IGL Working Paper Series

Better Flee From Freedom? The Effects of Structured Accountability on New Venture Performance

Michael Leatherbee, Juanita González-Uribe and Edgar E. Kausel

June 2023

IGL Working Paper No. 23/05

IGL Working Paper No. 23/05 June 2023

Better Flee From Freedom? The Effects of Structured Accountability on New Venture Performance

Michael Leatherbee*
Pontificia Universidad Católica de Chile
mleatherbee@ing.puc.cl

Edgar E. Kausel
Pontificia Universidad Católica de Chile
ekausel@uc.cl

Juanita González-Uribe
London School of Economics
j.gonzalez-uribe@lse.ac.uk

ABSTRACT

Accountability decreases the performance of people conducting creative, unstructured tasks, but helps routine tasks that lack uncertainty. What about for new ventures? Does it benefit them to develop structures that make them accountable for their plans? We explore these questions in the context of a business accelerator, conducting a randomized controlled trial on 361 ventures. Despite an effective treatment intervention and sample power, we find evidence that structured accountability affects venture performance heterogeneously. Our post-hoc analyses suggest that structured accountability can aid or harm startups as a function of the founders' level of formal education. Moreover, we find that founders value the existence of accountability structures within accelerators. Furthermore, we find a contradiction between what entrepreneurs need and what they want. Our qualitative analysis shows that, even though structured accountability has detrimental (or at least null) effects on the ventures of founders with high education, these founders appear to want more structured accountability instead of less of it. In contrast, low education founders seem to prefer less structured accountability, although they tend to benefit a great deal from it.

***Acknowledgments**: The authors thank Anna Valero and participants of the Innovation Growth Lab 2020 Global Conference for thoughtful and constructive feedback. We gratefully acknowledge the financial support of IGL, Kauffman Foundation and ANID (The Chilean Agency for Research and Development) through the FONDECYT Iniciación Project number 11170537. The support of Start-Up Chile and CORFO was invaluable for a successful implementation of the experiment. We recognize the support of David Fernández, Rocío Fonseca, Sebastián Díaz, Ángeles Romo and all the Start-Up Chile team. We are particularly grateful to Macarena Gonzalez, for her outstanding research assistance. Also, José Guridi, Rocío Ortega and Santiago Reyes provided helpful assistance. The collaboration of so many participants, mentors and scribes was key and is greatly appreciated.*

*Corresponding author

This research was funded in part by the IGL Grants Programme, the Ewing Marion Kauffman Foundation, the Argidius Foundation, and Nesta. The contents of this publication are solely the responsibility of the authors.

1. INTRODUCTION

Business accelerators have become a relevant and effective organizational form to shape entrepreneurial ecosystems and support new ventures. In addition to becoming a prevalent institution in most entrepreneurship hubs across the globe (as of 2021 CrunchBase lists almost 3,000 accelerators worldwide), accelerators have garnered much scholarly interest (Cohen *et al.*, 2019b; Gonzalez-Uribe and Leatherbee, 2018; Hallen, Cohen, and Bingham, 2020; Yu, 2020). Early work focused on the effects of accelerators on new venture performance, finding consistent evidence to support their role as an organizational sponsor that supports and promotes entrepreneurship (Armanios *et al.*, 2017; Gonzalez-Uribe and Leatherbee, 2018; Hallen *et al.*, 2020). This was important work because entrepreneurship is responsible for a large proportion of new job creation (Haltiwanger, Jarmin, and Miranda, 2013), and providing evidence about the positive effects of accelerators is a fundamental building block for further research. After establishing the causal effects of accelerators as an instrument for endowing participant ventures with entrepreneurial capital,[2] scholars shifted attention towards understanding the specific mechanisms that drive this phenomenon. This was important because of the mounting evidence that design features of accelerators are responsible for the heterogeneity in the impact they have on new venture performance (Chan, Patel, and Phan, 2020; Hallen *et al.*, 2020), such as broad, intensive and paced consultation practices (Hallen *et al.*, 2020), customized advice and visibility (Gonzalez-Uribe and Reyes, 2021), regular updates with program managers (Cohen *et al.*, 2019b), and feedback about the viability of business ideas (Leatherbee and Katila, 2020; Yu, 2020).

[2]While managerial capital has been found to improve the exploitation of existing businesses (Bloom and Van Reenen, 2010; Bruhn, Karlan, and Schoar, 2010), entrepreneurial capital improves the discovery and capture of new ventures (Gonzalez-Uribe and Leatherbee, 2018).

Prior literature has also highlighted the parallels between accelerators and business schools, as both have broadly similar features of capability-building and certification (Armanios *et al.*, 2017; Dutt *et al.*, 2016; Gonzalez-Uribe and Leatherbee, 2018; Hallen *et al.*, 2020): a competition for admittance (sorting); acquisition of knowledge about organizational practices (learning); access to unique information (networking); periodic examinations and a graduation challenge (accountability structures); and certification (signaling). While this prior literature has explored the broad features of sorting, learning, networking, and signaling, little is known about the feature of accountability on new venture performance.

We focus on studying the causal relation between structured accountability (as a business accelerator feature) and new venture performance. Specifically, we ask, "How does structured accountability affect new venture performance?" and "What are its boundary conditions?" By structured accountability we mean supervisory structures that prime or coerce entrepreneurs to report on their periodic progress, and to commit to execute specific tasks. It is important to answer this question because prior research has theorized about features that are conceptually close to structured accountability – such as regularly describing to program managers what founders have learned (Cohen *et al.*, 2019b), providing temporal structures that encourage entrepreneurs to converge quicker towards decisions and actions (Hallen *et al.*, 2020), and standardizing programmatic activities (Cohen, Bingham, and Hallen, 2019a) – and have called for stronger evidence about the features' causal effect on venture performance. Moreover, our paper distinguishes itself from prior accelerator literature that focuses primarily on the relation between learning and decision-making. We do so by answering calls to better understand the psychology-based features that accelerators can leverage for the benefit of new ventures (McKenzie, 2020) and contributing to the literature studying the psychological mechanisms that affect strategic decision-making (Grégoire, Corbett, and Mcmullen, 2011; Grimes, 2017).

Our paper is also relevant for practice. Firstly, because the race of accelerators to establish a reputation of actually increasing new venture performance attracts higher quality applicants, thus pushing accelerators into a self-sustaining virtuous cycle (Hallen *et al.*, 2020). Secondly, many accelerator managers are experimenting with new feature designs, few of which are evidence-based. This trial-and-error approach can lead to superstitious learning and unwittingly perpetuate the use of ineffective acceleration practices.

To answer our research question, we conduct a two-year randomized controlled trial field experiment with 361 business accelerator startups. Because new venture performance can express itself in different ways depending on its goals and business model, we keep with prior literature (Gonzalez-Uribe and Leatherbee, 2018; Hallen *et al.*, 2020) and we measure multiple ex-post performance variables. We complement our empirical results with a qualitative analysis based on in-depth semi-structured interviews of a random subsample of 22 startups. Moreover, since prior literature has found that accelerator features can affect new ventures in different ways, for example based on their inherent quality (Gonzalez-Uribe and Reyes, 2021), we conduct a post-hoc exploratory analysis and find treatment heterogeneity across founder education level.

Despite the intuitive appeal for policy-makers and the strong theoretical support for using structured accountability in firms (Garg, 2013; Lerner, 1995; Sapienza and Gupta, 1994; Westphal, 1999), we do not find a causal relation between accountability structures and new venture performance. However, our findings do suggest that structured accountability increases the performance of ventures whose founders have relatively lower education levels (i.e., up to undergraduate studies), but seems to be detrimental for new ventures whose founders' have higher levels of education (i.e., graduate studies). Our findings open the door for future research to explore potential heterogeneous effects of structured accountability.

Interestingly, there appears to be a contradiction between what entrepreneurs need and what they want. Our qualitative analysis shows that, even though structured accountability has detrimental (or at least null) effects on the ventures of founders with high education, these founders appear to want more structured accountability instead of less of it. In contrast, low education founders seem to prefer less structured accountability, although they tend to benefit a great deal from it.

Our findings contribute to at least four literature streams. Firstly, to the growing literature around business accelerators. Specifically, how to better design the features of programs aimed at fostering entrepreneurship (Gonzalez-Uribe and Leatherbee, 2018; Hallen *et al.*, 2020), by showing the first causal estimates of the impact of structured accountability on venture performance, and delving into the use of psychological mechanisms to drive entrepreneurial performance (McKenzie, 2020). Secondly, to a better understanding of the boundary conditions of supervisory structures (Barney *et al.*, 1996; Daily *et al.*, 2002; McGrath, 2001), sharpening the limitations of new venture oversight. While there is growing literature on new venture boards and their interaction with founder-CEOs (Garg and Eisenhardt, 2017; Garg and Furr, 2017; Garg, John Li, and Shaw, 2019), there is little empirical and causal evidence about ways to improve such interactions. Thirdly, to the discussion about new venture founding team characteristics (Beckman, 2006; Boeker and Wiltbank, 2005; Eisenhardt and Schoonhoven, 1990; Furr, Cavarretta, and Garg, 2012; Vissa and Chacar, 2009). Prior studies have extensively examined how new venture performance is related to the characteristics of individual entrepreneurs (Brockhaus, 1980; Hsu, Roberts, and Eesley, 2007; Shane, 2000), and we contribute to the emerging literature about the relation between new venture teams' educational background and the entrepreneurial process (Leatherbee and Katila, 2020). Fourthly, to recent calls to provide more evidence-based insights to strategic management (Chatterji *et al.*, 2016), by providing, to the best of our

knowledge, the first randomized controlled trial to test the effects of structured accountability on new venture performance.

The rest of our paper is structured as follows. We provide a background to frame our work, after which we describe the research setting and methods, and empirical setup. Next, we present our results, followed by a discussion and concluding remarks.

2. BACKGROUND AND MOTIVATION

2.1. Business Accelerators and New Venture Performance

How to improve entrepreneurial performance is a question that intrigues new venture stakeholders across the board. Despite the importance of new ventures for economic growth (Haltiwanger *et al.*, 2013), and the relevant public and private resources spent to foster entrepreneurial activity, little is known about which specific interventions actually speed up new venture performance.

Currently, one of the most popular institutional forms aimed at supporting new ventures are business accelerators. They typically are a fixed-term, cohort-based, financial intermediary that offer startups cash, shared office space, and business education. From only one in 2005—Y Combinator—thousands now exist worldwide, with governments sponsoring approximately 18% of the programs (Lewis, Harper-Anderson, and Molnar, 2011). Accelerators distinguish themselves by their strong emphasis on the business-education component (Cohen and Hochberg, 2014). The main purpose of this component is to guide entrepreneurs down the path that will lead to greater performance faster, a path that some entrepreneurs will not naturally follow on their own accord, possibly due to information constraints (e.g., Leatherbee and Katila, 2020).

While testing the causal effects of accelerators has recently drawn the attention of entrepreneurship scholars (Bernthal, 2015; Fehder and Hochberg, 2014; Gonzalez-Uribe and Leatherbee, 2018; Hallen *et al.*, 2020; Yu, 2020), questions still remain regarding the specific

programmatic features or mechanisms through which business accelerators affect new venture performance. Recent work has found causal evidence about the performance-enhancing effects of accelerator's role as "entrepreneurship school" (Gonzalez-Uribe and Leatherbee, 2018; Gonzalez-Uribe and Reyes, 2021). However, the entrepreneurship school has, thus far, been treated as a "black box" and it is not yet clear what are the exact mechanisms of the school that drive the superior performance.

Based on the education literature, we distinguish two broad potential mechanisms within the entrepreneurship school (see Figure 1): productivity increases (Becker, 1975) and certification (Arrow, 1973; Spence, 1973), both of which have been found to be present in similar entrepreneurship-promotion institutions (Armanios *et al.*, 2017). Productivity may increase via the instruction of entrepreneurship know-how from peers and staff (Hallen *et al.*, 2020; Lerner and Malmendier, 2013), access to valuable social networks (Granovetter, 1973; Ketchen, Ireland, and Snow, 2007), the structured accountability imposed by regular meetings (Cialdini and Goldstein, 2004) and increases in the self-efficacy of founders (Bandura, 1982; Forbes, 2005).

Regarding the accountability component of entrepreneurship schools, Gonzalez-Uribe and Leatherbee (2018) found suggestive evidence that structured accountability could play a role, but had no rigorous evidence. This paper takes a step towards opening the black box of the entrepreneurship school embedded in the business accelerator model, by specifically testing the causal effects of structured accountability on new venture performance. This focus is important because it contributes to an unexplored feature in the literature of business accelerators, and delves into the psychological mechanisms that could be used to improve new venture performance (McKenzie, 2020). Moreover, it can be a cheaper mechanism to implement relative to other features, such as mentors.

2.2. Structured Accountability

By “structured accountability” we mean the process by which entrepreneurs are exogenously encouraged to articulate to a third party (i.e., an accelerator staff member, a board member, or a peer), on a regular basis, the strategic activities they consider important to work on during a given time period (e.g., a month), and how they fared with the tasks they committed to during the previous period (i.e., their progress, achievements and challenges). Thus, by relying on preemptive self-criticism in anticipation of opinions by others (Tetlock and Henik, 2005), social pressure to execute the declared committed tasks (Thaler, 2000), and periodic deadlines (Ariely and Wertenbroch, 2002; Latham and Locke, 2006), the structured accountability policy may guide entrepreneurs to perform better than the alternative of leaving them to their own volition.

Accountability has been defined as the “expectation that one may be called on to justify one’s beliefs, feelings, and actions to others” (Lerner and Tetlock, 1999: p. 255). As this definition suggests, accountability is a broad term that encompasses different actions or situations, such as the mere presence of others, identifiability, explanation-giving, and evaluation (Tetlock and Mellers, 2011). In established organizations, accountability is typically reflected in performance evaluations, employment contracts, and reward systems, among other procedures (Frink and Klimoski, 2004). In the case of new ventures, such accountability structures are rare, as organizational routines are seldom in place, and the main drivers of the organization are the founders, who, by construction, rarely have a superior to report to.

2.3. Should Structured Accountability Improve or Hinder New Venture Performance?

In many established organizations, accountability is taken for granted as a critical element of management. Monitoring, one of the main functions of board of directors and managers (Shapiro, 2015), creates accountability (Garg, 2013), which in turn helps keep organizational

members aligned with the organizational goals (Brickley, Zimmerman, & Smith, 2008; Zajac & Westphal, 1994). However, it is unclear whether new ventures benefit from structured accountability.

McGrath's (2001) findings shed some light on to the question. In the case of new projects in established organizations, managerial oversight decreased learning for those projects dealing with higher levels of uncertainty. Indeed, there are reasons to believe that imposing accountability structures on founders of ventures that have yet to become consolidated may not be a good idea. Venture CEOs do not always find board advice valuable (Barney *et al.*, 1996; Ehrlich *et al.*, 1994), which suggests that perhaps, under certain conditions, founders would be better off spending their time running the firms rather than reporting to a board. Moreover, prior literature has found that accountability negatively affects important entrepreneurial behaviors. For example, it fosters 'tunnel vision,' whereby individuals focus on specific issues accounted for while omitting other non-accounted factors of equal importance (Ossege, 2012). As such, accountability often reduces creativity and knowledge exchange (Son, Cho, and Kang, 2017). When startups are conducting exploratory efforts, which are important for learning processes and typically antecede exploitation processes needed once the business idea has been validated (Leatherbee and Katila, 2020), managerial oversight appears to obstruct the important learning that is required during the discovery process.

In contrast to the negative perspective of accountability on new venture performance, the psychology literature views it in a more positive light. An interesting perspective in accountability comes from judgment and choice research, which shows that human decisions are prone to a number of cognitive biases (Bazerman and Moore, 2013; Kahneman, 2011). In this context, there are two ways by which accountability can enhance new venture performance. First, accountability has been proposed as a prescriptive strategy to debias

decisions (Larrick, 2004). Indeed, accountable decision makers tend to use more systematic decision strategies (Kausel *et al.*, 2015; Tetlock, 2000), decrease their overconfidence (Tetlock and Kim, 1987), and make fewer sunk costs errors (Simonson and Staw, 1992). In particular, accountable individuals engage in "preemptive self-criticism" by evaluating their judgments and decisions critically and anticipating counterarguments of potential critics (Connolly, Reb, and Kausel, 2013; Tetlock, 1983). As a result, they generally improve their judgment and decision-making processes. Second, accountability is important when people plan tasks and set deadlines (Latham and Locke, 2006). For example, studies by Ariely and Wertenbroch (2002) suggest that when individuals are accountable of delivering their work progressively (i.e., at regular intervals), they tend to improve their performance. Moreover, having externally imposed deadlines and assessments can help individuals avoid spending effort on activities that may appear urgent or easy to complete, but are strategically trivial.

These two apparently competing views—which claim that structured accountability can impair or boost performance—make it hard to anticipate the effects of structured accountability on new venture performance.

3. RESEARCH SETTING

We worked closely with Start-Up Chile, an accelerator sponsored by the Chilean government that was introduced in 2010. It was one of the first ecosystem accelerators world-wide, and it has been recognized as one of the biggest accelerators in terms of the number of startups supported. Its main aim is the attraction of early-stage, high-potential entrepreneurs from across the globe, and the transformation of the domestic entrepreneurship ecosystem. Towards the end of 2022, approximately 2,200 startups had participated in the program, and nearly 16,000 had applied.

Like other business accelerators worldwide, Start-Up Chile is a fixed-term, cohort-based program. Once per semester, Start-Up Chile accepts roughly 90 startups into the

program, who coexist in the accelerator during a six-month tenure. Each startup receives an equity free grant of roughly $30,000 US dollars from the government. The selection process is based on the relative quality of the submitted application, as evaluated by external judges. At the end of their term, participating startups "graduate" through a "demo day" competition (i.e., a formal presentation of the companies to external investors).

3.1. The provision of structured accountability as a feature of accelerators

One of the features of Start-Up Chile that we exploit in this study is a standardized activity (see Cohen *et al.*, 2019a) provided regularly to participants. We focus on a specific type of standardized activity: roughly seven monthly meetings throughout the program (see Table 1). These meetings were held in English, which is the most common language spoken by participants in the accelerator. Participation rates are high and considered good "citizenship" (Section 4 discusses participation in detail).

The first type of meeting is known as *Platoons*. These meetings are moderated by a program staff *executive*, include four to eight program peers, and occur in roughly two-month intervals (see Table 1). Executives are assigned roughly 2 platoons per cohort. The assignment occurs at the start of the program based on industry and does not change during program roll-out. The purpose of the Platoons is to encourage the exchange of experiences and lessons among the entrepreneurs and are moderated by the corresponding executive.

The second type of meeting is known as *Boards.* These meetings are held between the lead founder and a panel of two to three industry experts who act as proxy board members or "advisors." Industry advisors are matched to startups such that participants meet with the same industry experts each time. No other program peers attend these meetings. The purpose of the Boards is to provide team leaders with customized feedback. Boards are also held at two-month intervals, interspersed with the Platoon meetings such that entrepreneurs participate in either the Platoon or the Board activity roughly once per month (every 3-4

weeks). Due to program-level considerations, boards were implemented starting at cohort 19, replacing half the Platoon sessions such that all cohorts had roughly the same number of standardized meetings.

Participant startups are not allowed to change from their assigned platoon nor attend other platoon meetings. Thus, for a given entrepreneur, its executive, board advisors and Platoon peers remain constant throughout the program. Platoon and board sessions last 83 minutes on average, but there is variation across cohorts. The average length in cohorts 18 and 21 was roughly 110 and 70 minutes on average (see Table 1).

3.2. Sample

Our sample includes 369 startups across five cohorts of the program (cohorts 17-21). The experiment took place between February 2017 and August 2019 (the application and graduation dates for each cohort are summarized in Table 1). On average, each cohort includes 5 executives and the average number of startups per executive is 10. Each cohort includes between 12 to 14 Platoons, and each Platoon includes between 5 to 8 startups.

We pooled five cohorts to increase statistical power: the number of startups that participated in each cohort ranged between 54 and 85. Pooling makes sense in our context because the standardized activities offered by Start-Up Chile did not vary substantially across cohorts. Naturally, it is still important to control for potential differences across cohorts (for example, the average quality of participants or skills of the program staff participating in each cohort), and we do this in several ways as we explain in more detail in our empirical strategy (Section 4).

Start-Up Chile provided us with all the application data, including application scores, for the participants in these five cohorts. Based on the program's records, we constructed six covariates to use as controls: age of the lead founder (Age Team Leader), indicator variables for domestic and female applicants (Chilean, Female), the natural logarithm of the number of

employees (Initial Employees), the natural logarithm of the (monthly) sales (Initial Sales), a discrete variable indicating the stage of the applicant (Initial Stage),[3] the application score (Score) and indicator variables for capital raised before application to the program (Capital Raised Before), and for highly educated founders (High Education)—i.e., Masters and above.

Table 3 provides summary statistics of our sample. Roughly 47% of participant startups have raised external financing prior to their application, the average number of employees is 4.83, and has monthly sales of $9.52 thousand US dollars. Participants are concentrated in information technology related sectors—IT & Communication (17%), Education (11%), and Health (10%). The proportion of Chilean participants is 31% (see Appendix 2 for a more detailed industry breakdown). Most founders are between 25 and 35 years old (average age is 31), and the proportion of women in the program is 22%.

Our sample is comparable to prior research on early-stage ventures, particularly in terms of the number of employees (e.g., Haltiwanger *et al.*, 2013) and industry representation (e.g., Puri and Zarutskie, 2012). Our sample is also comparable to the ecosystem business-accelerator genre (Clarysse, Wright, and Hove, 2015). Using information from the Entrepreneurship Database (ED) program at Emory University, which has records of multiple ecosystem accelerators worldwide, we report in Appendix 1 at-application comparisons between the startups (founders) in our sample and those of the ED database (reported under the heading "ED"). The tables show that, relative to average applicants in other ecosystem accelerators worldwide, the average Start-Up Chile participant is younger, less likely to be female, has a younger and more underdeveloped business, and is less likely to have raised capital prior to potential participation.

[3] A self-reported variable of 4 categories: Concept, Prototype, Functional Product, Scaling Sales.

4. EXPERIMENT DESCRIPTION

We begin by describing the assignment of participants to groups, then discuss details of the treatment and control offerings. Finally, we describe the mechanisms we used to ensure the proper implementation of the experiment.

4.1. Random assignment

We used a stratified randomization (Edovald and Firpo, 2016) with a 1:1 allocation ratio (see Figure 2), based on the executive assigned to the startup. The executive assignment was done according to industry – a criteria required by Start-Up Chile. The benefit of this randomization approach is the reduction of the intrablock variance, which makes treatment estimates more accurate because of increased statistical power and precision of test statistics (Ariel and Farrington, 2010).

All participants in each cohort were classified by Start-Up Chile according to industry and assigned to executives on that basis. Executives can have more than one industry assigned to them, and all founders in the same industry were not necessarily assigned to the same executive because of executives' capacity constraints. For a given executive, we randomly allocated the founders to the Platoon treatment and control groups. The treatment assignment was also extended to Boards. That is, all participants in each Platoon have the same treatment assignment, and all participants in the treatment Platoons also were assigned treatment Boards.[4]

[4] The only exception to this assignment rule occurred during cohort 18 where the assignment to treatment was done at the Platoon rather than at the individual level. This occurred because unbeknownst to the principal investigators, Start-Up Chile assigned participants jointly to executives and Platoons at entry. Therefore, when we implemented the random assignment, we had to randomize the entire Platoon, rather than the individual participants as we did in the other cohorts. As a robustness check, we show in the Appendix that results are similar if we drop cohort 18 from the regressions.

Table 1 (panel A) shows the distribution of startups into treatment and control groups across cohorts (row 7). The assignment is roughly 1:1 (except in the cases where the total number of participants is not an even number). The number of Platoons per executive varies across cohorts: 3 for the first two cohorts, and 2 for the rest. Note that Platoons have different sizes in each cohort, and the average size varies across cohorts too (row 6). However, the number of Platoons and size of Platoons across treatment groups is virtually identical in every cohort (rows 4 and 6). For cohorts 17 and 18 there is an imbalance in the number of Platoons by treatment assignment per executive: some executives had 2 treatment and 1 control groups, whereas some had 2 treatment and 1 control group. The executives in both cohorts are the same individuals, so we made sure to flip the imbalance from one cohort to the next.

4.2. Randomization checks

We use the application data to provide evidence that the process resulted in treatment and control groups that are comparable in terms of their initial characteristics. Table 4 shows that the groups are balanced in terms of: lead founder education, gender and age, firm location, and capital raised at application. But some differences remain in spite of the randomization: the treatment group has a slightly lower level of initial sales, initial number of employees and application scores. Table 4 reports differences in characteristics at application across groups controlling by strata (we include cohort-executive fixed effects and industry fixed effects—as we explain in more detail in Section 6 below; results are quantitatively similar if we use simple t-tests that do not control by strata). The coefficients in the table are based on log transformations of the variables. Relative to the unconditional means, they imply that treated firms have on average $6.5K lower initial sales and 0.5 fewer initial employees.

There are several explanations for these small, yet statistically significant differences. First, the randomization was blind to founder and firm characteristics: we had no access to the application data at the moment of the randomization. Second, the table pools results from

several individually randomized cohorts. In unreported analysis, we find that the differences are driven by cohort 19.

In the main analysis, we show that the differences in initial sales, employees and scores appear immaterial: results are robust to the inclusion of these variables as controls in the regression. In addition, we show in robustness checks that results remain quantitatively similar if we exclude cohort 19 from the sample (see Section 7.3).

4.3. The structured accountability intervention

We build our experiment on the customized feedback structure provided by Start-Up Chile. Both groups were offered the Platoons and Board meetings and were blinded to treatment status. The key difference was that treated lead founders (those subject to structured accountability) were asked to articulate the strategic tasks to be completed during the following weeks until the next platoon (board) meeting, and report about their progress since the last platoon (board) meeting.

In detail, during each Platoon meeting, executives asked treated founders two specific questions: "how was your progress on the committed tasks since our last meeting?" and "what would you say are the key strategic tasks you need to work on until our next meeting?" During Board meetings, these questions were asked by the board members as requested by Start-Up Chile. The only role program executives played in this case was to provide board members with the founder's list of previously committed tasks. Treatment participants are expected to think about these questions freely, and provide answers based on their own understanding of what "good progress" (or the "best strategy") actually means. Executives (board members) did not ask these questions to the Platoon (Board) control group. This means that while both control and treatment groups have the chance to get customized feedback from the program peers and the advisory board, only the treated entrepreneurs were asked these two specific questions in their meetings.

There were several reasons for offering the control group the standardized meetings rather than not providing any meetings at all. Firstly, from Start-Up Chile's point of view, the regular meetings provide a control mechanism to oversee the proper use of the grants by participants, and therefore not meeting control participants was out of the question.

Secondly, from a causal assessment standpoint, offering both groups the standardized meetings lowers the risk of Hawthorne and John Henry effects, since both groups were told they were provided the same service in the Platoon and Board meetings. By having similar regular meetings that had a subtle but important difference (the questions and discussion about the progress of the committed tasks), we minimized the risk of treatment spillover.

Finally, the provision of structured accountability within the context of a meeting should be more powerful than on its own (e.g., by simply filling out a form), due to the way the intervention should induce preemptive self-criticism in anticipation of opinions by others (Hoorens, 1993; Leary and Kowalski, 1990), and social pressure to execute the declared committed tasks (Thaler, 2000).

4.4. Mechanisms to ensure proper implementation of the experiment

Proper implementation of the structured accountability experiment depends on two main success conditions. First, participation of treatment and control participants in the Platoon and Board meetings. Second, correct implementation of the structured accountability treatment (and absence of it) in the treatment (and control) groups, conditional on meeting participation.

The conditioning of grants' disbursement to meetings' participation sets strong compliance incentives for the first condition. Consistent with this intuition, Table 1 shows that participation compliance is very high and is no different between treatment and control groups. Panel A (row 10) shows that average participation rates almost always match the mandatory number of meetings per cohort. Attendance to all meetings was mandatory in

cohort 17 (Table 1 shows virtually perfect compliance for that batch). Starting at cohort 18, lead founders were allowed to skip one meeting, and consequently the average difference between number of attended meetings relative to planned meetings (6) is one. Starting at cohort 19, lead founders could skip two meetings—one for each type of meeting (recall Boards are introduced in cohort 19); thus, the difference between planned and attended meetings in that cohort is two. All participants, regardless of participation compliance, are included in the sample to avoid participation bias.[5]

To ensure the compliance of the second condition—correct implementation of the structured accountability treatment (conditional on participation in the meeting), we used several diverse mechanisms that we now describe in detail. First, we ran a pilot intervention on Start-Up Chile's 16th cohort to align the expectations between the research team and Start-Up Chile's staff. The pilot lasted 6 months between August 2016 and January 2017. In the trial, we assigned two Platoons to the treatment condition and used two as the control. Based on the Pilot's experience, we initiated several protocols to ensure proper implementation that we used throughout the experiment, including separate training sessions with the program staff (the executives) and the board members to explain the implementation design. In addition, we developed an *Operations Manual* to be used as guidance by the staff and that included answers to frequently asked questions.

Second, we hired an experiment implementation team of local assistants to implement and oversee the execution of the experiment. The team included a local coordinator that ran the training sessions with executives and board members. The scribes were present at each

[5] In post-hoc analysis, we split the sample between high and low-education lead founders. Panel B of Table 1 shows no differences in participation between firms in the subsamples with high and low education founders. No differences are visible either between treatment and control firms for both the high and low education sub-samples.

one of the meetings, taking notes for both control and treatment groups, and registering and reporting whether the treatment (control) group entrepreneurs were (not) asked the structured accountability questions. In addition, scribes took detailed notes of the articulated tasks, which were revised and discussed during the next meeting. The control group participants did not have any kind of explicit tracking of commitments that were spontaneously brought up by entrepreneurs during the meetings. Scribes were trained for the job ahead of each cohort and required to sign a confidentiality agreement to safeguard the privacy of the entrepreneurs and avoid disclosure about the experiment.

Third, we designed and measured a series of key performance indicators (KPIs), which we revised on a regular basis. Table 1 summarizes the average KPIs for each cohort. Panel A shows that errors in implementation were uncommon (row 11): virtually all meetings held were correctly implemented. The same panel also shows that the few implementation errors are equally distributed across treatment and control groups.[6]

Panel A in Table 1 summarizes other KPIs for the whole sample. For reference, we note that the average meeting duration is not substantially different across treatment and control groups (row 12). Consistent with the correct implementation statistics, the number of committed tasks is higher for treatment than control firms.

The fourth mechanism to ensure the proper implementation of the experiment, was taking corrective actions in the few instances that control group entrepreneurs did receive the treatment (because the platoon leader mistakenly asked the treatment question to the control group). These instances were reported by the scribes to the principal investigator, who took

[6] By "implementation errors" we mean that a control group was given the treatment and vice versa.

swift corrective action (such as talk with the platoon moderator or the program director to avoid future mistakes).

Finally, we conducted novel analyses to ensure the correct implementation of the structured accountability treatment by testing the perceptions of participants. Starting on cohort 18, towards the end of venture's tenure in the program, we asked all participants a series of questions meant to elicit their perceptions and experiences regarding their participation in both *Platoons* and *Boards*. Table 2 provides compelling evidence that the mechanisms to ensure proper implementation were successful. Panel A shows evidence that treated participants perceive greater structured accountability than control participants. By contrast, treated and control participants report similar perceptions in terms of feedback. It is interesting to note that control group participants perceived a request to enunciate goals and follow-up on goals (4.3 and 4.2 on a 1-5 Likert scale), even though control-group executives did no such thing, and even though the treatment group perceived a significantly higher structured accountability than the control group. Panel B repeats the analysis splitting the sample into the subsample of firms with high-education and with low-education lead founders. The results in these panels provide compelling evidence that the mechanisms to ensure proper implementation in both groups were equally successful in both sub-samples.

5. OUTCOME VARIABLES

Collecting performance measures for all participants to the accelerator is challenging. The vast majority of participants are not registered in standard (local or foreign) business data sources. Moreover, the program collects participant data only irregularly. Therefore, we use three strategies to address this challenge. First, similar to prior research, we hand-collected internet-based performance measures for all participants (Goldfarb, Kirsch, and Miller, 2007; Gonzalez-Uribe and Leatherbee, 2018; Hallen *et al.*, 2020; Kerr, Lerner, and Schoar, 2014). Second, we relied on surveys that we co-designed and co-implemented with program staff.

Greater details about this data-collection strategy and the definitions of each outcome variable can be found in Appendix 4. Finally, we conducted in-depth interviews with a sub-sample of participants. All outcomes are measured within 2.75 years since application to the accelerator.

5.1. Quantitative measures of venture performance

For our internet-based measures, we searched through the Crunchbase and LinkedIn platforms at the end of 2020. Because participants in cohorts 17-21 applied to the program between February 2017 to February 2019, these metrics represent new venture performance outcomes between 0.75 and 2.75 years since application into the program.

We also conducted surveys in October 2019 (cohort 17) and November (cohorts 18-20) of 2019, and January (cohort 21) of 2020 (see Appendix 4 for a detailed explanation). In each survey, we asked participants about their performance after graduation at six-month intervals. For example, for cohort 17 participants that were surveyed in October 2019, we asked about their performance by 0.5, 1, 1.5 and 2 years after graduating in August 2017. Depending on the cohort, we have performance data for a varying number of semesters after graduation from the program (1 semester for cohort 21, and 4 semesters for cohort 17).

For each data source, we constructed five new venture performance proxies: the natural logarithm of the number of *Employees*; *Capital Raised* measured as a binary variable for securing capital after participation in the accelerator; the natural logarithm of the *Amount Raised* after participation and excluding the seed capital provided by the program; *Market Traction* measured as the natural logarithm of the sales (or LinkedIn Followers-results) during the six preceding months; and a binary variable to indicate *Survival*. We used logarithmic transformations of continuous outcome variables to mitigate the potential impact of outliers; we add one to the variable before the transformation and in unreported analysis

show that results are robust to adding the minimum (non-zero) value of the respective variable. Table 3 presents the summary statistics of the five internet-based outcome measures.

5.2. Interviews

To better understand the underlying mechanisms at play, we complement our quantitative results by conducting in-depth, semi-structured interviews with 22 participants of cohort 19. We randomly selected 15 entrepreneurs from the treatment group and 15 from the control group. Four selected participants from each group declined our invitation to participate in the interviews, leaving us with 11 interviews for each group. Seventeen interviews were conducted in English and five in Spanish.

One of the research team members conducted all 22 interviews, unbeknownst to the assignment to treatment of each interviewee. Interviews were roughly one hour long, and a conversation guide was used to ensure that all issues of interest were covered while maintaining flexibility to tailor questions depending on the interviewees' responses. The interviews included general questions about the experience in the program, as well as the experience in the platoon and board meetings. As the interview progressed, the interviewer made more specific questions about the effectiveness of the meetings and gathered the interviewee's opinion about the content and structure at the meetings. In addition, the interviewer inquired about the entrepreneur's feelings regarding the sharing of experiences and the listening to the experiences of others, as well as receiving positive or negative feedback. The interviewer also asked entrepreneurs about the usefulness of the meetings in terms of how these meetings shaped the outcome of their ventures. Because the interviews were semi-structured, not all questions were explicitly articulated, particularly if the interviewee addressed the question on their own volition. Moreover, interviewees were allowed to discuss other aspects that may not have been capturable by the predefined questionnaire, thus providing further potentially valuable insights.

All interviews were recorded and transcribed by experts external to the research team. The data was analyzed using a four-stage coding technique employing MAXQDA 12 software. Firstly, we generated codes organically as specific concepts surfaced from reading the transcripts. Secondly, we classified these codes into broad topics. Six topics clearly stand out: positive and negative aspects of platoon and board meetings, comparison between both types of meetings, and general sentiments towards the business accelerator. The third stage consisted of teasing apart categories within the broad topics, in the cases where such a separation was justified. Finally, we iteratively analyzed the recurrence of the codes and merged codes whenever increases in the robustness of the categorization ensued, after which we counted the frequency with which the code emerged in the transcripts. While other concepts also emerged from the interviews, we only conserved codes with a frequency of 5 or higher. Table 8 shows the resulting six topics and 13 categories.

Only after members of the research team were comfortable with the codification and categorization of the interviews, was the sample separated between treatment and control groups. We then compared the frequency between the experimental groups for each category to identify contrasting patterns between both groups. We discuss the findings in the results section.

6. EMPIRICAL STRATEGY

In this section we discuss the empirical strategy we use to measure the effect of the joint provision of structured accountability and customized feedback, relative to customized feedback alone, on venture outcomes. We start by describing the strategy to estimate the average effect of randomly offering the structured accountability intervention to participants. We next present the strategy used in our post-hoc exploration of the heterogeneity of results across ventures with high and low education lead founders.

6.1. Average effects of treatment assignment

To estimate the average impact of structured accountability and customized feedback relative to customized feedback on its own on firm outcomes, we use the following base specification:

(1) $y_i = \alpha + \beta Treat_i + CohortFE \times Executive\ FE + IndustryFE + Controls_i + \varepsilon_i$,

where the sub-index i denotes firm i, and the variable $Treat_i$ equals one if the firm receives the structured accountability treatment. We include cohort fixed effects interacted with executive fixed effects to reflect the level at which the treatment was randomized among firms. We also include industry fixed effects to control for any residual variation across industries. In some specifications, we add controls for initial sales (in logs), initial employees (in logs), and score to absorb the unintended differences in baseline despite the randomization (see Section 4). We use robust standard errors throughout.

The coefficient of interest is β that measures the impact of being assigned to the treatment group and being offered the possibility to participate in the structured accountability meetings rather than the control meetings. That is, β measures the Intention-to-Treat—ITT (Angrist and Pischke, 2008). Because we include cohort fixed effects interacted with executive fixed effects, β is estimated as the average differences in outcomes between treated and control groups for a given executive in a given cohort. Therefore, the estimate controls for any potential differences across executives in specific cohorts. Likewise, the estimate also controls for any differences across cohorts.

We implement several robustness checks as we explain in detail in Section 7.3. We show that results are robust to a number of changes in the specification of equation (1) or changes in the estimation sample.

6.2. Heterogeneity across high and low education lead founders

We explore the heterogeneity of results across firms with high-education and low-education team leaders. We estimate the potential differential impact of offering structured accountability and customized feedback, relative to customized feedback alone, on high-education and low-education firm outcomes using the following specification:

$$(2)\quad y_i = \alpha + \beta_{Low} Treat_i + \alpha_{High} High_i + \beta_{High-Low} High_i \times Treat_i + CohortFE \times Executive\ FE \times High_i + IndustryFE \times High_i + Controls_i + \varepsilon_i,$$

where the variable $High_i$ equals one if the firm has a high-education team leader. We include interactions of the fixed effects in equation (1) with the variable $High_i$ to control for education differences among entrepreneurs assigned to different executives in a given cohort. Note that such interactions absorb the level effect of $High_i$, which is why the results reported in Table 5 do not report this variable.

We also interact the industry fixed effects with the variable $High_i$ to control for differences in founder education across industries. The variables in $Controls_i$ vary across specifications, but across all models we control for lead founder age. This way the regression captures differences across high and low education lead founders, after absorbing potential differences in age across the groups. In some specifications, we add equation (1)'s basic controls for initial sales (in logs), initial employees (in logs), and score to absorb the unintended baseline differences between treatment and control units despite the randomization. In other specifications, we interact the basic controls with the variable $High_i$ to allow for differences in baseline controls across different education levels, although results in Table 4 show no such differences exist in sales, employees, or scores. Finally, in other specifications, we include controls for the initial stage interacting with the education dummy to absorb the only difference in baseline characteristics for treatment relative to control firms of different founder education (see Table 4).

The coefficients of interest are β_{Low} that measures the average effect of offering the possibility to participate in the structured accountability meetings for firms with low-education founders, and $\beta_{High-Low}$ that measures the difference in treatment assignment between firms with high and low education founders. We also report in Table 5 (and Table 4) the estimate (and p-value) of the assignment to treatment for firms with high-education founders, estimated as $\beta_{High} = \beta_{Low} + \beta_{High-Low}$.

Similar to the base regressions, we implement several robustness checks as we explain in detail in Section 7.3. We show that results are robust to a number of changes in the specification of equation (2) or changes in the estimation sample.

7. RESULTS

In this section we discuss the results from implementing the experiment described in Section 4. We show that the results point to no *average* meaningful differences in outcomes across ventures randomly assigned to treatment and control groups. Moreover, results from the qualitative analysis using the in-depth interviews, show important differences between treatment and control groups in terms of their expectations and general sentiment towards the business accelerator. Furthermore, we show that the average null effects mask substantial heterogeneity of impact across firms with high- and low-education lead founders. The intervention has remarkable positive effects on firms with low-education founders, but little (and sometimes negative) effects on firms with high-education lead founders.

7.1. Effects of structured accountability on venture performance

Table 5 summarizes results based on the internet-based outcome variables. Panel A presents the results from estimating equation (1), while Panel B presents results from estimating equation (2). For each variable, we report results from the various specifications varying the control variables as specified in the last columns of each row. To conserve space, we only

report the estimate for the variables of interest: treatment in Panel A, and treatment for low- and high-education and their difference in Panel B.

The results in Panel A show no meaningful average differences in performance across treatment and control firms. We cannot reject the hypothesis of no effects across the different specifications, and for all outcome variables. Results are similar if we use survey-based variables.

The pattern of results in Panel B show that the lack of results in Panel A mask a large heterogeneity. We find consistent evidence of positive effects on treated firms with low-education founders (column 1), and typically negative, yet not statistically significant effects for high-education founders (column 2). Moreover, column 3 shows that the estimated effects for low- and high-education lead founders are statistically significantly different to each other. The only exception are the results for survival: while they are only statistically significant for the low-education founders, we cannot reject the null that the results are the same across high- and low-education lead founders.

The results in Panel B provide compelling evidence that assignment to treatment has positive and significant venture performance effects for firms with low-education lead founders. This pattern of heterogeneity of results is robust across all outcome variables and specifications considered in the panel. The only exception is the most saturated model for the variable capital raised: the estimate for the effect on low education founders is positive and almost statistically significant (with a p-value of 0.106). This pattern of heterogeneity is also visible in Table 6 that summarizes results using the survey variables. To ease comparison, Table 6 only summarizes results of the most saturated model (last row for any variable in Table 5). Across the different survey-based outcome variables, the point estimate for the effects on firms with low (high) education lead founders are positive (negative), although they are never statistically significant. However, for many variables the estimates in column 3

show that the difference in the effects for firms with high- and low-education lead founders are statistically significant.

As we respectively show in Appendix 5, 6 and 7, these results are robust to (a) excluding industry fixed effects from the regressions, (b) excluding cohort 19, for which the average differences across employees and scores at application between treatment and control groups stronger, and (c) excluding cohort 18 for which the randomization was incorrectly implemented.

In terms of magnitudes, the estimated effects of assignment to treatment on the firms with low-education founders are sizable. The results in Panel B in Table 5 roughly show that assignment to treatment increases, within 2.25 years of participation, the number of employees by 40%, the probability of raising capital by 10%, the amount raised by 32%, the market traction by 33% and survival by 14%. Relative to the unconditional averages in Table 3, these estimates imply an increase of 4 employees, 1 percentage points in the probability of raising capital (from 16% to 17%) and $27K in the amount raised.

The economic magnitude of the effect for firms with low-education lead founders is similar across the different survey- and internet-based proxies. To facilitate comparisons, Tables 6 and 7 present results based on standardized variables. For each variable the results in Table 6 are most directly comparable to the results of the last specification in Table 7. The average of the coefficients in column 1 of the first four rows of Table 6 and the fourth row in Table 7 show that assignment to treatment increases the number of employees by 0.31 (average across surveys) and 0.40 standard deviations, respectively, as measured by the survey and internet proxies. In terms of capital raised, the assignment to treatment increases the amount of capital according to survey proxies and internet proxies by similar amounts: 0.25 (average across surveys) and 0.28 standard deviations, respectively. Finally, the effects

on market traction increases by 0.35 (average across surveys) and 0.36 standard deviations, respectively, for survey- and internet-based proxies.

7.2. How structured accountability is perceived by founders

Results from the analysis of the 22 in-depth semi-structured one-hour interviews are consistent with the results from the quantitative analysis and provide additional insights. Table 8 lists the patterns that emerged from the interviews, as described in Section 5.2.

The second column shows the relative statement frequencies between treatment and control groups. Positive (negative) percentages indicate that treatment (control) participants expressed listed concepts more than control (treatment) participants. The greater the absolute magnitude of the values, the greater the relative difference between groups.

Treated participants were more likely to perceive the presence of structured accountability than control participants (row 2). This is consistent with our experiment implementation tests in Section 4.4. Moreover, treated participants had more positive opinions as well as more constructive criticisms about platoons and board meetings. In contrast, control group participants systematically expressed negative criticisms. For example, treated participants valued the support of the board members and platoon peers (rows 1 and 7), and appreciated the (or desired more) structured accountability (rows 2 and 10). However, they were critical about the lack of organization (row 3 and 8), the value of the feedback (row 4), and the lack of domain expertise (row 3 and 8). In contrast, control group participants felt strong negative emotions about the board meetings, such as frustration, anxiety, and lack of sincerity from the board (rows 6 and 9).

Regarding the platoons, treated participants appear to appreciate the opportunities to interact and learn from their peers, generate new business leads, friendship, and emotional support. However, they were also more critical regarding the organization, lack of clear objectives and heterogenous expertise of their peers. In contrast, control group participants

appear to be very critical about the value of the platoon meetings. Finally, treatment group participants had strong opinions about the favorable support that Start-Up Chile provided regarding the efforts to expand their ventures internationally but felt that the program required too many obligations – likely due to the accountability structure. In contrast, the control group had a strong negative opinion about the high levels of bureaucracy. Overall, participants appear to appreciate structured accountability, want more of it, and have more positive sentiments towards the accelerator.

When comparing the opposing perspectives between high- and low-education founders relative to structured accountability (rows marked with *), a consistent pattern emerges: high-education founders appreciate structured accountability and want more of it, while low-education founders appreciate it less and want less of it. For example, treated high-education founders appreciate structured accountability, while treatment and control low-education founders do not substantially differ (row 2). Also, treated high-education founders expected more from the boards (row 3), even though they were receiving greater structure relative to control high-education founders, and expected much more accountability (row 5). Moreover, structured accountability in the platoons appears to be more appreciated by high-education founders relative to their low-education peers (row 7) and had fewer negative opinions about the organization of platoons (row 8). Furthermore, treated high-education founders did not perceive the structured accountability to be onerous, while their low-education peers appear to find the structured accountability an unwanted obligation (row 11).

Surprisingly, while structured accountability appears to quantitatively benefit low-education founders, these founders appear less appreciative of its existence, relative to high-education founders who qualitatively appreciate it but quantitatively do not benefit from it. We discuss this conundrum in the Discussion section.

8. DISCUSSION

Overall, our results indicate that structured accountability is a tool that can play an important role in new venture performance and in how founders perceive the institutions that support them. On average, we found no effects of structured accountability on new venture performance. However, this null result hides substantial heterogeneity, regarding both venture performance as well as the perceptions of entrepreneurs about the role of the organizational sponsor. We find that a source of this heterogeneity is the education level of venture founders.

In terms of venture performance, our quantitative results point to strong positive effects for low-education founders from the provision of structured accountability. The evidence for high-education lead founders is less conclusive: the estimated effects of treatment assignment are generally negative but are not statistically significantly different from zero. That said, the effects for high-education lead founders are statistically significantly different to the estimated effects for the low-education lead founders.

In terms of perceptions and sentiments towards the organizational sponsor, on average, entrepreneurs in our sample appear to expect the provision of structured accountability. Moreover, they develop negative perceptions and feelings towards the organizational sponsor when it is not provided. However, these perceptions are strongly heterogeneous between low- and high-education founders. Postgraduate founders have positive perceptions about the provision of structured accountability, while low-education founders are more prone to reject or resist it.

Together, our quantitative and qualitative results present an interesting conundrum. While structured accountability improves the performance of ventures led by low-education founders and is detrimental (or at least useless) to high-education founders, the former want less of it while the latter want more. This is an important insight, because it suggests that founders are unaware of (or even incorrect) about what sort of oversight structures are

beneficial for improving the performance of their new ventures. They desire more of something they should have less of.

One way to explain the heterogeneity in treatment effects is the idea that higher levels of education, relative to lower levels, require students to have higher levels of autonomy and self-accountability to be successful (Ariely and Wertenbroch, 2002; Arrow, 1973; Artino and Stephens, 2009). Undergraduate programs tend to have higher structure and "hand-holding" compared to graduate programs which typically demand that students "navigate" curricula with multiple self-guided research assignments. Thus, low-education lead founders may have, on average, lower self-accountability skills, and therefore can benefit more from the structured accountability. In contrast, for high-education founders that already possess certain degree of self-accountability, the provision of structured accountability has no added benefits, and in fact can be detrimental for performance.

A plausible explanation for why some entrepreneurs resist the provision of structured accountability can be that they seek freedom, independence and autonomy. Perhaps that is precisely the reason for their decision to become entrepreneurs. What remains to be answered, however, is why do founders who do not need structured accountability want more of it.

The role of business accelerators

Improving new venture success is key for job and wealth creation in societies (Haltiwanger *et al.*, 2013). An emerging stream of literature has found that organizational sponsors, such as business accelerators, can play an important role in this success by inoculating new ventures from the liability of newness (Stinchcombe, 1965) through the certification and capacity-building (Armanios *et al.*, 2017; Gonzalez-Uribe and Reyes, 2021) provided by entrepreneurship schooling (Gonzalez-Uribe and Leatherbee, 2018). Moreover, it has provided important insights about how specific business accelerator features can help to

mitigate bounded rationality (Cohen *et al.*, 2019a), reduce uncertainty about the viability of business ideas (Yu, 2020), inspire venture teams to come up with new business ideas (Leatherbee and Katila, 2020) and improve venture performance through consultation practices (Hallen *et al.*, 2020) and visibility (Gonzalez-Uribe and Reyes, 2021).

Our findings provide important insights for business accelerators. While the reputation of accelerators improves among participants when structured accountability is provided, this relatively expensive tool is objectively not useful for everyone. Accelerators must learn to navigate the balance between *offering* a service that is expected and appreciated by participating founders, and *providing* such a service only to those who really need it. Identifying those founders who do not need structured accountability and communicating this to them is key to avoid spending time and resources on something that hurts founders even though it makes them happy.

Entrepreneurial freedom or voluntary subjection?

An important characteristic of entrepreneurs is their desire for independence (Aldridge, 1997; Hisrich, 1984; Hornaday and Aboud, 1971; Shane, Locke, and Collins, 2003). However, as Wasserman (2012) describes, entrepreneurs commonly face a tradeoff between independence and venture performance. Our findings provide important new insights about this tradeoff.

For some types of entrepreneurs (relatively lower-education founders in our setting), subjection to accountability structures improves new venture performance. For other types of entrepreneurs (higher-education founders) freedom from such accountability structures appears to be the best choice. Ironically, our qualitative analysis suggests that founders desire more of that which is detrimental to their goals. Founders that are most likely to benefit from structured accountability appear to shun it, while founders that are least likely to benefit from such subjection appear to desire more of it.

Our findings shed light on a specific type of organizational feature, namely structured accountability, that can be very relevant for founders' goal of greater performance, but that can be at odds with founders' desires. It is interesting to compare this conundrum to what Schwartz (2004) called 'the paradox of choice.' Schwartz argued that, although autonomy and freedom of choice are critical to people's wellbeing, societies and individuals who enjoy the most freedom and autonomy do not seem to be benefiting from it. In this paper, we found that while low-education founders had negative attitudes toward structured accountability and lack of autonomy, their ventures were positively affected by it. This conundrum begs an important question: what other types of paradoxes are founders facing (see Ramoglou, 2021)? We believe this is an interesting avenue for future research.

New venture boards

New venture boards of directors have been posited by prior literature as consequential to the outcome of the companies they supervise (Daily *et al.*, 2002). Venture directors are extensively involved in monitoring (Garg, 2013; Lerner, 1995; Sapienza and Gupta, 1994), and board monitoring and venture performance appear to go hand in hand, particularly in established firms (Westphal, 1999). However, it appears that overseeing new ventures is not a one-size-fits-all approach.

Our findings provide important prescriptions for new venture boards. Understanding the boundary conditions and tradeoffs of monitoring new ventures can be consequential for the mission of boards of directors. Knowing when and how to provide structured accountability can benefit both founders and boards.

Limitations

We cannot rule out that the lack of significantly negative effects for high education founders is driven by power issues. We have fewer firms with high-education lead founders than with low-education founders (see Table 1). Moreover, our post-hoc analysis of

heterogeneous treatment effects was not defined ex-ante, which surrounds our findings with a mantle of doubt.

Our results should be considered within the bounds of our research setting, whereas structured accountability (or the lack thereof) was bundled with the provision of customized feedback. Inferences about the potential effects of structured accountability on its own must be treated with caution. Our experiment helps to understand the effects of structured accountability within the context of business accelerators, and care should be used while extrapolating outside our setting.

Self-selection of founders into organizational sponsors should be considered. Founders who rather not apply to business accelerators may have a different relation to structured accountability, both in terms of their preferences and performance. It is not clear whether founders who flee from the accountability structures provided by accelerators are making a self-inflicting decision.

9. CONCLUSION

Through a randomized controlled trial of 361 new ventures in a business accelerator, our study provides evidence-based insights for improving new venture performance, specifically regarding the use of structured accountability. Despite some founders' resistance against–and others' desire for–greater structured accountability, governance structures aimed at supporting new venture performance can be consequential to achieving this goal. We show that the proper application of structured accountability can, albeit counterintuitive to founders, help achieve greater performance. Our findings provide new insights and tools for entrepreneurs, policymakers and new venture boards.

REFERENCES

Aldridge JH. 1997. An occupational personality profile of the male entrepreneur as assessed by the 16PF Fifth Edition. *Dissertation Abstracts International: Section B: The Sciences and Engineering*.

Angrist JD, Pischke J-S. 2008. *Mostly harmless econometrics: An empiricists' companion*. Princeton University Press.

Ariel B, Farrington D. 2010. Randomized block designs. In *Handbook of quantitative criminology*. Springer: New York: 437–454.

Ariely D, Wertenbroch K. 2002. Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science* **13**(3): 219–224.

Armanios DE, Eesley CE, Li J, Eisenhardt KM. 2017. How Entrepreneurs Leverage Institutional Intermediaries in Emerging Economies to Acquire Public Resources. *Strategic Management Journal* **38**: 1373–1390.

Arrow KJ. 1973. Higher education as a filter. *Journal of Public Economics* **2**(3): 193–216.

Artino AR, Stephens JM. 2009. Academic motivation and self-regulation: A comparative analysis of undergraduate and graduate students learning online. *Internet and Higher Education*. Elsevier B.V. **12**(3–4): 146–151.

Ashton RH. 1992. Effects of justification and a mechanical aid on judgment performance. *Organizational Behavior and Human Decision Processes* **52**(2): 292–306.

Bandura A. 1982. Self-Efficacy Mechanism in Human Agency. *American Psychologist* **37**(2): 122–147.

Barney JB, Busenitz LW, Fiet JO, Moesel DD. 1996. New venture teams' assessment of learning assistance from venture capital firms. *Journal of Business Venturing* **11**(4): 257–272.

Bazerman M, Moore D. 2013. *Judgment in Managerial Decision Making*, 8th ed. Wiley: New York, NY.

Becker GS. 1975. Human Capital. A Theoretical and Empirical Analysis with special Reference to Education. *The National Bureau of Economic Research* **3**: 1–22.

Beckman CM. 2006. The Influence of Founding Team Company Affiliations on Firm Behavior. *Academy of Management Journal* **49**(4): 741–758.

Bernthal B. 2015. *Investment Accelerators*. Available at SSRN: http://ssrn.com/abstract=2642436.

Boeker W, Wiltbank R. 2005. New Venture Evolution and Managerial Capabilities. *Organization Science* **16**(2): 123–133.

Brockhaus RH. 1980. Risk Taking Propensity of Entrepreneurs. *Academy of Management Journal* **23**(3): 509–520.

Cao L. 2012. Differences in procrastination and motivation between undergraduate and graduate students. *Journal of the Scholarship of Teaching and Learning* **12**(2): 39-64.

Chan CSR, Patel PC, Phan PH. 2020. Do differences among accelerators explain differences in the performance of member ventures? Evidence from 117 accelerators in 22 countries. *Strategic Entrepreneurship Journal* **14**(2): 224–239.

Chatterji AK, Findley M, Jensen NM, Meyer S, Nielson D. 2016. Field experiments in strategy research. *Strategic Management Journal* **37**(1): 116–132.

Cialdini RB, Goldstein NJ. 2004. Social Influence: Compliance and Conformity. *Annual review of psychology* **55**(1974): 591–621. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14744228.

Clarysse B, Wright M, Hove J Van. 2015. A look inside Accelerators: Building Businesses. *Nesta* (February 2015): 24.

Cohen SL, Bingham CB, Hallen BL. 2019a. The Role of Accelerator Designs in Mitigating

Bounded Rationality in New Ventures. *Administrative Science Quarterly* **64**(4): 810–854.
Cohen SL, Fehder DC, Hochberg Y V., Murray F, States U. 2019b. The design of startup accelerators. *Research Policy*. Elsevier **48**(7): 1781–1797.
Cohen SL, Hochberg Y V. 2014. Accelerating Startups: The Seed Accelerator Phenomenon. *SSRN Journal* : 1–16.
Connolly T, Reb J, Kausel EE. 2013. Regret salience and accountability in the decoy effect. *Judgment and Decision Making* **8**(2): 136–149.
Daily CM, McDougall PP, Covin JG, Dalton DR. 2002. Governance and strategic leadership in entrepreneurial firms. *Journal of Management* **28**(3): 387–412.
Dutt N *et al.* 2016. How open system intermediaries address institutional failures: The case of business incubators in emerging-market countries. *Academy of Management Journal* **59**(3): 818–840.
Edovald T, Firpo T. 2016. *Running Randomised Controlled Trials in Innovation, Entrepreneurship and Growth: An Introductory Guide*. London.
Ehrlich SB, de Noble AF, Moore T, Weaver RR. 1994. After the cash arrives: A comparative study of venture capital and private investor involvement in entrepreneurial firms. *Journal of Business Venturing* **9**(1): 67–82.
Eisenhardt KM, Schoonhoven CB. 1990. Organizational Growth: Linking Founding Team, Strategy, Environment, and Growth Among U.S. Semiconductor Ventures, 1978-1988. *Administrative Science Quarterly* **35**(3): 504–529.
Fehder DC, Hochberg Y V. 2014. *Accelerators and the Regional Supply of Venture Capital Investment*.
Forbes DP. 2005. The Effects of Strategic Decision Making on Entrepreneurial Self-Efficacy. *Entrepreneurship Theory and Practice* **29**(September): 599–627.
Frink DD, Klimoski RJ. 2004. Advancing accountability theory and practice: Introduction to the human resource management review special edition. *Human Resource Management Review* **14**(1): 1–17.
Furr NR, Cavarretta F, Garg S. 2012. Who Changes Course? The Role of Domain Knowledge and Novel Framing in Making Technology Changes. *Strategic Entrepreneurship Journal* **6**(3): 236–256.
Garg S. 2013. Venture boards: Distinctive monitoring and implications for firm performance. *Academy of Management Review* **38**(1): 90–108.
Garg S, Eisenhardt KM. 2017. Unpacking the CEO-Board relationship: How strategy making happens in entrepreneurial firms. *Academy of Management Journal* **60**(5): 1828–1858.
Garg S, Furr N. 2017. Venture Boards: Past Insights, Future Directions, and Transition to Public Firm Boards. *Strategic Entrepreneurship Journal* **11**(3): 326–343.
Garg S, John Li Q, Shaw JD. 2019. Entrepreneurial firms grow up: Board undervaluation, board evolution, and firm performance in newly public firms. *Strategic Management Journal* **40**(11): 1882–1907.
Goldfarb B, Kirsch D, Miller DA. 2007. Was there too little entry during the Dot Com Era? *Journal of Financial Economics* **86**: 100–144.
Gonzalez-Uribe J, Leatherbee M. 2018. The effects of business accelerators on venture performance: Evidence from Start-Up Chile. *Review of Financial Studies*. Oxford University Press **31**(4): 1566–1603.
Gonzalez-Uribe J, Reyes S. 2021. Identifying and boosting "Gazelles": Evidence from business accelerators. *Journal of Financial Economics* **139**(1): 260–287.
Granovetter M. 1973. The Strength of Weak Ties. *The American Journal of Sociology* **78**(6): 1360–1380.
Grégoire DA, Corbett AC, Mcmullen JS. 2011. The Cognitive Perspective in

Entrepreneurship: An Agenda for Future Research. *Journal of Management Studies* **48**(6): 1443–1477.
Grimes MG. 2017. The Pivot: How Founders Respond To Feedback Through Idea and Identity Work. *Academy of Management Journal* **61**(5): 1692–1717.
Hallen BL, Cohen SL, Bingham CB. 2020. Do accelerators work? If so, how? *Organization Science* **31**(2): 378–414.
Haltiwanger JC, Jarmin RS, Miranda J. 2013. Who creates jobs? Small versus large versus young. *The Review of Economics and Statistics* **XCV**(2): 347–361.
Hisrich RD. 1984. The Woman Entrepreneur in the United States and Puerto Rico: A Comparative Study. *Leadership & Organization Development Journal* **5**(5): 3–8.
Hoorens V. 1993. Self-enhancement and superiority biases in social comparison. *European Review of Social Psychology* **4**(1): 113–139.
Hornaday JA, Aboud J. 1971. Characteristics of Successful Entrepreneurs. *Personnel Psychology* **24**(2): 141–153.
Hsu DH, Roberts EB, Eesley CE. 2007. Entrepreneurs From Technology-Based Universities: Evidence from MIT. *Research Policy* **36**(5): 768–788.
Kahneman D. 2011. *Thinking, fast and slow*. Farrar, Straus and Giroux: New York, NY.
Kausel EE, Culbertson SS, Leiva PI, Slaughter JE, Jackson AT. 2015. Too arrogant for their own good? Why and when narcissists dismiss advice. *Organizational Behavior and Human Decision Processes*. Elsevier Inc. **131**: 33–50.
Kerr WR, Lerner J, Schoar A. 2014. The consequences of entrepreneurial finance: Evidence from angel financings. *Review of Financial Studies* **27**(1): 20–55.
Ketchen DJ, Ireland RD, Snow CC. 2007. What makes a process a capability? Heuristics, Strategy and Effective Capture of Opportunities. *Strategic Entrepreneurship Journal* **1**(1): 371–385.
Larrick R. 2004. Debiasing. In *Blackwell handbook of judgment and decision making*, D. Koehler & N. Harvey (ed). Blackwell Publishers: Oxford, UK: 316–337.
Latham GP, Locke EA. 2006. Enhancing the Benefits and Overcoming the Pitfalls of Goal Setting. *Organizational Dynamics* **35**(4): 332–340.
Leary MR, Kowalski RM. 1990. Impression management: A literature review and two-factor model. *Psychological Bulletin* **107**(I): 34–47.
Leatherbee M, Katila R. 2020. The lean startup method: Early-stage teams and hypothesis-based probing of business ideas. *Strategic Entrepreneurship Journal* **14**(4): 570–593.
Lerner J. 1995. American Finance Association Venture Capitalists and the Oversight of Private Firms Author (s): Josh Lerner Source : The Journal of Finance , Vol . 50 , No . 1 (Mar ., 1995), pp . 301-318 Published by : Wiley for the American Finance Association Stabl. *The Journal of Finance* **50**(1): 301–318.
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.
Lerner JS, Tetlock PE. 1999. Accounting for the Effects of Accountability. *Psychological Bulletin* **125**(2).
Lewis DA, Harper-Anderson E, Molnar LA. 2011. *Incubating Success. Incubation Best Practices That Leat to Successful New Ventures*.
Locke EA, Latham GP. 2002. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *The American Psychologist* **57**(9): 705–717.
McGrath RG. 2001. Exploratory Learning, Innovative Capacity, and Managerial Oversight. *The Academy of Management Journal* **44**(1): 118–131.
McKenzie D. 2020. Small Business Training to Improve Management Practices in Developing Countries: Reassessing the Evidence for “Training Doesn’T Work”. *Small*

Business Training to Improve Management Practices in Developing Countries: Reassessing the Evidence for "Training Doesn'T Work" : 1–28.
Murphy R. 1994. The Effects of Task Characteristics on Covariation Assessment: The Impact of Accountability and Judgment Frame. *Organizational Behavior and Human Decision Processes*.
Puri M, Zarutskie R. 2012. On the Life Cycle Dynamics of Venture Capital and Non Venture-Capital Financed Firms. *The Journal of Finance* **67**(6): 2247–2293.
Ramoglou, S. 2021. Knowable opportunities in an unknowable future? On the epistemological
paradoxes of entrepreneurship theory. *Journal of Business Venturing, **36**(2),* 106090.
Sapienza HJ, Gupta AK. 1994. Impact of Agency Risks and Task Uncertainty on Venture Capitalist–CEO Interaction. *Academy of Management Journal* **37**(6): 1618–1632.
Shane S, Locke EA, Collins CJ. 2003. Entrepreneurial motivation. *Human Resource Management Review* **13**(2): 257–279.
Shane SA. 2000. Prior Knowledge and the Discovery of Entrepreneurial Opportunities. *Organization Science* **11**(4): 448–469.
Simonson I, Staw BM. 1992. Deescalation Strategies: A Comparison of Techniques for Reducing Commitment to Losing Courses of Action. *Journal of Applied Psychology* **77**(4): 419–426.
Son SY, Cho DH, Kang SW. 2017. The impact of close monitoring on creativity and knowledge sharing: The mediating role of leader-member exchange. *Creativity and Innovation Management* **26**(3): 256–265.
Spence M. 1973. Job Market Signaling. *The Quarterly Journal of Economics* **87**(3): 355–374.
Stinchcombe AL. 1965. Social structure and organizations. In *Handbook of Organizations*, March JG (ed). Rand McNally: Chicago: 142–193.
Tetlock PE. 1983. Accountability and complexity of thought. *Journal of Personality and Social Psychology* **45**(1): 74–83.
Tetlock PE. 1985. Accountability : A Social Check on the Fundamental Attribution Error Author (s): Philip E . Tetlock Source : Social Psychology Quarterly , Vol . 48 , No . 3 (Sep ., 1985), pp . 227-236 Published by : American Sociological Association Stable URL : http. *Social Psychology Quarterly* **48**(3): 227–236.
Tetlock, P. E., & Henik, E. 2005. Theory- versus imagination-driven thinking about historical counterfactuals.In D. R. Mandel, D. J. Hilton, & P. Catellani (Eds.),*The psychology of counterfactual thinking* (pp. 199–216). London: Routledge.
Tetlock P. E., Mellers B. A. 2011. Intelligent management of intelligence agencies: Beyond accountability ping-pong. *American Psychologist, 66,* 542–554.
Tetlock PE, Kim J. 1987. Accountability and Judgment Processes in a Personality Prediction Task. *Journal of personality and social psychology* **52**(4): 700–709.
Thaler RH. 2000. From homo economicus to homo sapiens. *Journal of Economic Perspectives* **14**(1): 133–141.
Vissa B, Chacar AS. 2009. Leveraging ties: the contingent value of entrepreneurial teams' external advice networks on Indian software venture performance. *Strategic Management Journal* **30**(11): 1179–1191.
Wasserman N. 2012. *The Founder's Dilemmas*. Princeton University Press: Princeton, NJ.
Westphal JD. 1999. Collaboration in the boardroom: Behavioral and performance consequences of ceoboard social ties. *Academy of Management Journal* **42**(1): 7–24.
Yu S. 2020. How do accelerators impact the performance of high-technology ventures? *Management Science* **66**(2): 530–552.

TABLES

Table 1: Descriptive statistics of experimental intervention

Panel A: Pooled dataset

	Cohort	17	18	19	20	21
1	Start date	Feb-17	Jul-17	Mar-18	Jul-18	Feb-19
2	End date	Aug-17	Dec-17	Aug-18	Dec-18	Aug-19
3	Count of executives in cohort	4	4	7	7	6
4	Count of platoons (treatment/control)	6/6	6/6	7/7	7/7	6/6
5	Count of platoons per executive	3	3	2	2	2
6	Count of startups per platoon (t/c)	6.8/7	7.1/6.8	5.3/5.6	4.7/4.2	4.5/4.5
7	Count of startups (t/c)	41/42	43/41	37/39	33/31	27/27
8	Count of low education startups (t/c)	25/24	28/25	22/27	20/20	16/20
9	Count of high education startups (t/c)	16/18	15/16	15/12	13/11	11/7
10	Average number of meetings attended (t/c/total)	7/6.8/7	4.7/4.5/6	5.6/5.3/7	5.9/5.3/7	4.1/4.1/7
11	Average of correctly implemented meetings (t/c)	n.r.	4.6/4.5	4.9/5.2	5.5/5	3.9/3.3
12	Average meeting length in minutes (t/c)	n.r.	112/109	75/75	77/74	75/69
13	Average time allocated to reviewing past tasks in platoons	5.2	8.1	11.1	6.4	5.6
14	Average time allocated to enunciating future tasks in platoons	4.0	3.1	4.5	2.7	3.3
15	Committed tasks (t/c)	2.9/0.3	3.1/0.3	2.9/1.1	2/0.6	2.3/2.6
16	Completed tasks	1.9	1.8	1.9	1.8	1.2
17	Failed tasks	1.1	1.3	1.1	0.9	1.1

Note: Panel A describes the characteristics of the experiment for cohorts 17 through 21. Time allocated to discussing tasks and the count of tasks is at the founder level. Cohorts 17 and 18 only platoons, as boards were implemented starting at cohort 19, replacing half the platoon sessions. All cohorts had the same number of instances of structured accountability. According to accelerator policy, meeting attendance was mandatory in cohort 17, founders could skip one meeting in cohorts 18, and starting at cohort 19, founders could skip two meetings (one of each type). Time spent reviewing past tasks and enunciating future tasks was only registered for platoons, as founders had to share the meeting with peers. In contrast, board meetings were exclusive for each founder, and lasted one hour on average. Completed and failed tasks were only registered for treatment participants, because control participants were not held accountable for their tasks. We use "n.r." to denote "not recorded."

Panel B: Descriptive statistics separated by education level

	High Educated					High Educ. Avg. Treated	High Educ. Avg. Control	Diff	p-value
Cohort	**17**	**18**	**19**	**20**	**21**				
Average number of meetings attended (t/c)	7/7	4.8/4.5	5.3/5	6.2/6.2	3.6/4.1	5.5	5.6	0.1	0.80
Average of correctly implemented meetings (t/c)	n.r.	4.7/4.5	5/4.7	6/6	3.7/3.5	5.3	5.5	0.2	0.50
Average meeting length in minutes (t/c)	n.r.	109/109	72/77	82/76	72/75	86	88	-2	0.65
Average time allocated to reviewing past tasks in platoons	5.6	8.9	11.9	8.1	8.0	8.4	n.a.	-	-
Average time allocated to enunciating future tasks in platoons	4.8	2.4	4.1	2.9	2.1	3.4	n.a.	-	-
Committed tasks (t/c)	1.4/0	3/0.2	2.7/0.7	2.3/0.7	2.1/2.3	2.6	1.6	2	0.00
Completed tasks	1.8	1.8	1.6	1.9	1.0	1.7	n.a.	-	-
Failed tasks	0.8	1.2	1.1	0.8	1.1	1.0	n.a.	-	-

	Low Educated					Low Educ. Avg. Treated	Low Educ. Avg. Control	Diff	p-value
Cohort	**17**	**18**	**19**	**20**	**21**				
Average number of meetings attended (t/c)	7/6.7	4.7/4.6	5.7/5.4	5.7/4.8	4.5/4.2	5.6	5.2	0.4	0.04
Average of correctly implemented meetings (t/c)	n.r.	4.6/4.5	5.2/5.3	5.3/4.5	3.2/4.1	5.3	5	0.3	0.11
Average meeting length in minutes (t/c)	n.r.	114/109	76/73	78/70	76/70	89	83	6	0.11
Average time allocated to reviewing past tasks in platoons	5	8.1	10.1	7.9	5.9	7.4	n.a.	-	-
Average time allocated to enunciating future tasks in platoons	3.5	3.6	4	2.8	4.0	3.6	n.a.	-	-
Committed tasks (t/c)	1.7/0	3.2/0.1	3.1/1.2	1.9/0.5	2.3/2.6	2.9	0.9	2	0.00
Completed tasks	2.0	1.8	2.1	1.7	1.2	1.9	n.a.	-	-
Failed tasks	1.3	1.4	1.2	1.0	1.1	1.2	n.a.	-	-

Note: Panel B describes the characteristics of the experiment, separated by education level. We use (t/c) to denote (treatment/control). We used t-tests to compare high- and low-educated means. We use “n.r.” and “n.a.” to denote “not recorded” and “not applicable” respectively.

Table 2: Experiment implementation check

Panel A: Pooled dataset

	Treatment Group				
Question	**Treated**	**N**	**Control**	**N**	**p-value**
Attended at least one meeting (Y/N)	.98	203	.99	186	0.21
Number of meetings attended (1–7)	3.94	199	4.1	184	0.31
Received feedback by executive (Y/N)	.95	198	.97	185	0.19
Received feedback by others (mentors or other startups) (Y/N)	.98	198	.99	181	0.36
Executive asked to enunciate goals (1–5 Likert)* - platoons	4.7	96	4.3	91	0.00
Executive requested to follow-up on goals (1–5 Likert)* - platoons	4.6	96	4.2	91	0.01
Commitment with enunciated goals (1–5 Likert)	4.6	199	4.2	185	0.00
Competence in meeting enunciated goals (1–5 Likert)	4.5	199	4.3	185	0.03

* Not available for Gen 17

Note: Panel A presents the statistics of our analysis to ensure the correct implementation of the structured accountability treatment by testing the perceptions of startup team participants. We sent individual participants an email with an invitation to answer an online survey with questions to the eight items listed in the table. We received a maximum of 389 responses from our sample of 369 startups, since more than one team member per startup could participate in the platoon and board meetings. Treatment group participants perceived the treatment significantly more than the control group. We used Fisher's exact tests for binary (Y/N) responses and t-tests for Likert responses.

Panel B: Implementation check separated by education level

	Treated					Control				
Question	**High Educ**	**N**	**Low Educ**	**N**	**p-value**	**High Educ**	**N**	**Low Educ**	**N**	**p-value**
Attended at least one meeting (Y/N)	1.0	92	1.0	111	0.41	1	62	.99	124	0.48
Number of meetings attended (1–7)	3.8	91	4	108	0.49	4.4	62	3.9	122	0.02
Received feedback by executive (Y/N)	0.9	91	1.0	108	0.19	1.0	62	1.0	123	0.71
Received feedback by others (mentors or other startups) (Y/N)	1.0	91	1.0	107	0.06	1.0	59	1.0	122	0.48
Executive asked to enunciate goals (1–5 Likert)* - Board	4.7	79	4.6	94	0.38	4.3	48	4.5	106	0.14
Executive requested to follow-up on goals (1–5 Likert)* - Board	4.5	79	4.4	94	0.68	4.1	48	4.3	106	0.36
Executive asked to enunciate goals (1–5 Likert)* - Platoon	4.8	45	4.6	51	0.06	4.1	31	4.4	60	0.17
Executive requested to follow-up on goals (1–5 Likert)* - Platoon	4.7	45	4.5	51	0.46	4.1	31	4.2	60	0.79
Commitment with enunciated goals (1–5 Likert)	4.7	91	4.5	108	0.13	4.3	62	4.2	123	0.73
Competence in meeting enunciated goals (1–5 Likert)	4.7	91	4.4	108	0.00	4.4	62	4.2	123	0.18

* Not available for Gen 17

Note: Panel B replicates Panel A but separates results by education level. P-values compare responses between high- and low-education founders. Interestingly, in the treatment group, platoon executives were more likely to ask high education level founders to enunciate goals and low education founders were less likely to feel competent in meeting the enunciated goals. In the control group, as expected, we find no significant differences between high and low education founders when it comes to perceiving the treatment, as no treatment was given in this group. We used Fisher's exact tests for binary (Y/N) responses and t-tests for Likert responses.

Table 3: Descriptive statistics for baseline and performance variables stats

	Pooled Sample			Treatment						Control					
				High Education			Low Education			High Education			Low Education		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Baseline															
High Education	.37	.48	361	-	-	-	-	-	-	-	-	-	-	-	-
Female Leader	.22	.42	361	.29	.46	70	.20	.40	111	.19	.39	64	.22	.42	116
Age Team Leader	31.37	6.60	359	33.58	6.78	68	29.79	5.95	111	33.42	6.84	64	30.46	6.43	116
HQ in Chile	.39	.49	361	.4	.49	70	.41	.49	111	.34	.48	64	.41	.49	116
Chilean Entrepreneur	.31	.47	361	.29	.46	70	.34	.48	111	.27	.45	64	.34	.47	116
LatAm Entrepreneur	.37	.48	361	.3	.46	70	.35	.48	111	.36	.48	64	.43	.50	116
Capital Raised	.47	.50	361	.49	.50	70	.43	.50	111	.39	.49	64	.53	.50	116
Initial Sales*	9.52	9.42	347	2.61	11.43	65	1.92	4.96	108	5.79	20.91	62	22.92	164.41	112
Initial Stage	3.09	.80	347	2.97	.88	65	3.14	.68	108	3.13	.8	62	3.07	.87	112
Initial Employees	4.83	3.95	360	4.3	3.32	70	4.41	3.02	111	4.66	4.42	64	5.65	4.66	115
Score	3.56	.42	361	3.56	.42	70	3.5	.43	111	3.62	.44	64	3.58	.41	116
Internet Variables															
Employees	10.68	16.19	361	9.39	16.63	70	9.9	15.01	111	15.88	23.55	64	9.34	10.8	116
Capital Raised	.16	.36	361	.17	.38	70	.14	.34	111	.23	.43	64	.12	.33	116
Amount Raised*	86.88	551.71	360	94.45	457.46	70	37.73	226.59	111	154.02	509.63	63	92.89	794.13	116
Market Traction	537.74	3,838.67	361	476.39	1,512.78	70	358.89	1,728.75	111	1,582.72	8,653.4	64	169.35	311.04	116
Survival Online	.80	.40	361	.81	.39	70	.80	.40	111	.81	.39	64	.77	.42	116
Survey Variables															
Employees S1	5.78	7.59	260	5.11	6.97	54	5.84	8.35	76	5.92	8.58	48	6.09	6.67	82
Employees S2	9.86	65.9	190	4	2.8	35	20.25	121	56	6.54	9.39	37	5.77	7.48	62
Employees S3	4.66	7.34	146	2.88	2.95	25	3.73	5.31	45	7.45	11.42	29	4.77	7.16	47
Employees S4	3.55	7.25	100	2.37	3.65	19	3.06	7.14	32	6.16	11.73	19	3.17	5.02	30
Amount Raised S1*	27.26	100.09	247	20.48	67.15	51	8.55	25.14	70	42.56	167.49	47	39.12	104.77	79
Amount Raised S2*	51.21	199.83	180	9.16	22.65	32	37.92	207.27	53	148.43	345.55	36	26.65	69.1	59
Amount Raised S3*	42.29	309.31	140	12.42	27.76	22	7.31	22.52	43	185.63	681.35	28	2.87	10.97	47
Amount Raised S4*	97.67	909.73	98	0.78	3.23	17	3.13	17.68	32	497.83	2060	19	0	0	30
Market Traction S1*	7.68	23.24	249	8.66	30.6	52	4.24	9.53	71	10.68	25.29	48	8.33	25.07	78
Market Traction S2*	8.55	22.03	185	4.98	9.02	34	7.33	18.57	54	15.89	35.52	37	7.16	18.55	60
Market Traction S3*	6.85	18.41	142	4.76	11.23	24	8.98	27.49	43	11.83	19.11	29	2.82	5.1	46
Market Traction S4*	9.17	29.11	99	2.49	6.24	18	12.59	42.75	32	14.52	26.36	19	6.14	19.65	30

Note: Table 3 shows the mean, standard deviation, and sample size of baseline (at application) and post-treatment performance variables. Internet Variables correspond to performance measures found on the world wide web, while Survey Variables correspond to answers from four survey waves (S1 through S4). Appendix 4 describes the development of performance measures in more detail. Variables marked with an (*) are reported in thousands of USD. Variables are reported for the pooled sample and divided between high- and low-education founders according to their treatment condition.

Table 4: Baseline regressions relative to the control group

	(1)			(2)				
Estimate	**Treatment**	**N**	**R^2**	**Treatment Low Educ.**	**Treatment High Educ.**	**Difference**	**N**	**R^2**
High Education	0.037	357	0.10	-	-	-	-	
	(0.870)							
Female	0.024	357	0.12	-0.011	0.068	0.079	346	0.24
	(0.623)			(0.878)	(0.43)	(0.478)		
HQ in Chile	-0.006	357	0.20	-0.054	0.043	0.097	346	0.31
	(0.916)			(0.493)	(0.66)	(0.437)		
Chilean Entrepreneur	-0.020	357	0.20	-0.045	0.039	0.084	346	0.30
	(0.691)			(0.558)	(0.65)	(0.465)		
LatAm Entrepreneur	-0.074	357	0.19	-0.036	-0.073	-0.037	346	0.29
	(0.175)			(0.657)	(0.45)	(0.768)		
Age Team Leader	-0.721	355	0.18	-1.708	1.220	2.928	344	0.32
	(0.351)			(0.077)	(0.46)	(0.128)		
Capital Raised Before App.	-0.055	357	0.18	-0.115	0.054	0.169	346	0.29
	(0.332)			(0.141)	(0.61)	(0.201)		
Initial Sales	-0.797	343	0.21	-0.664	-1.472	-0.808	331	0.32
	(0.034)			(0.217)	(0.02)	(0.327)		
Initial Stage	-0.012	343	0.20	0.177	-0.398	-0.575	331	0.33
	(0.897)			(0.143)	(0.02)	(0.006)		
Initial Employees	-0.122	356	0.19	-0.149	-0.142	0.007	345	0.30
	(0.031)			(0.067)	(0.11)	(0.954)		
Score	-0.077	357	0.29	-0.034	-0.163	-0.129	346	0.39
	(0.067)			(0.548)	(0.05)	(0.204)		

Note: Table 4 shows the statistical differences for all baseline variables between treatment and control groups. Robust p-values are reported in parentheses. Model (1) includes fixed effects of Executive interacted with Cohort, and Industry. Model (2) include the same fixed effects interacted with the High Education dummy.

Table 5: Main results for performance measures

Panel A: Pooled results

				Control Variables	
	Treatment	**N**	**R^2**	**Basic**	**Age Team Leader**
Employees	0.088	343	0.28	Y	N
	(0.499)				
	0.122	341	0.30	Y	Y
	(0.348)				
Capital Raised	-0.006	343	0.17	Y	N
	(0.890)				
	0.001	341	0.17	Y	Y
	(0.983)				
Amount Raised	-0.072	343	0.15	Y	N
	(0.604)				
	-0.040	341	0.16	Y	Y
	(0.766)				
Market Traction	0.119	343	0.30	Y	N
	(0.278)				
	0.137	341	0.30	Y	Y
	(0.218)				
Survival	0.073	343	0.18	Y	N
	(0.139)				
	0.079	341	0.19	Y	Y
	(0.112)				

Note: Panel A reports the estimate of assignment to treatment as specified in section 6.1. Robust p-values reported in parentheses. Stars to signal statistical significance are omitted. Basic controls include Initial Employees, Initial Sales and Application Score.

Panel B: Results cut by education level

						Control Variables			
	Treatment Low Education	**Treatment High Education**	**Diff.**	**N**	**R^2**	**Basic**	**Basic*Above**	**Initial Stage**	**Age Team Leader**
Employees	0.394	-0.319	-0.713	331	0.38	Y	N	N	N
	(0.013)	(0.18)	(0.012)						
	0.412	-0.356	-0.767	331	0.39	N	Y	N	N
	(0.010)	(0.13)	(0.007)						
	0.398	-0.289	-0.687	331	0.41	N	Y	Y	N
	(0.015)	(0.19)	(0.012)						
	0.461	-0.308	-0.769	329	0.44	N	Y	Y	Y
	(0.007)	(0.17)	(0.007)						
Capital Raised	0.107	-0.164	-0.271	331	0.32	Y	N	N	N
	(0.062)	(0.08)	(0.015)						
	0.103	-0.151	-0.253	331	0.32	N	Y	N	N
	(0.074)	(0.11)	(0.023)						
	0.090	-0.139	-0.229	331	0.33	N	Y	Y	N
	(0.132)	(0.14)	(0.041)						
	0.095	-0.128	-0.223	329	0.38	N	Y	Y	Y
	(0.106)	(0.20)	(0.056)						
Amount Raised	0.359	-0.572	-0.930	331	0.31	Y	N	N	N
	(0.026)	(0.09)	(0.013)						
	0.322	-0.496	-0.818	331	0.33	N	Y	N	N
	(0.041)	(0.14)	(0.029)						
	0.297	-0.450	-0.747	331	0.34	N	Y	Y	N
	(0.074)	(0.18)	(0.047)						
	0.331	-0.439	-0.770	329	0.35	N	Y	Y	Y
	(0.044)	(0.20)	(0.045)						
Market Traction	0.315	-0.367	-0.682	331	0.41	Y	N	N	N
	(0.019)	(0.12)	(0.012)						
	0.333	-0.428	-0.761	331	0.42	N	Y	N	N
	(0.014)	(0.08)	(0.007)						
	0.321	-0.382	-0.703	331	0.44	N	Y	Y	N
	(0.017)	(0.11)	(0.010)						
	0.360	-0.394	-0.754	329	0.45	N	Y	Y	Y
	(0.009)	(0.12)	(0.010)						
Survival	0.131	0.018	-0.114	331	0.31	Y	N	N	N
	(0.044)	(0.81)	(0.243)						
	0.140	0.001	-0.139	331	0.30	N	Y	N	N
	(0.036)	(0.99)	(0.145)						
	0.136	0.018	-0.118	331	0.32	N	Y	Y	N
	(0.040)	(0.78)	(0.211)						
	0.154	-0.002	-0.156	329	0.33	N	Y	Y	Y
	(0.024)	(0.98)	(0.114)						

Note: Panel B reports the estimate of assignment to treatment by education level as specified in section 6.2. Robust p-values in parentheses. Basic controls include Initial Employees, Initial Sales and Application Score. Basic*Above are the basic controls interacted with the High Educated dummy variable. Initial Stage interact with the High Education dummy, however Age of Team Leader is added as control without interaction.

Table 6: Robustness 1 - Standardized survey-based variables

	Treatment Low Education	Treatment High Education	Difference	N	R^2
Employees S1	0.227	-0.323	-0.550	235	0.61
	(0.233)	(0.13)	(0.057)		
Employees S2	0.373	-0.575	-0.948	164	0.55
	(0.128)	(0.02)	(0.008)		
Employees S3	0.142	-0.791	-0.933	119	0.60
	(0.581)	(0.00)	(0.011)		
Employees S4	0.494	-0.099	-0.593	69	0.81
	(0.026)	(0.90)	(0.478)		
Amount Raised S4	0.085	-0.265	-0.350	222	0.37
	(0.670)	(0.49)	(0.421)		
Amount Raised S2	0.217	-1.000	-1.217	156	0.40
	(0.349)	(0.01)	(0.008)		
Amount Raised S3	0.384	-0.582	-0.966	115	0.51
	(0.101)	(0.39)	(0.177)		
Amount Raised S4	0.302	-2.779	-3.082	66	0.73
	(0.375)	(0.09)	(0.071)		
Market Traction S1	0.082	-0.357	-0.439	223	0.56
	(0.660)	(0.28)	(0.245)		
Market Traction S2	0.318	-0.366	-0.684	160	0.51
	(0.112)	(0.32)	(0.104)		
Market Traction S3	0.444	-0.797	-1.241	116	0.51
	(0.100)	(0.03)	(0.008)		
Market Traction S4	0.563	-1.027	-1.590	66	0.74
	(0.119)	(0.33)	(0.158)		

Note: Table 6 provides a robustness to Table 5, Panel B by using standardizing survey-based variables. To ease comparison, it only summarizes results of the most saturated model of Table 5B. S1 through S4 correspond to answers from the four survey waves. Robust p-values in parentheses. All models include Initial Employees, Initial Sales, Score, Initial Stage (interacted with High Education), and Age Team Leader as control variables.

Table 7: Robustness 2 - Standardized internet-based variables (main results)

						Control Variables			
	Treatment Low Education	**Treatment High Education**	**Difference**	**N**	**R^2**	**Basic**	**Basic*Above**	**Initial Stage**	**Age Team Leader**
Employees	0.343	-0.278	-0.622	331	0.38	Y	N	N	N
	(0.013)	(0.18)	(0.012)						
	0.359	-0.310	-0.669	331	0.39	N	Y	N	N
	(0.010)	(0.13)	(0.007)						
	0.347	-0.252	-0.669	331	0.42	N	Y	Y	N
	(0.015)	(0.19)	(0.007)						
	0.402	-0.269	-0.599	329	0.44	N	Y	Y	Y
	(0.007)	(0.17)	(0.012)						
Capital Raised	0.107	-0.164	-0.271	331	0.32	Y	N	N	N
	(0.062)	(0.08)	(0.015)						
	0.131	-0.151	-0.253	331	0.32	N	Y	N	N
	(0.074)	(0.11)	(0.023)						
	0.090	-0.139	-0.229	331	0.33	N	Y	Y	N
	(0.132)	(0.14)	(0.041)						
	0.095	-0.128	-0.223	329	0.34	N	Y	Y	Y
	(0.106)	(0.20)	(0.056)						
Amount Raised	0.308	-0.491	-0.800	331	0.31	Y	N	N	N
	(0.026)	(0.09)	(0.013)						
	0.277	-0.426	-0.703	331	0.33	N	Y	N	N
	(0.041)	(0.14)	(0.029)						
	0.255	-0.387	-0.642	331	0.34	N	Y	Y	N
	(0.074)	(0.18)	(0.047)						
	0.284	-0.377	-0.661	329	0.35	N	Y	Y	Y
	(0.044)	(0.20)	(0.045)						
Market Traction	0.315	-0.367	-0.671	331	0.41	Y	N	N	N
	(0.019)	(0.12)	(0.007)						
	0.333	-0.428	-0.682	331	0.42	N	Y	N	N
	(0.014)	(0.08)	(0.012)						
	0.321	-0.382	-0.761	331	0.44	N	Y	Y	N
	(0.017)	(0.11)	(0.007)						
	0.360	-0.394	-0.703	329	0.45	N	Y	Y	Y
	(0.009)	(0.12)	(0.010)						
Survival	0.131	0.018	-0.114	331	0.31	Y	N	N	N
	(0.044)	(0.81)	(0.243)						
	0.140	0.001	-0.139	331	0.30	N	Y	N	N
	(0.036)	(0.99)	(0.145)						
	0.136	0.018	-0.118	331	0.32	N	Y	Y	N
	(0.040)	(0.78)	(0.211)						
	0.154	-0.002	-0.156	329	0.33	N	Y	Y	Y
	(0.024)	(0.98)	(0.114)						

Note: Table 7 replicates Table 6, but using the internet-based variables. Robust p-values in parentheses. Basic controls include Initial Employees, Initial Sales and Application Score. Basic*Above are the basic controls interacted with the High Educated dummy variable. Initial Stage interact with the High Education dummy, however Age of Team Leader is added as control without interaction.

Table 8: Qualitative analysis from interviews about standardized meetings

		Weight: Treatment (+) Control (-)	**High-Educ:** Treatment (+) Control (-)	**Low-Educ:** Treatment (+) Control (-)	
Topic 1: Positive aspects of the board					
1	Positive perception about the role of the board included: access to new contacts, ideas, constructive criticism, domain knowledge, and encouragement.	+100%	+167%	+78%	
2	Structured accountability was appreciated as it helped founders to plan, focus and work within a given structure.	+52%	+139%	-8%	*
Topic 2: Critical aspects of the board					
3	Negative perception about the role of the board members, primarily due to unfulfilled expectations about what a board should do. Critiques included: boards were not relevant domain experts; board participation was discontinuous and sometimes absent; board members were not chosen by the entrepreneurs; instead of constructive feedback it was negative.	+20%	+67%	-16%	*
4	Feedback was perceived as useless, meetings were disorganized, and the tasks requested were not consistent with the development stages of the startups.	+32%	+29%	+35%	
5	Very little accountability was perceived, tasks requested seemed unreal or contradictory to founder's own goals.	0%	+167%	-93%	*
6	Emotions of frustration, anxiety, and perceptions of lack of honesty.	-94%	-91%	-100%	
Topic 3: Positive aspects of the platoon					
7	Founders appreciate platoons because of the emotional support from sharing experiences, the support from peers, generation of new contacts and leads, interaction with domestic entrepreneurs (for foreigners), friendships.	+22%	+93%	-22%	*
Topic 4: Critical aspects of the platoon					
8	Negative perception associated primarily to disorganization, unclear objectives, heterogeneous domain experience of peers.	+22%	-100%	+43%	*
9	Perceived as useless gatherings that did not add value, and conversations were forced.	-58%	-50%	-67%	
Topic 5: Comparison between platoons and boards					
10	Platoons are less formal, and thus there is a lower commitment in executing the enunciated tasks. Peers are too forgiving.	+116%	+200%	+86%	
Topic 6: General sentiments about Start-Up Chile					
11	Too many obligations.	+143%	-200%	+200%	*
12	Bureaucratic.	-120%	-67%	-200%	
13	Support to expand internationally.	+111%	+100%	+100%	

Note: Statement frequencies were counted from the 22 transcribed interviews. Weight is calculated as the percentage deviation of the frequencies over the mean frequency for the given statements. Positive (negative) percentages indicate that treatment (control) participants expressed listed concepts more than control (treatment) participants. The greater the absolute magnitude of the values, the greater the relative difference between groups. Nine (thirteen) interviewees were categorized as high-education (low-education) founders. Rows marked with a * highlight opposing perspectives between high- and low-education founders.

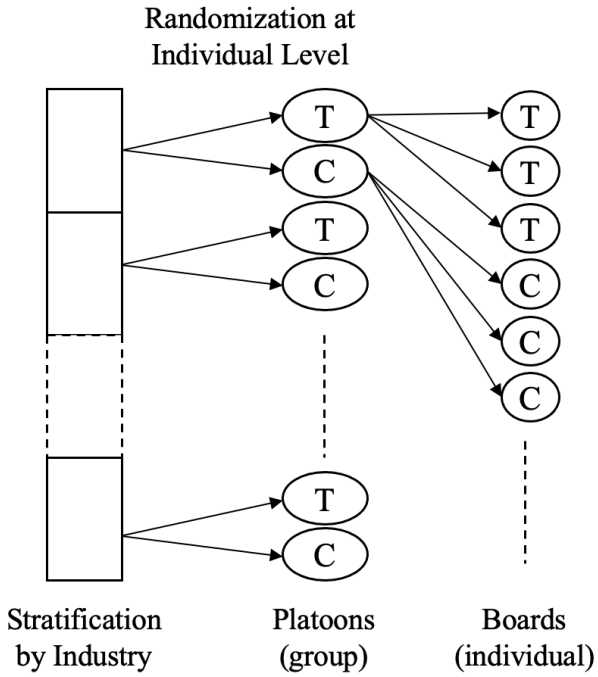
FIGURES

Figure 1: Parallel between Business Schools and Business Accelerators

Sources positive returns to schooling	Mechanism	Business School	Business Accelerators
Signalling (Spence, 1973, Arrow, 1973)	**Reputation** (Rao1994; Zott and Huy, 2007)	Certification from selection, graduation from business school, diploma.	Certification from selection, graduation from entrepreneurship school, exposure to community.
Productivity (Becker, 1964)	**Know-how** (Lerner and Malmendier, 2013)	Developing and growing a company through classes, professors, guest speakers, career office, advisors, fellow classmates.	Developing and growing a startup through workshops, staff, guest speakers, industry experts, mentors, fellow participants.
	Social Networks (Granovetter, 1973; Ketchen, Ireland and Snow, 2007)	Preferential access to peer and professor networks.	Preferential access to peer and staff networks.
	Self-efficacy (Bandura, 1982; Forbes, 2005)	Self-confidence from selection and graduation (in the form of business self-efficacy)	Self-confidence from selection and graduation (in the form of entrepreneurial self-efficacy)
	Structured Accountability (Locke and Latham, 2002; Cialdini and Goldstein, 2004)	Setting learning goals, class work, homework, exams.	Setting strategic tasks, monthly follow-up meetings, demo-day

Note: Figure 1 provides a conceptual comparison between business schools and business accelerators. Five mechanisms are described. This paper dives into the mechanism of structured accountability.

Figure 2: Graphical representation of assignment to experimental groups



Note: Startups were first stratified by industry and then randomized individually to treatment or control groups. Platoon executives were matched to the industry strata, such that each executive had two or more treatment and control groups. The treatment assignment was then extended to Boards.

APPENDIX

Appendix 1: Startup and founder characteristics at application by cohort

Panel A: Applicants and participants

Cohort	Applicants	Rejections	Selections	Participants	Treatment	Control
17	619	519	100	83	41	42
18	719	619	100	84	43	41
19	582	491	91	76	37	39
20	467	394	73	64	33	31
21	156	96	60	54	27	27
Total	2,543	2,119	424	361	181	180

Panel B: Capital raised at application

	Start-Up Chile							ED
	17	18	19	20	21	Total	%	%
No (Bootstrapped)	55	58	30	28	22	193	53.4	79.3
Yes	28	26	46	36	32	168	46.5	20.7
Total	83	84	76	64	54	361		

Panel C: Number of full-time workers at application

	Start-Up Chile							ED
	17	18	19	20	21	Total	%	%
-	0	2	7	2	2	13		
<5	61	60	45	41	21	228	65.5	68.5
5-10	20	15	19	20	24	98	28.2	16.9
10+	2	7	5	1	7	22	6.3	14.6
Total	83	84	76	64	54	348		

Panel D: Startup age at application

	Start-Up Chile							ED
	17	18	19	20	21	Total	%	%
-	83	0	0	0	0	83		
Less than 6 months	-	18	17	10	3	48	17.3	21.9
6-12 months	-	41	33	30	26	130	46.8	29.4
12-24 months	-	22	18	23	22	85	30.6	17.1
More than 2 years	-	3	8	1	3	15	5.4	31.7
Total	-	84	76	64	54	278		

Panel F: Startup development stage

	Start-Up Chile							ED
	17	18	19	20	21	Total	%	%
-	1	2	10	0	1	14		
Concept	2	4	2	3	4	15	4.3	22.1
Prototype in Development	11	12	10	11	9	53	15.3	4.1
Functional Product with Users	37	41	27	34	28	167	48.1	22.5
Scaling Sales	32	25	27	16	12	112	32.3	51.3
Total	83	82	66	64	53	361		

Panel G: Location of lead founder

	Start-Up Chile							ED
	17	18	19	20	21	Total	%	%
-								
Africa	0	2	1	1	2	6	1.7	19.1
Asia	7	10	8	8	4	37	10.2	19.1
Europe	15	6	6	9	5	41	11.4	6.6
N. America	11	8	12	5	6	42	11.6	34.8
Oceania	1	1	0	1	1	4	1.1	0.4
S. America (exc. Chile)	31	22	26	19	19	117	32.4	19.4
Chile	18	35	23	21	17	114	31.6	0.6
Total	83	84	76	64	54	361		

Panel H: Age of the lead founder

	Start-Up Chile							ED
	17	18	19	20	21	Total	%	%
-								
Younger than 25	13	8	7	4	5	37	10.2	10.6
Between 25 and 30	30	36	34	25	23	148	41.0	21.7
Between 30 and 35	23	21	20	17	14	95	26.3	21.6
Between 35 and 40	10	9	11	10	5	45	12.5	15.2
Older than 40	7	10	4	8	7	36	10.0	30.8
Total	83	84	76	64	54	361		

Panel I: Gender of the lead founder

	Start-Up Chile							ED
	17	18	19	20	21	Total	%	%
Female	19	18	21	12	10	80	22.2	28.6
Male	64	66	55	52	44	281	77.8	71.4
Total	83	84	76	64	54	361		

Note: Panels A through I describe the composition of the sample, totalized and also separated into cohorts 17 through 21. The distribution of the study population (%) is compared to the average applicants to ecosystem accelerators worldwide under the heading "ED," based on information from the Emory Entrepreneurship Database. Panels A through F are startup level variables, while Panels G through I are founder-level variables. For those applicant startups with multiple founders, only the characteristics of the founder leader (self-reported in application) are described The total number of startups in the experiment was 361. Tables with fewer observations are the result of non-response questions at the application stage. Percentages are calculated over the number of non-missing responses.

Appendix 2: Industry breakdown of applicant startups

Industry (aggregated)	Freq	Percent
Other	71	19.67%
Information and communication tech.	61	16.90%
Education & Culture	38	10.53%
Biotechnology & Health	35	9.70%
Financial Services	24	6.65%
IT & Software	21	5.82%
Agriculture & Natural Resource	19	5.26%
Retail	13	3.60%
Tourism	10	2.77%
Energy, Sustainability & Environment	9	2.49%
SaaS	8	2.22%
Environment & CleanTech	7	1.94%
Technical assistance services	5	1.39%
Fintech	4	1.11%
Infrastructure/ facilities development	4	1.11%
Manufacturing	4	1.11%
Social Impact	4	1.11%
Supply chain services	4	1.11%
Food, Beverage & Restaurants	3	0.83%
Logistics & Transportation	3	0.83%
Consulting	2	0.55%
Entertainment & Gaming	2	0.55%
Hardware	2	0.55%
Housing and Development	2	0.55%
Security	2	0.55%
CRM & ERP	1	0.28%
General Services	1	0.28%
Import/Export	1	0.28%
Social Media/Social Network	1	0.28%
Total	361	100%

Note: This table shows the distribution of applicant startups classified by industry.

Appendix 3: Baseline without 19th cohort

	(1)			(2)				
Estimate	**Treatment**	**N**	**R^2**	**Treatment Low Educ.**	**Treatment High Educ.**	**Difference**	**N**	**R^2**
High Education	0.011	281	0.11	-	-	-	-	-
	(0.870)							
Female	-0.073	281	0.15	-0.161	0.053	0.214	268	0.25
	(0.193)			(0.035)	(0.63)	(0.109)		
HQ in Chile	-0.020	281	0.23	-0.053	0.026	0.079	268	0.35
	(0.753)			(0.563)	(0.82)	(0.594)		
Chilean Entrepreneur	-0.034	281	0.21	-0.076	0.030	0.106	268	0.39
	(0.577)			(0.401)	(0.78)	(0.452)		
LatAm Entrepreneur	-0.045	281	0.19	0.027	-0.086	-0.113	268	0.30
	(0.474)			(0.773)	(0.47)	(0.452)		
Age Team Leader	-1.551	279	0.21	-2.135	-0.808	1.327	266	0.32
	(0.099)			(0.083)	(0.68)	(0.565)		
Capital Raised Before App.	-0.084	281	0.18	-0.172	0.083	0.255	268	0.30
	(0.195)			(0.053)	(0.49)	(0.091)		
Initial Sales	-0.742	277	0.20	-0.737	-1.175	-0.437	263	0.32
	(0.083)			(0.230)	(0.11)	(0.648)		
Initial Stage	0.024	277	0.23	0.173	-0.371	-0.544	263	0.33
	(0.818)			(0.232)	(0.07)	(0.029)		
Initial Employees	-0.094	281	0.22	-0.086	-0.206	-0.119	268	0.38
	(0.133)			(0.346)	(0.03)	(0.365)		
Score	-0.068	281	0.31	-0.007	-0.186	-0.179	268	0.43
	(0.160)			(0.915)	(0.07)	(0.134)		

Note: This table shows the statistical differences for all baseline variables between treatment and control groups while excluding cohort 19. Robust p-values are reported in parentheses. Model (1) includes fixed effects of Executive interacted with Cohort, and Industry. Model (2) include the same fixed effects interacted with the High Education dummy.

Appendix 4: Development of outcome variables

Given the fledgling nature of startups, the standard metrics used to establish firm performance for more mature businesses (e.g., profits or stock price) are not generally available, nor are they particularly useful in new venture settings (cf. Puri and Zarutskie, 2012). For example, Facebook purchased Instagram for roughly $1 billion when it was only one and a half years old and had neither revenues nor profits. However, it had over 100 million active users. Therefore, in keeping with prior literature (Eisenhardt and Schoonhoven, 1990; Maurer and Ebers, 2006), we construct performance measures that proxy for venture fundraising, venture scale, and venture survival. We use two methods: internet searches and surveys.

Our first internet-based search (conducted during the months of October and November of 2020) focused on CB Insights and LinkedIn. Our second web search (conducted November 2020) focused on the Facebook and Twitter platforms. Our first survey (conducted during the fourth quarter of 2019 and first quarter of 2020) was focused on participants. Logarithmic transformations of the survey responses are used to reduce the impact of outliers. Following, we describe details of each of the data collection methods.

Internet-based Measures:

Variable	Description	Construction logic
Employees	Number of employees	LinkedIn reports the number of employees in ranges (e.g., 1-10 employees), which we transform into point estimates using the median employee size in the range (i.e., we assigned an employment level of 5 when the reported range was 1-10 employees). We confirmed that the transformation rule is immaterial for the results.

Capital Raised	*Capital Raised* equals 1 if the startup has a post-application fundraising record, and 0 otherwise.	If a startup has relevant fundraising activity, that activity is most likely to appear on CB Insights. By construction, we also code this variable with zero for those that do not have a profile on CB Insights. We use detailed information about the fundraising date in the platform, together with the startups' application date, to classify fundraising rounds as post-application.
Amount Raised*	*Amount Raised* is the natural log of the value of capital raised.	This variable equals zero if the startup has no post-application fundraising record on CB Insights, if such a record exists but does not specify an amount of capital raised, or if the startup has no profile on CB Insights.
Market Traction	*Market Traction* is the natural log of the number of LinkedIn followers, standardized.	*
Survival Online	*Survival Online* equal 1 if the startup has online presence.	This variable equals zero if the startup does not have a profile on LinkedIn or Crunchbase.

Survey based Variables:

We sent an email to all the 361 participants between the months of October 2019 and January 2020 asking for the survival (and last time the startup was active if the answer was negative), the average monthly values of revenue, the amount of capital raised, and the number of employees of each semester since graduation. We received 265 responses (73.4% response rate).

Cohort	**17**	**18**	**19**	**20**	**21**
Participation Dates	Feb-2017 to Aug-2017	Jul-2017 to Dic-2017	Mar-2018 to Aug-2018	Jul-2018 to Dic-2018	Feb-2018 to Aug-2018
Survey date	Oct-19	Oct-19	Nov-19	Nov-19	Jan-20
Participants	83	84	76	64	54
Responses	52	63	60	54	36
Response Rate	62.7%	75.0%	78.9%	84.4%	66.7%
Times Surveyed	4	3	2	1	1
Semesters Surveyed	2017-2 to 2019-1	2018-1 to 2019-1	2018-2 to 2019-1	2019-1	2019-2
Semesters since graduation	4	3	2	1	1

Consistency Across Instruments:

To assess the consistency our performance proxies, we compare our internet-based measures against the survey-based measures. Despite being different and independent instruments, all performance variables are strongly and significantly correlated, except for employees which shows a weak correlation, as seen in the following table:

	Internet-based Variables	**Survey Variables**	**Correlation**	**Observations**
Employment	LinkedIn Employees	Employees	0.06	270
			(0.28)	
Capital Raised	Capital Raised dummy	Capital Raised	0.32	286
			(0.00)	
Amount Raised	Funding After Start-Up Chile	Capital Raised	0.55	286
			(0.00)	
Market Traction	LinkedIn Followers	Revenue	0.21	262
			(0.00)	
Survival	Survival Online	Survival	0.26	295
			(0.00)	

Note: The table presents correlations across Internet-based and survey-based venture performance metrics. Robust p-values are reported in parenthesis.

Appendix 5: Main results without industry fixed effects

	Treatment					Controls			
Estimate	**Low Education**	**High Education**	**Difference**	**N**	**R^2**	**Raw**	**Above**	**Initial Stage**	**Age Team Leader**
Employees	0.249	-0.265	-0.514	347	0.303	Y	N	N	N
	(0.086)	(0.23)	(0.048)						
	0.273	-0.293	-0.566	347	0.308	N	Y	N	N
	(0.064)	(0.18)	(0.032)						
	0.254	-0.254	-0.508	347	0.336	N	Y	Y	N
	(0.091)	(0.20)	(0.041)						
	0.285	-0.270	-0.556	345	0.348	N	Y	Y	Y
	(0.067)	(0.16)	(0.026)						
Capital Raised	0.066	-0.113	-0.179	347	0.183	Y	N	N	N
	(0.180)	(0.19)	(0.071)						
	0.063	-0.103	-0.166	347	0.185	N	Y	N	N
	(0.206)	(0.24)	(0.098)						
	0.053	-0.098	-0.151	347	0.193	N	Y	Y	N
	(0.305)	(0.26)	(0.135)						
	0.054	-0.085	-0.139	345	0.197	N	Y	Y	Y
	(0.287)	(0.33)	(0.173)						
Amount Raised	0.193	-0.441	-0.634	347	0.198	Y	N	N	N
	(0.165)	(0.15)	(0.064)						
	0.163	-0.404	-0.567	347	0.205	N	Y	N	N
	(0.230)	(0.19)	(0.095)						
	0.138	-0.379	-0.517	347	0.218	N	Y	Y	N
	(0.326)	(0.22)	(0.127)						
	0.153	-0.348	-0.502	345	0.223	N	Y	Y	Y
	(0.265)	(0.26)	(0.137)						
LinkedIn Followers	0.193	-0.233	-0.426	347	0.337	Y	N	N	N
	(0.111)	(0.28)	(0.083)						
	0.205*	-0.263	-0.469	347	0.341	N	Y	N	N
	(0.090)	(0.23)	(0.061)						
	0.184	-0.231	-0.415	347	0.368	N	Y	Y	N
	(0.127)	(0.25)	(0.079)						
	0.201	-0.251	-0.452	345	0.371	N	Y	Y	Y
	(0.104)	(0.24)	(0.070)						
Survival Online	0.087	0.017	-0.070	347	0.199	Y	N	N	N
	(0.151)	(0.81)	(0.443)						
	0.097	0.002	-0.095	347	0.205	N	Y	N	N
	(0.118)	(0.97)	(0.306)						
	0.091	0.013	-0.078	347	0.222	N	Y	Y	N
	(0.145)	(0.84)	(0.374)						
	0.100	-0.002	-0.102	345	0.229	N	Y	Y	Y
	(0.118)	(0.98)	(0.257)						

Note: This table replicates Table 5, Panel B, by reporting the estimate of assignment to treatment by education level as specified in section 6.2, but excluding industry fixed effects as described in section 7.1. Robust p-values in parentheses. Basic controls include Initial Employees, Initial Sales and Application Score. Basic*Above are the basic controls interacted with the High Educated dummy variable. Initial Stage interact with the High Education dummy, however Age of Team Leader is added as control without interaction.

Appendix 6: Main results without 19th cohort

	Treatment					Controls			
Estimate	**Low Education**	**High Education**	**Difference**	**N**	**R^2**	**Raw**	**Above**	**Initial Stage**	**Age Team Leader**
Employees	0.468	-0.404	-0.872	263	0.40	Y	N	N	N
	(0.013)	(0.12)	(0.006)						
	0.463	-0.369	-0.832	263	0.40	N	Y	N	N
	(0.015)	(0.15)	(0.009)						
	0.432	-0.325	-0.756	263	0.44	N	Y	Y	N
	(0.026)	(0.17)	(0.014)						
	0.498	-0.261	-0.759	261	0.46	N	Y	Y	Y
	(0.014)	(0.29)	(0.018)						
LinkedIn Followers	0.349	-0.257	-0.606	263	0.45	Y	N	N	N
	(0.017)	(0.35)	(0.051)						
	0.345	-0.289	-0.634	263	0.46	N	Y	N	N
	(0.020)	(0.31)	(0.050)						
	0.301	-0.255	-0.555	263	0.49	N	Y	Y	N
	(0.037)	(0.36)	(0.077)						
	0.346	-0.216	-0.562	261	0.51	N	Y	Y	Y
	(0.022)	(0.46)	(0.091)						
Capital Raised	0.098	-0.191	-0.289	263	0.34	Y	N	N	N
	(0.164)	(0.07)	(0.023)						
	0.095	-0.172	-0.267	263	0.35	N	Y	N	N
	(0.178)	(0.11)	(0.038)						
	0.084	-0.164	-0.248	263	0.36	N	Y	Y	N
	(0.251)	(0.13)	(0.057)						
	0.086	-0.138	-0.224	261	0.37	N	Y	Y	Y
	(0.231)	(0.23)	(0.103)						
Amount Raised	0.296	-0.744	-1.040	263	0.35	Y	N	N	N
	(0.117)	(0.04)	(0.013)						
	0.267	-0.626	-0.893	263	0.37	N	Y	N	N
	(0.142)	(0.08)	(0.027)						
	0.256	-0.595	-0.851	263	0.38	N	Y	Y	N
	(0.180)	(0.10)	(0.039)						
	0.273	-0.531	-0.804	261	0.39	N	Y	Y	Y
	(0.145)	(0.16)	(0.061)						
Survival	0.160	-0.004	-0.164	263	0.27	Y	N	N	N
	(0.043)	(0.96)	(0.157)						
	0.160	-0.011	-0.171	263	0.28	N	Y	N	N
	(0.046)	(0.90)	(0.137)						
	0.145	0.002	-0.832	263	0.31	N	Y	Y	N
	(0.069)	(0.98)	(0.009)						
	0.166	-0.002	-0.168	261	0.32	N	Y	Y	Y
	(0.045)	(0.98)	(0.151)						

Note: This table replicates Table 5, Panel B, by reporting the estimate of assignment to treatment by education level as specified in section 6.2, but excluding cohort 19 as described in section 7.1. Robust p-values in parentheses. Basic controls include Initial Employees, Initial Sales and Application Score. Basic*Above are the basic controls interacted with the High Educated dummy variable. Initial Stage interact with the High Education dummy, however Age of Team Leader is added as control without interaction.

Appendix 7: Main Results without 18th cohort

	Treatment					Controls			
Estimate	**Low Education**	**High Education**	**Difference**	**N**	**R^2**	**Raw**	**Above**	**Initial Stage**	**Age Team Leader**
Employees	0.287	-0.346	-0.633*	248	0.384	Y	N	N	N
	(0.128)	0.25	(0.071)						
	0.316*	-0.384	-0.700**	248	0.391	N	Y	N	N
	(0.099)	0.20	(0.048)						
	0.285	-0.316	-0.601*	248	0.417	N	Y	Y	N
	(0.147)	0.28	(0.088)						
	0.359*	-0.347	-0.706*	247	0.432	N	Y	Y	Y
	(0.083)	0.25	(0.058)						
Capital Raised	0.123*	-0.181	-0.305**	248	0.360	Y	N	N	N
	(0.078)	0.12	(0.026)						
	0.133*	-0.189	-0.321**	248	0.371	N	Y	N	N
	(0.060)	0.12	(0.022)						
	0.096	-0.174	-0.271*	248	0.395	N	Y	Y	N
	(0.193)	0.14	(0.054)						
	0.105	-0.167	-0.272*	247	0.399	N	Y	Y	Y
	(0.154)	0.17	(0.060)						
Amount Raised	0.357*	-0.427	-0.784*	248	0.335	Y	N	N	N
	(0.068)	0.31	(0.090)						
	0.372*	-0.466	-0.838*	248	0.344	N	Y	N	N
	(0.060)	0.28	(0.076)						
	0.284	-0.419	-0.703	248	0.362	N	Y	Y	N
	(0.174)	0.33	(0.140)						
	0.331	-0.427	-0.758	247	0.368	N	Y	Y	Y
	(0.113)	0.33	(0.121)						
LinkedIn Followers	0.337**	-0.265	-0.603*	248	0.420	Y	N	N	N
	(0.042)	0.35	(0.068)						
	0.376**	-0.294	-0.670*	248	0.435	N	Y	N	N
	(0.025)	0.35	(0.061)						
	0.356**	-0.253	-0.608*	248	0.448	N	Y	Y	N
	(0.035)	0.41	(0.085)						
	0.394**	-0.274	-0.669*	247	0.452	N	Y	Y	Y
	(0.023)	0.41	(0.083)						
Survival	0.067	0.031	-0.036	248	0.337	Y	N	N	N
	(0.377)	0.72	(0.747)						
	0.079	0.019	-0.060	248	0.348	N	Y	N	N
	(0.300)	0.82	(0.600)						
	0.071	0.037	-0.035	248	0.361	N	Y	Y	N
	(0.349)	0.66	(0.757)						
	0.090	0.021	-0.069	247	0.369	N	Y	Y	Y
	(0.255)	0.80	(0.556)						

Note: This table replicates Table 5, Panel B, by reporting the estimate of assignment to treatment by education level as specified in section 6.2, but excluding cohort 18 as described in section 7.1. Robust p-values in parentheses. Basic controls include Initial Employees, Initial Sales and Application Score. Basic*Above are the basic controls interacted with the High Educated dummy variable. Initial Stage interact with the High Education dummy, however Age of Team Leader is added as control without interaction.