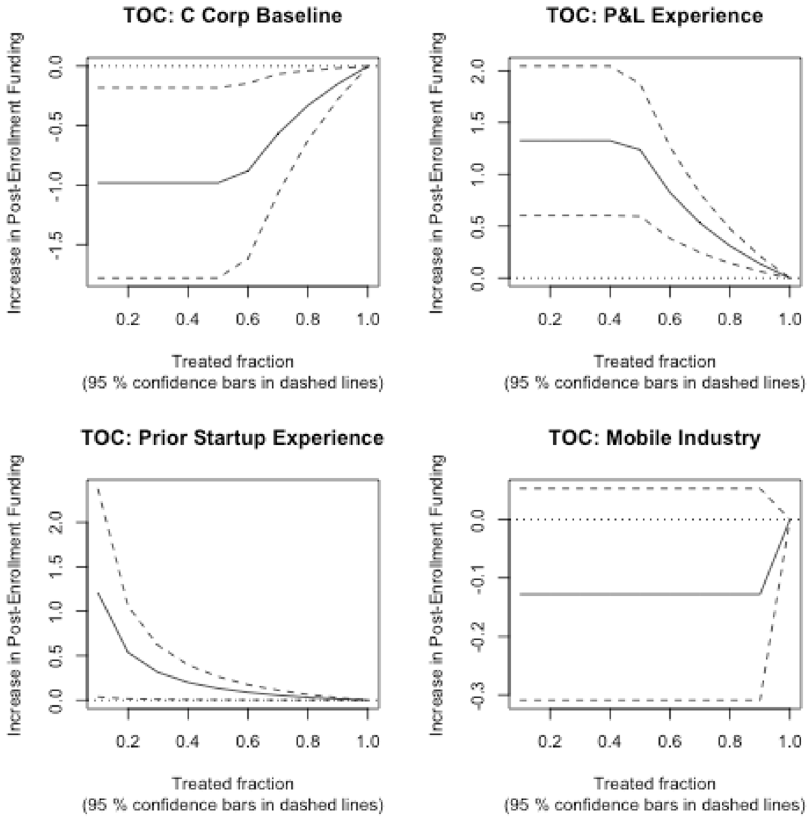
Notes. This figure provides kernel density plots of the distribution of log funding at endline (Log (1+Total Funding Dollars Raised by Endline)) for all non-zero observations of funding for treatment and control groups. Kolmogorov-Smirnov test of the equality of distributions provides evidence differences in the distribution of the treatment and control groups (p=0.023).

Figure 5: TOC Curves for Survival



Notes. Each of the plots in this figure plots the Targeting Operator Characteristic (TOC) curves, as proposed in Yadlowsky et al. (2025), which provides information on the change to the average treatment effect if treatment was prioritized by the covariate. Observations are ranked by the covariate value (in decreasing order), and the horizontal axis reports the fraction of treated units included, starting from those with the highest predicted values of the covariate. The vertical axis shows the change in the dependent variable relative to the average treatment effect of the entire population. The solid line provides the estimates of the differential treatment effect if treatment was restricted to companies at higher or lower values of the covariate. The dotted lines represent 95% confidence intervals for the point estimates along the distribution. Plots are provided for statistically significant covariates in Table 13 column (1).

Figure 6: TOC Curves for Funding



Notes. Each of the plots in this figure plots the Targeting Operator Characteristic (TOC) curves, as proposed in Yadlowsky et al. (2025), which provides information on the change to the average treatment effect if treatment was prioritized by the covariate. Observations are ranked by the covariate value (in decreasing order), and the horizontal axis reports the fraction of treated units included, starting from those with the highest predicted values of the covariate. The vertical axis shows the change in the dependent variable relative to the average treatment effect of the entire population. The solid line provides the estimates of the differential treatment effect if treatment was restricted to companies at higher or lower values of the covariate. The dotted lines represent 95% confidence intervals for the point estimates along the distribution. Plots are provided for statistically significant covariates in Table 13 column (2).

TABLES

Table 1: Sample Size by Location

	(1)	(2)	(3)
Site	Total	Control	Treatment
CIC - Boston	29	16	13
CIC - Cambridge	20	11	9
CIC - St. Louis	20	10	10
Capital Factory - Austin	60	31	29
1871 - Chicago	82	43	39
Union Hall - Cincinnati	44	25	19
Galvanize - Denver	34	18	16
Station - Houston	17	8	9
Cross Campus - Los Angeles	79	41	38
Matter - NYC	87	47	40
American Underground - Raleigh/Durham	48	26	22
1776 - Washington D.C.	33	17	16
Observations	553	293	260

Notes. Summary statistics for sample size (total, treatment group, and control) by experiment site.

Table 2: Survey Response Rates over Time

	Treatment Status		
	Control	Treated	Total
Survey			
Baseline	1.00	1.00	1.00
Follow-up 1 - 6 months	0.97	0.90	0.93
Follow-up 2 - 12 months	0.85	0.80	0.83
Follow-up 3 - 18 months	0.84	0.79	0.82
Follow-up 4 - 24 months	0.81	0.75	0.78

Notes. Survey response rates for treatment, control, and the full sample by period.

Table 3: Summary Statistics

Panel A: Outcome Variables

	Observations	Mean	Median	Std. Dev.	Min	Max
At Endline						
Time to Closure	553	820.93	1047.00	335.18	39.00	1095.00
Time to Funding	553	914.29	1095.00	350.67	1.00	1095.00
I{Raise}	553	0.23	0.00	0.42	0.00	1.00
Logged Funding	553	3.11	0.00	5.72	0.00	17.67
Founded Startup Next	430	0.46	0.00	0.50	0.00	1.00
Per Period (Panel)						
I{Closure}	3318	0.28	0.00	0.45	0.00	1.00
I{Raise}	3318	0.06	0.00	0.24	0.00	1.00
Logged Funding	3318	0.82	0.00	3.18	0.00	17.46
I{Pivot}	1593	0.05	0.00	0.21	0.00	1.00
Employment	1438	3.91	2.00	6.72	0.00	80.00
Revenue	1252	6.12	8.01	5.07	0.00	15.76

Panel B: Baseline Covariates

	Observations	Mean	Median	Std. Dev.	Min	Max
Company Covariates						
Corporation	553	0.43	0	0.50	0	1
Delaware Registration	553	0.43	0	0.50	0	1
Baseline Funding	553	0.39	0	1.58	0	22
Baseline Revenue	553	0.24	0	1.28	0	25
Baseline Employment	553	2.66	2	3.26	0	19
Startup Founder Demographic Covariates						
Founder with Children	553	0.46	0	0.50	0	1
BIPOC Founder	553	0.25	0	0.43	0	1
Founder with P&L Experience	553	0.53	1	0.50	0	1
Married Founder	553	0.56	1	0.50	0	1
Female Founder	553	0.37	0	0.48	0	1
Prior Entre. Experience	553	0.98	1	0.13	0	1
STEM Education	553	0.53	1	0.50	0	1
Business Education	553	0.46	0	0.50	0	1
Top 10 Degree	553	0.18	0	0.39	0	1
Top 11-20 Degree	553	0.11	0	0.31	0	1
Average Age	553	34.38	33.5	8.37	18	61
Startup Founder Behavioral Covariates						
Max Founder Optimism	553	19.95	21	3.48	6	24
Max Founder Time Preference	553	0.65	1	0.48	0	1
Max Founder Risk Preference	553	3.60	4	0.72	1	4

Panel C: Industry and Stage of Development

	Observations	Mean	Median	Std. Dev.	Min	Max
Industry Fixed Effects						
Education	553	0.07	0	0.25	0	1.00
Finance	553	0.06	0	0.23	0	1.00
Healthcare/Biotechnology	553	0.09	0	0.29	0	1.00
E-Commerce	553	0.16	0	0.37	0	1.00
IT & Enterprise Software	553	0.21	0	0.41	0	1.00
Media	553	0.04	0	0.20	0	1.00
Mobile & Wireless	553	0.05	0	0.21	0	1.00
Natural Resources/Energy	553	0.01	0	0.09	0	1.00
Social Enterprise	553	0.03	0	0.18	0	1.00
Social Media/Social Networking	553	0.05	0	0.22	0	1.00
Tourism	553	0.02	0	0.15	0	1.00
Other	553	0.20	0	0.40	0	1.00
Stage of Development Fixed Effects						
Concept	553	0.05	0	0.21	0	1.00
Working Prototype in development	553	0.38	0	0.49	0	1.00
Functional product with users	553	0.40	0	0.49	0	1.00
Scaling sales	553	0.17	0	0.38	0	1.00

Notes. The table presents summary statistics for our analysis sample. Panel A presents summary statistics for our outcome variables, Panel B for baseline covariates, and Panel C for industry and stage indicators. Variables are described in detail in Section 3 of the paper, and detailed notes on their construction are provided in Online Appendix A.

Table 4: Summary Statistics by Treatment Status and Test of Balance

Panel A: Company, Demographic, and Behavioral Covariates

	Treatment Status			
	Control	Treated	Total	Balance
Company Covariates				
Corporation	0.39	0.47	0.43	0.074*
	(0.49)	(0.50)	(0.50)	(0.043)
Delaware Registration	0.41	0.44	0.43	0.030
	(0.49)	(0.50)	(0.50)	(0.043)
Baseline Funding	0.29	0.50	0.39	0.022
	(1.05)	(2.01)	(1.58)	(0.013)
Baseline Revenue	0.27	0.20	0.24	-0.010
	(1.60)	(0.79)	(1.28)	(0.017)
Baseline Employment	2.83	2.47	2.66	-0.008
	(3.55)	(2.88)	(3.26)	(0.007)
Startup Founder Demographic Covariates				
Female Founded	0.35	0.38	0.37	0.035
	(0.48)	(0.49)	(0.48)	(0.044)
Prior Entrepreneurship	0.98	0.98	0.98	-0.030
	(0.13)	(0.14)	(0.13)	(0.160)
STEM Education	0.56	0.5	0.53	-0.060
	(0.5)	(0.5)	(0.5)	(0.043)
Business Education	0.42	0.51	0.46	0.088**
	(0.49)	(0.5)	(0.5)	(0.042)
Top 10 Degree	0.17	0.19	0.18	0.036
	(0.38)	(0.39)	(0.39)	(0.055)
Top 11-20 Degree	0.11	0.11	0.11	-0.013
	(0.32)	(0.31)	(0.31)	(0.068)
Founder with P&L Experience	0.49	0.56	0.53	0.067
	(0.5)	(0.5)	(0.5)	(0.042)
Married Founder	0.55	0.58	0.56	0.032
	(0.5)	(0.49)	(0.5)	(0.043)
Founder with Children	0.45	0.47	0.46	0.023
	(0.5)	(0.5)	(0.5)	(0.043)
BIPOC Founder	0.23	0.26	0.25	0.040
	(0.42)	(0.44)	(0.43)	(0.049)
Average Age	34.03	34.77	34.38	0.003
	(8.19)	(8.56)	(8.37)	(0.003)
Startup Founder Behavioral Covariates				
Max Founder Optimism	20.04	19.85	19.95	-0.004
	(3.34)	(3.65)	(3.48)	(0.006)
Max Founder Time Preference	0.63	0.67	0.65	0.046
	(0.48)	(0.47)	(0.48)	(0.045)
Max Founder Risk Preference	3.58	3.63	3.6	0.024
	(0.76)	(0.67)	(0.72)	(0.030)

Panel B: Industry and Stage

	Treatment Status			
	Control	Treated	Total	Balance
Industry Categories				
Education	0.06	0.07	0.07	0.017
	(0.25)	(0.25)	(0.25)	(0.085)
Finance	0.08	0.04	0.06	-0.167*
	(0.26)	(0.19)	(0.23)	(0.091)
Healthcare/Biotechnology	0.08	0.10	0.09	0.077
	(0.27)	(0.31)	(0.29)	(0.074)
E-Commerce	0.17	0.15	0.16	-0.037
	(0.38)	(0.36)	(0.37)	(0.057)
IT & Enterprise Software	0.24	0.18	0.21	-0.081
	(0.43)	(0.39)	(0.41)	(0.052)
Media	0.03	0.05	0.04	0.145
	(0.17)	(0.23)	(0.20)	(0.106)
Mobile & Wireless	0.04	0.05	0.05	0.094
	(0.19)	(0.23)	(0.21)	(0.102)
Natural Resources/Energy	0.01	0.01	0.01	0.131
	(0.08)	(0.11)	(0.09)	(0.225)
Social Enterprise	0.04	0.03	0.03	-0.105
	(0.20)	(0.16)	(0.18)	(0.117)
Social Media/Social Networking	0.05	0.05	0.05	-0.023
	(0.23)	(0.22)	(0.22)	(0.095)
Tourism	0.01	0.03	0.02	0.201
	(0.12)	(0.17)	(0.15)	(0.146)
Other	0.17	0.22	0.20	0.072
	(0.38)	(0.41)	(0.40)	(0.054)
Stage of Development Categories				
Concept	0.05	0.04	0.05	-0.049
	(0.22)	(0.20)	(0.21)	(0.100)
Working prototype in development	0.39	0.38	0.38	-0.009
	(0.49)	(0.49)	(0.49)	(0.044)
Functional product with users	0.40	0.40	0.40	0.001
	(0.49)	(0.49)	(0.49)	(0.043)
Scaling sales	0.17	0.17	0.17	0.004
	(0.38)	(0.38)	(0.38)	(0.056)

Notes. The table presents the distribution of the baseline covariates for the startups and their founders in our sample across treatment status as well as the results of a balance test. The tables is broken into two panels due to its length. The first column shows the mean and standard deviation (in parentheses) of the baseline covariates of the startups in the control group. The first column shows the mean and standard deviation (in parentheses) of the baseline covariates of the startups in the treatment group. The third column shows the mean and standard deviation (in parentheses) of the baseline covariates for all the startups in the sample. The fourth column shows the coefficient estimate of separate regressions for each covariate of the form $Treat = \beta_0 + \beta_1\, Covar + \epsilon$. In the column, we show the estimates for each covariate of β_1 and its standard error (in parentheses). * indicates significance at the 10% level, ** at 5%, and ***at 1%.

Table 5: Raw Survival by Follow-up Period

	(1)	(2)	(3)	(4)	(5)
Time Period	Full Sample	Control	Treatment	Diff. (3)-(2)	P-Val
6 Months	96.38	96.93	95.77	-1.16	0.467
12 Months	85.35	87.71	82.69	-5.02	0.096
18 Months	76.67	81.91	70.77	-11.14	0.002
24 Months	66.55	73.04	59.23	-13.81	0.001
30 Months	59.13	65.87	51.54	-14.33	0.001
36 Months	49.01	56.31	40.77	-15.54	0.000
Observations	553	293	260	553	553

Notes. This table shows the raw differences in the percentage of the sample surviving during each of the 6-month periods in our sample as well as individual tests of the statistical significance of the difference at each 6-month interval. Acquired startups ("positive" exits) are treated as survivors. Column (1) shows the percentage of the total sample that survived in each 6-month follow-up period. Column (2) shows the survival percentage for the control group in each 6-month follow-up period. Column (3) shows the survival percentage for the treatment group in each 6-month follow-up period. Column (4) shows the difference in survival percentage between treatment and control groups. Column (5) shows the p-value associated with the t-statistic associated with a t-test of the difference in the distribution of surviving startups in the treatment and control groups in each 6-month follow-up period.

Table 6: Startup Survival

Panel A: Panel Regression Models of Startup Closure

	(1) Closure	(2) Closure	(3) Closure
Treated	0.105***	0.098***	0.094***
	(0.028)	(0.029)	(0.028)
Observations	3318	3318	3318
Site Fixed Effects	X	X	X
Followup-Period Fixed Effects	X	X	X
Only Unbalanced Covariates	-	X	-
All Covariates	-	-	X

Panel B: Hazard Models of Startup Closure

	(1) Closure	(2) Closure	(3) Closure
Treated	1.581***	1.537***	1.519***
	(0.200)	(0.241)	(0.241)
Observations	553	553	553
Site Fixed Effects	X	X	X
Only Unbalanced Covariates	-	X	-
All Covariates	-	-	X

Notes. The table presents analysis of startup survival over the study period. Panel A presents the estimates from panel regression models, where an observation is a startup x followup period, and the outcome variable is an indicator as to whether the startup closed (negative exit) during that period. All models include site fixed effects and followup-period fixed effects. Column (2) adds unbalanced covariates as controls, while column (3) adds all covariates as controls. Standard errors are clustered at the startup level. Panel B presents the estimates from Cox proportional hazard regressions predicting the hazard of closure (negative exit) and includes the key variable of interest, treatment, as well as site and industry fixed effects. Columns (2)-(4) include additional control covariates. All models include site fixed effects. Column (2) adds unbalanced covariates as controls, while column (3) adds all covariates as controls. Standard errors are clustered at the site level. Standard errors are presented in parentheses. Covariates are described in detail in Section 3.2 of the paper. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 7: Pivoting

Panel A: Panel Regressions of Pivoting

	(1) Pivot	(2) Pivot	(3) Pivot
Treated	-0.026**	-0.019*	-0.022*
	(0.011)	(0.011)	(0.012)
Observations	1593	1593	1593
Site Fixed Effects	X	X	X
Follow-up Period Fixed Effects	X	X	X
Only Unbalanced Covariates	-	X	-
All Covariates	-	-	X

Notes. The table presents panel models of pivoting status for startups over the follow-up periods of the study. The observation level is a startup x followup-period. The outcome variable is an indicator for whether the startup conducted a pivot in that period. The panel is unbalanced, and observation inclusion is conditional on not having shutdown or exited via acquisition. All specifications include site fixed effects and follow-up period fixed effects. Column (2) adds unbalanced covariates as controls, while column (3) adds all covariates as controls. Standard errors are clustered at the startup level. Standard errors are presented in parentheses. Covariates are described in detail in Section 3.2 of the paper. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 8: Funding Outcomes

Panel A: Panel Regression Models of Funding

	(1) I{Raise}	(2) I{Raise}	(3) I{Raise}	(4) Logged $	(5) Logged $	(6) Logged $
Treated	0.027**	0.023**	0.024**	0.347**	0.298**	0.312**
	(0.011)	(0.011)	(0.011)	(0.148)	(0.144)	(0.140)
Observations	3318	3318	3318	3318	3318	3318
Site Fixed Effects	X	X	X	X	X	X
Followup-Period Fixed Effects	X	X	X	X	X	X
Only Unbalanced Covariates	-	X	-	-	X	-
All Covariates	-	-	X	-	-	X

Panel B: Hazard Models of Time to First Fundraising

	(1) Funding	(2) Funding	(3) Funding	(4) Funding	(5) Funding	(6) Funding
Treated	1.698***	1.634***	1.813***	1.673***	1.620***	1.799***
	(0.318)	(0.268)	(0.337)	(0.314)	(0.271)	(0.341)
Observations	553	553	553	553	553	553
Site Fixed Effects	X	X	X	X	X	X
Only Unbalanced Covariates	-	X	-	-	X	-
All Covariates	-	-	X	-	-	X

Panel C: Funding Levels at Endline

	(1) I{Raise}	(2) I{Raise}	(3) I{Raise}	(4) Logged $	(5) Logged $	(6) Logged $
Treated	0.106**	0.096**	0.097***	1.496**	1.356***	1.368***
	(0.038)	(0.032)	(0.031)	(0.497)	(0.426)	(0.398)
Observations	553	553	553	553	553	553
Site Fixed Effects	X	X	X	X	X	X
Only Unbalanced Covariates	-	X	-	-	X	-
All Covariates	-	-	X	-	-	X

Panel D: Cragg Hazard Models

	Logged Funding		
	(1)	(2)	(3)
Treated	0.422	0.527**	0.595*
	(0.329)	(0.232)	(0.340)
Survival Time	0.001*	0.001**	0.001
	(0.001)	(0.001)	(0.001)
	I{Raise}		
	(1)	(2)	(3)
Treated	0.473***	0.449***	0.521***
	(0.134)	(0.136)	(0.137)
Survival Time	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)
Observations	553	553	553
Site Fixed Effects	X	X	X
Only Unbalanced Covariates	-	X	-
All Covariates	-	-	X

Notes. The tables presents estimates of models of fundraising outcomes for the sample. Panel A presents estimates from panel regression models where an observation is at the startup x followup-period level, and the outcome variables are an indicator for whether the startup raised funding in that period and the natural logarithm of one plus the funding dollars raised. All specifications include site fixed effects and follow-up period fixed effects. Columns (2) and (5) add unbalanced covariates as controls, while columns (3) and (6) add all covariates as controls. Standard errors are clustered at the startup level. Panel B presents estimates from Cox Proportional Hazard models of the time to first fundraising by the startup, and the unit of observation is the startup. All specifications include site fixed effects. Columns (2) and (5) add unbalanced covariates as controls, while columns (3) and (6) add all covariates as controls. Standard errors are clustered at the site level. Panel C presents estimates from regression models where the unit of observation is the startup, and the outcome variables are an indicator for whether the startup had raised funding by the end of the study or the natural logarithm of one plus the total amount of funding raised by the startup over the study period. All specifications include site fixed effects. Columns (2) and (5) add unbalanced covariates as controls, while columns (3) and (6) add all covariates as controls. Standard errors are clustered at the site level. Panel D presents estimates from regression models which jointly estimate the extensive and intensive margin of treatment while incorporating survival time in a Cragg double hurdle model (Cragg 1971). The double hurdle model jointly estimates the probability of raising zero funding and the total amount of funding conditional on being a non-zero observation. All specifications include site fixed effects. Columns (2) and (5) add unbalanced covariates as controls, while columns (3) and (6) add all covariates as controls. Standard errors are clustered at the site level. Standard errors are presented in parentheses. Covariates are described in detail in Section 3.2 of the paper. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 9: Panel Regression Models of Startup Employment

	(1) Employment	(2) Employment	(3) Employment
Treated	1.571***	1.579***	1.436***
	(0.511)	(0.497)	(0.463)
Baseline Employment	0.471***	0.445***	0.338***
	(0.071)	(0.070)	(0.070)
Observations	1438	1438	1438
Followup-Period Fixed Effects	X	X	X
Site Fixed Effects	X	X	X
Only Unbalanced Covariates	-	X	-
All Covariates	-	-	X

Notes. The table presents estimates from panel regression models where an observation is a startup x followup period, and the outcome variable is the number of startup non-founder employees during that period. All models include site fixed effects and followup-period fixed effects. Column (2) adds unbalanced covariates as controls, while column (3) adds all covariates as controls. Standard errors are clustered at the startup level. Standard errors are presented in parentheses. Covariates are described in detail in Section 3.2 of the paper. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 10: Panel Regression Models of Revenue

	(1) Logged Revenue	(2) Logged Revenue	(3) Logged Revenue
Treated	0.622*	0.627*	0.554*
	(0.344)	(0.339)	(0.304)
Logged Baseline Revenue	0.611***	0.597***	0.384***
	(0.033)	(0.035)	(0.045)
Observations	1252	1252	1252
Site Fixed Effects	X	X	X
Time-Period Fixed Effects	X	X	X
Only Unbalanced Covariates	-	X	-
All Covariates	-	-	X

Notes. The table presents estimates from panel regression models where an observation is a startup x followup period, and the outcome variable is the last quarter's revenue as reported by the startup for that period. All models include site fixed effects and followup-period fixed effects. Column (2) adds unbalanced covariates as controls, while column (3) adds all covariates as controls. Standard errors are clustered at the startup level. Standard errors are presented in parentheses. Covariates are described in detail in Section 3.2 of the paper. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 11: Entrepreneurship Choice in Next Job Role

	Founded Startup Next		
	(1)	(2)	(3)
Treated	-0.094**	-0.085**	-0.095**
	(0.034)	(0.030)	(0.036)
Observations	428	428	428
Site Fixed Effects	X	X	X
Only Unbalanced Covariates	-	X	-
All Covariates	-	-	X

Notes. The table presents estimates of regressions where the observation level is an individual founder, and the outcome variable whether their next employment after shutting down their startup was to found another startup. The sample for this table is all founders of shutdown startups in our sample that had a LinkedIn profile we could use to observe their choice of entrepreneurship after shutdown of the startup enrolled in our study. All regressions are linear probability models. Standard errors are clustered at the site level. Column (2) adds unbalanced covariates as controls, while column (3) adds all covariates as controls. Standard errors are presented in parentheses. Covariates are described in detail in Section 3.2 of the paper. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 12: Treatment on Treated

Panel A: First Stage

	Sessions Attended			Attended Any Sessions		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	2.163***	2.206***	2.201***	0.653***	0.659***	0.651***
	(0.155)	(0.153)	(0.154)	(0.038)	(0.037)	(0.037)
Average First-Stage F-Statistic	199.49	235.58	238.37	253.76	300.62	288.55
Observations	553	553	553	553	553	553
Site Fixed Effects	X	X	X	X	X	X
Only Unbalanced Covariates	-	X	-	-	X	-
All Covariates	-	-	X	-	-	X

Panel B: IV Estimates for Endline Variables

	Time to Shutdown					
	(1)	(2)	(3)	(4)	(5)	(6)
Sessions Attended	-49.632***	-45.655***	-44.324***			
	(13.447)	(15.214)	(14.360)			
Attended Any Sessions				-164.483***	-152.814***	-149.017***
				(48.571)	(53.545)	(48.804)
	I{Raise}					
	(1)	(2)	(3)	(4)	(5)	(6)
Sessions Attended	0.049***	0.043***	0.044***			
	(0.018)	(0.014)	(0.014)			
Attended Any Sessions				0.162***	0.145***	0.149***
				(0.060)	(0.049)	(0.047)
	Logged Funding					
	(1)	(2)	(3)	(4)	(5)	(6)
Sessions Attended	0.691***	0.615***	0.625***			
	(0.231)	(0.197)	(0.180)			
Attended Any Sessions				2.290***	2.057***	2.102***
				(0.790)	(0.659)	(0.613)
	Founded Startup Next					
	(1)	(2)	(3)	(4)	(5)	(6)
Sessions Attended	-0.051***	-0.047***	-0.054***			
	(0.017)	(0.015)	(0.017)			
Attended Any Sessions				-0.157***	-0.145***	-0.169***
				(0.054)	(0.046)	(0.053)
Site Fixed Effects	X	X	X	X	X	X
Only Unbalanced Covariates	-	X	-	-	X	-
All Covariates	-	-	X	-	-	X

Panel C: IV Estimates for Panel Variables

	Pivot					
	(1)	(2)	(3)	(4)	(5)	(6)
Sessions Attended	-0.011**	-0.007*	-0.009*			
	(0.005)	(0.004)	(0.005)			
Attended Any Sessions				-0.036**	-0.025	-0.029*
				(0.015)	(0.015)	(0.016)
	Employment					
	(1)	(2)	(3)	(4)	(5)	(6)
Sessions Attended	0.721**	0.713***	0.623***			
	(0.288)	(0.254)	(0.189)			
Attended Any Sessions				2.391***	2.377***	2.113***
				(0.854)	(0.784)	(0.599)
	Revenue					
	(1)	(2)	(3)	(4)	(5)	(6)
Sessions Attended	0.215	0.207	0.190			
	(0.174)	(0.173)	(0.144)			
Attended Any Sessions				0.716	0.692	0.641
				(0.561)	(0.567)	(0.477)
Followup-Period Fixed Effects	X	X	X	X	X	X
Site Fixed Effects	X	X	X	X	X	X
Only Unbalanced Covariates	-	X	-	-	X	-
All Covariates	-	-	X	-	-	X

Notes. The table provides estimates of the effect of treatment on treated by using assignment to treatment as an instrument for attendance of training intervention sessions. Panel A shows the first stage of the IV. Panel B shows the second stage estimates for outcome variables measured at endline. Panel C presents the estimates for the second stage for panel-level variables, where each outcome variable observation is a startup x followup period. Standard errors are clustered at the site level. Column (2) adds unbalanced covariates as controls, while column (3) adds all covariates as controls. Standard errors are presented in parentheses. Covariates are described in detail in Section 3.2 of the paper. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 13: Best Linear Predictor of Covariates on Estimated CATE

Variable	(1) Survival	(2) Funding
Industry Fixed Effects		
Finance	77.725	3.547
	(112.863)	(2.695)
Healthcare/Biotechnology	-112.095*	3.646
	(57.125)	(2.712)
IT & Enterprise Software	53.938	0.348
	(49.004)	(2.396)
Media	-36.382	-1.583
	(209.641)	(2.286)
Mobile & Wireless	171.672	5.028*
	(120.354)	(2.995)
Natural Resources/Energy	185.665	-4.672
	(335.502)	(5.272)
Social Enterprise	-243.123**	-1.660
	(103.608)	(2.881)
Social Media/Social Networking	96.895	1.765
	(158.829)	(3.269)
Tourism	-109.273	0.704
	(628.891)	(4.076)
Other	-36.933	1.298
	(71.442)	(1.908)
Stage of Development Fixed Effects		
Concept	23.680	0.620
	(213.674)	(4.135)
Working Prototype in development	12.721	-2.041
	(118.815)	(2.875)
Functional product with users	7.239	-2.877
	(139.505)	(2.9)
Scaling sales	29.218	0.432
	(169.041)	(2.854)

Variable	(1) Survival	(2) Funding
Corporation	14.691	3.131***
	(63.562)	(0.908)
Delaware Registration	-74.351	-0.707
	(94.531)	(1.765)
Baseline Funding	-24.400	-0.624
	(28.396)	(0.646)
Baseline Employment	5.958	-0.061
	(9.249)	(0.189)
Baseline Revenue	-0.106	0.002
	(0.322)	(0.005)
Age	1.557	-0.026
	(3.291)	(0.084)
Female Founded	-3.689	1.647
	(62.31)	(1.521)
Prior Entrepreneurship	8.316	-8.769*
	(267.708)	(5.006)
STEM Education	71.432	1.847
	(64.187)	(1.204)
Business Education	80.508	0.475
	(61.347)	(1.73)
Top 10 Degree	111.235	2.031
	(79.971)	(1.897)
Top 11-20 Degree	50.092	0.489
	(68.658)	(2.468)
Founder with P&L Experience	-70.806	-2.579***
	(53.901)	(0.96)
Married Founder	85.012*	-0.231
	(49.573)	(0.957)
Founder with Children	-72.291*	0.266
	(41.41)	(1.63)
BIPOC Founder	-30.782	0.960
	(62.179)	(1.617)
Max Founder Optimism	15.117*	0.119
	(8.395)	(0.194)
Max Founder Time Preference	-99.885*	-1.050
	(52.596)	(1.064)
Max Founder Risk Preference	8.294	0.943
	(38.003)	(0.619)
Observations	553	553

Notes. This table explores potential heterogeneity by measuring the relationship between the predicted conditional average treatment effect (CATE) from our calibrated generalized random forest models (GRF) and the baseline covariates of the startup. To do so, we use a best linear prediction of the covariates onto the distribution of our predicted CATE. The results in the two columns correspond to the GRF models for the two main outcome variables of our study, funding and survival and their standard errors (in parentheses). * indicates significance at the 10% level, ** at 5%, and ***at 1%.