Understanding Probabilistic Reasoning in Entrepreneurship

Andrea Coali, Alfonso Gambardella, Elena Novelli

Preliminary Draft
Please do not circulate without authors' permission
November 2021

Abstract

Decision-making processes in the context of innovation and entrepreneurship are characterized by high uncertainty and prone to decision-making biases. In this paper we explore the implications of adopting what we call a *scientific approach to decision making*, based on probabilistic reasoning. We develop a structural model to disentangle and identify two separate but complementary effects of this approach. The estimation of our structural model, based on data from two randomized control trials (RCTs) involving early stage start-ups, shows that scientific entrepreneurs tend to be more conservative in assessing the value over their ideas, an effect that we call *debiasing effect*. It also shows that, conditional on their decision to remain operational, scientific entrepreneurs tend to perform better, an effect that we call *learning effect*. We discuss the implications for future research and entrepreneurial practice.

1 Introduction

An astonishing 90% of newly-born start-ups fail within ten years, with around 21% of them failing already in their first year (NationalBusinessCapital, 2020). Part of the reasons behind this pattern relates to the fact that entrepreneurs, and innovators developing new projects more in general, face a decision-making process that is characterized by high uncertainty along multiple dimensions (McGrath, 1997, Folta, 1998). In the presence of uncertainty, the assessment on the value of novel ideas, and therefore, entrepreneurial decision-making becomes difficult.

One way to deal with uncertainty would be to use a probabilistic decision-making process, making decisions on uncertain outcomes based on probabilistic information. However, research shows that decision makers often deviate from "rational" or probabilistic reasoning. Prior research has shown that entrepreneurs often do not follow a systematic decision-making process (Bloom and Van Reenen, 2007) and even ignore important information (Bennett and Chatterji, 2019, Tversky and Kahneman, 1974). Alternatively, they rely on heuristic principles to reduce the complex tasks of probability assessment and value prediction to a simpler task (Tversky and Kahneman, 1974). Whereas this is certainly a useful process and can lead to good results (Bingham and Eisenhardt, 2011), research shows that it can also lead to a plethora of severe and systematic biases (Tversky and Kahneman, 1974).

Differently from prior research that has largely explored deviations from a probabilistic approach to decision making, in this paper we aim to understand more about what happens when entrepreneurs are induced to employ probabilistic decision-making processes more systematically. Specifically, we address the following research question: *What are the implications of probabilistic reasoning on entrepreneurial decision making?*

This question has been so far under investigated, but this is perhaps unsurprising given the research-design challenges that addressing this question involves. First, answering this question requires an exogenous shock that induced entrepreneurs to reason and make decisions in probabilistic terms. Second, it requires observing the decision-making process of entrepreneurs in detail as well as the outcomes originating from that process, such as the specific decisions made and their performance. Third, it requires comparing entrepreneurs using a probabilistic approach to a proper counterfactual.

To respond to these challenges, we employ a randomized control trial (RCT) design, where we teach a sample of entrepreneurs to reason in probabilistic terms, developing a theory of their business idea and the problems it would likely solve, developing hypotheses flowing logically from it, designing tests that can provide them with signals regarding the probabilities of those hypotheses being supported with data, and evaluating those results in a disciplined way against their prior theoretical expectations. Following related work (Ashraf, Banerjee, and Nourani, 2021; Camuffo et al., 2020), we call this "a scientific approach to innovation management" as this process resembles the one followed by scientists in developing new knowledge. We maintain the other half of the sample in a control condition, where they are delivered equivalent management content but without a scientific approach. We then monitor these entrepreneurs for a variety of months, collecting detailed data on their decision-making process, choices, and performance.

We dig one step deeper compared to prior research (e.g. Camuffo et al., 2020) and develop a structural model that allows us to disentangle and identify precisely the two different effects that the exposure to a scientific approach has on entrepreneurial decision making. The estimation of our structural model, based on data from two randomized control trials (RCTs) conducted in Milan (2017) and Turin (2018), involving 377 early stage start-ups, shows that entrepreneurs following a scientific approach to decision making perform on average better, an effect that we call the *learning effect*; but they also make an earlier and faster downward adjustment of their business' expected values, ultimately showing a higher rate of project termination, an effect that we call the *debiasing effect*. This in line with the intuition that entrepreneurs tend to sometimes pursue "falsely positive ideas" and that probabilistic reasoning can help them to reach a more conservative evaluation of their ideas. But that it can also help them understand the problem faced better and identify a better solution, achieving superior performance.

The co-existence of these two effects leads to a natural follow up question. Given that a scientific approach to decision making leads both to a more conservative assessment of ideas but also to superior learning and performance, are scientific entrepreneurs excessively cautious when terminating their projects, effectively discarding ideas that could, instead, be successful? In other words, is it possible that while this method leads entrepreneurs to reduce the number of ideas that others would have falsely seen as positive, it also leads entrepreneurs to discard too many (falsely negative) ideas?

Answering this question is no easy task, as it would require the determination of the value of the terminated ideas, were they not terminated, which is clearly not possible. However, in the attempt to nurture this important debate, we provide some suggestive evidence. Leveraging on the results, we identify different assumptions and use them to calculate two different scenarios at different end of the spectrum of possibilities. In what we call the *lower-bound* condition, we assume that the value model for firms that stayed in the market is exactly the same as for those firms that terminate, not considering the positive learning effect of "scientific" entrepreneurs. Results show that, under this condition, the reduction in false positives is compensated by the increase in false negatives. In what we call the *upper-bound* condition we consider, instead, the positive learning effect of the "scientific" approach. Results show that the selection has been a "positive" one, meaning that the reduction in false positives more than compensates the increase in false negatives.

The remainder of the paper is structured as follows. Section II elaborates on the scientific approach and its implications. Section III details the structural framework we develop. Section IV describes our methodology and data. Section V reports the estimation results. In Section VI we offer concluding reflections.

2 The Scientific Approach and Probabilistic Reasoning

Consider an entrepreneur with the goal of developing an innovative product or service, or willing to launch a new business. Typically, she starts with an intuition coming from observation of real-world phenomena, spotting a problem that would need an innovative solution to be solved. Before deciding to embark in a new project, our decision-maker will evaluate whether her idea is worth the development efforts and this assessment will be made at regular intervals throughout the life of the project. At every round of assessment, her decision can be represented as a choice between three mutually exclusive alternatives

(Kirtley and O'Mahony, 2020, Lieberman, Lee, and Folta, 2017, Eisenhardt and Bingham, 2017, Gans, Stern, and Wu, 2019): 1) *terminate* the project, if the entrepreneur believes it won't generate sufficient value; 2) change substantial elements of the idea or project to improve its value (what we refer to as *pivoting*); 3) *continue* the development of the project along its current trajectory.

Along the way her assessment will be based on considerations regarding the multiple potential scenarios she could face in the environment in which she operates, over which there is uncertainty. This uncertainty could originate, for instance, from the fact that she is not yet familiar with customers' preferences in the environment she targets; or from the fact that these preferences might be subject to change. She will also consider the actions that she can take to deal with the multiple scenarios she might be facing. At the very early stage of her process, actions could concern the development of the idea. At later stages, actions could be linked to the idea commercialization and could include, for example, the development of different versions of the same product, service, or business model, or the implementation of alternative marketing strategies.

Of course, every action she envisions might have a different value under different scenarios. Suppose for instance that our entrepreneur's idea is about developing an innovative service for car-sharing, but there is uncertainty regarding the extent to which cars are going to be relevant in the medium term in the context in which she is operating. If the context in which she operates is going through a massive drop in the use of cars and an increase in the use of bikes, the action of pursuing such car-sharing project could have a negative value. Instead, if renting cars is a valuable option in the context in which she operates, the action of pursuing such car-sharing project idea could have a high value. Depending on what her assessment on the scenario more likely to manifest itself and what value she envisions her actions to have, she could decide to terminate the project, or to pivot to a new version of the project, or to simply continue the development of the project along its natural trajectory.

Prior literature in management, entrepreneurship and innovation has shown that very often entrepreneurs make these decisions simply following their gut feelings as opposed to trying and predicting the likelihood of scenarios and the value of actions (Bennett and Chatterji, 2019). Other streams of research have instead documented the use of structured approaches that support entrepreneurial decision making, such as the use of structured practices (Bloom and Van Reenen, 2007; Yang et al., 2020), trial and error (von Hippel and Tyre, 1995), effectuation (Sarasvathy, 2001), experimentation and lean methods (Ries, 2011; Thomke, 1998), heuristics (Bingham and Eisenhardt, 2011; Bingham and Haleblian, 2012). Whereas these approaches can be beneficial and lead to superior performance (e.g. Bingham and Eisenhardt, 2011), existing research is also replete with examples that show that the use of these practices can also lead to important biases (Tversky and Kahneman, 1974; Felin et al., 2020)

The key question we, thus, address in this paper is: to what extent can entrepreneurs, in the face of uncertainty, use probabilistic reasoning to discover relevant scenarios and assess the value of their entrepreneurial ideas under those scenarios? And what would the implications for entrepreneurial decision making of using such an approach be?

To this end, we explore a decision making approach that is based on ”probabilistic” reasoning. Following related work (Camuffo et al., 2020, Ashraf, Banerjee, and Nourani, 2021), we refer to this approach as to *a scientific approach to decision making* due to its resemblance to the process followed by scientists when they approach a problem. The key tenet of this methods is that “scientific” entrepreneurs follow a five-step process in making decisions. They start from thinking about the problem in a broad way, effectively developing a theory of the problem and identifying the key elements on which they should focus when developing their projects, for instance the relevant scenarios they should take into account and the relevant actions they could consider in those scenarios (step 1). Scientific entrepreneurs then formulate testable and falsifiable hypotheses based on their theory (step 2) and they test them via carefully designed tests (step 3). The outcomes of these tests can be used as “signals” by the entrepreneur to assess the value of the idea. Signals will then be rigorously evaluated against their theory and prior beliefs. Such evaluation (step 4) will ultimately lead to a decision on the future of the idea (step 5).

For instance, if our entrepreneur approached the problem in a scientific way, she would start by developing a theory about the problem that her such car-sharing service addresses and the way in which it addresses it, and how the value of those actions would change under different relevant scenarios. She would then develop some core hypotheses regarding the scenarios she is facing and the value of actions in those scenarios, such as that car transportation in large cities is highly valued by certain categories of individuals, that owning a car in a large city is not practical due to the high fixed costs and the limited use per person, and that individuals consider sharing cars a viable option. She will test such hypotheses by collecting relevant information from a representative group of target customers. She will then evaluate the results obtained from the test against her theory, to ultimately reach a decision about whether to continue with the development of her idea, terminate the project, or pivot.

What is the value of a ”scientific approach” compared to other approaches? The thrust of our work is that the scientific approach has two main effects on entrepreneurial decision making. First, it improves entrepreneurs’ ability to develop a more objective and conservative assessment of the value of the business, reducing the impact of decision making biases such as, for instance, overconfidence. We call this the *debiasing effect* of the scientific method. The development of a general theory of the problem and its articulation into hypotheses, helps scientific entrepreneurs focusing on the relevant assumptions behind the business idea that need to hold for the value proposition to generate value, effectively leading to the formulation of more structured prior probabilities. This is complemented with the design of high quality tests that can provide them with signals about the extent to which their theory and hypotheses, and their priors more in general, are actually supported by data from the environment. Relating signals received back to the broad theory leads to a validation of the theory or to a rejection of it. This results in an update of their priors toward something more objective and to a more conservative expectation on the value of the idea. If this is too low, entrepreneurs may be thus choose to terminate the project. For instance, if our entrepreneur collected a negative signal on people's willingness to share cars due to hygiene concerns in a pandemic world, she would be more likely to form a negative value expectation and terminate the project. This effect is likely to lead scientific entrepreneurs to terminate their projects more often than non-scientific entrepreneurs.

The second effect that the scientific approach has on entrepreneurial decision making is that it likely improves the ability of scientific entrepreneurs to identify the changes to the business proposition would lead it to develop more or less value more easily or more rapidly. We call this the *learning effect* of the scientific method. The development of a theory and its articulation into hypotheses leads to a clear identification of the core elements or the problem and the relationships between them. This facilitates a quicker and more efficient search of the solution space, as it leads actors to identify ex-ante the characteristics of the solution (e.g., Camuffo et al., 2020, Felin, Kauffman, and Zenger, 2020). For example, if our entrepreneur obtained a positive signal on her hypotheses that car transportation in large cities is highly valued by households with young children who cannot use other types of transportation such as bikes easily, she would immediately be able to understand that this will also make the service appealing to households that include the elderly and could pivot in this direction. This effect is likely to lead scientific entrepreneurs to perform better, conditional on the fact that they do not terminate their project.

3 Structural Framework

Our theory suggests that there are two effects associated with the application of a "scientific approach": a *debiasing effect* and a *learning effect.* The goal of this paper is that of disentangling and identifying both effects. Research on the impact of decision making on performance has been limited by the fact that studying this issue requires facing many challenges. One of this is that firm performance is observable only for firms that have not terminated their activities, creating a source of selection bias to deal with. To address this challenge, we develop a structural decision-making framework and estimate it with a multi-equation simultaneous maximum-likelihood model. Another challenge is that the choice to use a specific decision-making process is endogenous to firm performance, and to other firm characteristics. In our setting, as we will explain in greater detail, this challenge is mitigated by recurring to a randomized setting and to an exogenous treatment offered to entrepreneurs, thus leading to an ideally unbiased parameter estimation.

We start with a *value equation.* We consider the realized value (v) of the business idea and model it as:

$$v = a + \theta T + \sigma \epsilon \tag{1}$$

Where $a = \gamma X$, with X being a set of controls recorded at the baseline period. We assume that the realized value of the business is a function of a set of controls X and of whether the entrepreneur has been trained with the scientific approach. Purposefully, the dummy T separates entrepreneurs trained with the "scientific approach" with entrepreneurs in the control group, hence θ identifies what we have labeled to be the *learning effect.*

Our model postulates that entrepreneurs explore their business ideas and form expectations of their potential value and probability of success over time. Let us denote these expectations by $\hat{v}$. We assume that entrepreneurs decide to keep developing their business if such expectations are higher than their outside option w. Therefore, in our framework, such estimations are crucial as the decision between continuing with the development of the business or terminating the project is based on the evaluation of that expectation with respect to an individual outside option.

In our model, we represent the entrepreneur's decision making process as characterized by four crucial points in time: (i) the baseline, before the training (0 - the *Baseline Evaluation*), (ii) after the training (E - the *Early Evaluation*), (iii) later in time after the training (L - the *Late Evaluation*), and (iv) at the time of the decision whether to remain active or terminate the business (F - the *Final Evaluation*). To clarify what we mean with the *Late Evaluation* period, it is worth explaining briefly the structure of our data collection process. We follow entrepreneurs for multiple data points, recording whether they are still operational in the market at each of them. Once an entrepreneur decides to terminate his/her project, our data collection reaches a natural end. Hence, we consider as *Late Evaluation* the last available data point before such decision. If an entrepreneur never terminates the project within the observation window, we consider as *Late Evaluation* the end of our observation window. Hence, we develop four equations:

$$\hat{v_0} = c + c_0 + \sigma_0 \epsilon \tag{2}$$

$$\hat{v_E} = c + c_0' + c_E + (c_{ET} + \theta)T + \sigma_E \epsilon \tag{3}$$

$$\hat{v_L} = c + c_0' + c_L + (c_{LT} + \theta)T + \sigma_L \epsilon \tag{4}$$

$$\hat{v_F} = c + c_0' + c_F + (c_{FT} + \theta)T + \sigma_F \epsilon \tag{5}$$

The baseline evaluation (Eq. 2) happens before the training and therefore it depends on a series of factors independent of the training, such as education levels, age or previous startup experience, which we include in the vector c_0, whereas c represents a constant term.

Once the intervention starts, we assume it to have an effect on the evaluation. Equations relating the early (Eq. 3), late (Eq. 4) and final (Eq. 5) evaluations include c, which is the constant term, c_0' that identifies constant idiosyncratic factors as above but that we assume could vary in terms of magnitude (as represented by the apostrophe), and c_j (where $j = E, L, F$), which identify contemporaneous factors affecting the value estimation.

In addition to this, we assume that the intervention has two different effects on the evaluation made by entrepreneurs, i.e.the *learning effect* θ and the *debiasing effect* c_{jT} (which is not restricted to be constant over time) cannot be identified empirically. This is because we assume that the scientific approach helps entrepreneurs understand the opportunity for positive performance since the beginning of its application, but that its debiasing effect might vary overtime. To achieve the goal of this paper of identifying these two effects, additional structure in our model is needed.

We first build on the previous steps, and generalize the decision-making process as:

$$\hat{v_j} = c + c_0' + c_j + (c_{jT} + \theta)T + \sigma_j \epsilon \geq w_j \tag{6}$$

With j representing the different time periods, and w_j representing the entrepreneur's outside option (which we assume vary over time and that we represent as $w_j{=}w_{j-1}{+}b_j$). This condition is verified if and only if:

$$\epsilon \geq \frac{w_j - c - c_0' - c_j - (c_j T + \theta)T}{\sigma_j} = z_j \tag{7}$$

We relabel the right hand side of equation (Eq. 7) z_j. When the decision of staying in the market is made (which we labeled with F), for entrepreneurs who choose to terminate their project, we cannot observe the values above. Rather, we only observe the final outcome. Hence, for F, we consider the following equation based on a latent model for the probability of remaining active in the market:

$$Pr(Stay) = \Phi(\frac{-w_F + c + c_0' + c_L + (c_{FT} + \theta)T}{\sigma_F}) \tag{8}$$

We can now re-arrange some equations to retrieve the structural parameters of interest. Let us rewrite Eq. 7 for the first three data points (0, E and L).

$$z_0 = \frac{w_0 - c - c_0}{\sigma_0} \tag{9}$$

$$z_E = \frac{w_E - c - c_0' - c_E - (c_{ET} + \theta)T}{\sigma_E} \tag{10}$$

$$z_L = \frac{w_L - c - c_0' - c_L - (c_{LT} + \theta)T}{\sigma_L} \tag{11}$$

Plugging Eq. 11 into Eq. 8, Eq. 10 into Eq. 11 and Eq. 9 into Eq. 10, we obtain:

$$z_E = \frac{\sigma_0 z_0 + b_E - (c_0' - c_0) - c_E - (c_{ET} + \theta)T}{\sigma_E} \tag{12}$$

$$z_L = \frac{\sigma_E z_E + b_L - (c_L - c_E) - (c_{LT} - c_{ET})T}{\sigma_L} \tag{13}$$

$$Pr(Stay) = \Phi(\frac{-\sigma_L z_L - b_F + (c_F - c_L) + (c_{FT} - c_{LT})T)}{\sigma_F}) \tag{14}$$

We now turn to the description of the data and methodology used to estimate the whole model and the structural coefficients of interest.

4 Methodology and Data

4.1 Experimental Design

To estimate the structural framework we leverage data from two field experiments, delivered in the context of a business support program that was offered to entrepreneurs in Milan and Turin (Italy). Both RCTs shared the same structure, type of intervention, and data collection process. The two RCTs were held asynchronously.

Both programs were advertised nationally over multiple offline and online channels. The advertisement campaign lasted for several weeks to ensure recruitment of at least 100 entrepreneurs per batch. The campaign promoted the program as a cutting edge business support program, offered free of charge to early stage entrepreneurs operating in any industry. The focus on early stage startups ensured that participants into the programs were highly involved in the decision making process. To apply, entrepreneurs were required to fill in an online survey and complete a telephone interview. In total, the first RCT (Milan) recruited 250 entrepreneurs, and the second (Turin) recruited 127.

Entrepreneurs were assigned to either a treatment or a control group through simple randomization. We checked that the randomization was successful with a set of balance checks across groups. Then, each group was broken down into smaller groups and assigned to an instructor, thus creating different classes of entrepreneurs. Entrepreneurs in both groups attended the same number of sessions. All the sessions were highly experiential and the division in small classes ensured that instructors provided feedback to each participant. Both groups of entrepreneurs were exposed to general managerial frameworks (such as the balance scorecard or the business model canvas) and to data gathering techniques (such as interview techniques, surveys and A/B testing). However, the treatment group was taught to apply this content using a scientific approach. Treated entrepreneurs learnt to develop a theory of the problem faced, to develop hypotheses that flow logically from it, and to use the evidence gathering techniques to test those hypotheses and relate the results back to the theory.

For instance, both groups were exposed to the Business Model Canvas (BMC), a highly used tool in entrepreneurship, which helps entrepreneurs graphically schematize a company's business model. Entrepreneurs in the control group were exposed to this method and taught to apply it to their business. Instead, treated entrepreneurs were taught to use the BMC as a starting point for their theoretical reasoning. Each component of the BMC was translated in an hypotheses to be tested. In later sessions entrepreneurs were exposed to different testing designs. Entrepreneurs in the control group were generally encouraged to apply these techniques to the problems they were facing in their business, whereas treated entrepreneurs were explicitly encouraged to use those techniques to test the hypotheses developed in the previous sessions.

Contamination between treatment groups was prevented by scheduling classes in different days or times of the week, according to the offered training. Moreover, all the communication was separated by treatment group and the research team checked whether applicants to the program knew other applicants, allocating them to the same experimental group.

4.2 Data Collection Process

We asked entrepreneurs to provide data on their decision-making processes and business performance throughout the training program for up to 66 weeks after the beginning of the training programs to a team of research assistants (RAs) via a set of phone interviews. RAs were purposefully trained by the research team and were responsible of conducting monthly telephone interviews with entrepreneurs. Overall, for each entrepreneur we collected the baseline and up to 18 data points.

Each phone interview was based on a standardized semi-structured interview script, including both open and closed-ended questions. Inquired topics included business performance, decision-making practices and any change introduced to the business idea. Each interview was recorded and stored in an encrypted storage, while RAs were also instructed to encode qualitative answers into quantitative information.

Each entrepreneur was interviewed up until the end of the project of up until the time they declared to have terminated the development of their business idea; thus, for firms that exited the market we have information only up to such exit decision. For firms that did not terminate before the end of our observation window, we have information up to 66 weeks after the beginning of the study.

4.3 Data and Estimation technique

We turn now to the description of the data employed in the empirical estimation of the structural model. To allow for a consistent estimation, we collapse our panel dataset into a cross-sectional form, creating distinct variables corresponding to the three mentioned data points.

To measure selection, i.e. entrepreneurs whose projects are still operational at the end of our investigation period, we create a dummy variable that takes value 1 for entrepreneurs that are still operating in the market and 0 for those that instead terminate their project at any point in time. For the former, we measure overall performance (or value) by computing the revenue growth between the first (baseline) and last available data point. To measure entrepreneurs' perceived value or estimation of future value, we rely on survey and interview data recording two main components. First, we asked entrepreneurs to provide a predicted probability of termination at the baseline, early and late data points on a scale from 0 to 100. Second, we asked entrepreneurs to directly estimate the minimum and maximum potential future value of their business ideas, on a scale from 0 to 100. To compute the estimated value, we take the logged average of the two. Also in this case, we record entrepreneurs' own estimations at three main data points: before the start of the training (baseline), after eight weeks from the first class and in the last available data point. The latter means that, for entrepreneurs that remained active in the market, we have the full set of information. Instead, for entrepreneurs that terminated, we have information up to the data point prior to which they declared having terminated, which we treat as our "last available" data point. Finally, as to capture idiosyncratic factors that could affect both the project value and entrepreneurs' estimations, we employ pre-training data on team size (number of people in the founding team), team average age, weekly hours worked, years of experience with startups and the team-average education levels.

Table 1 includes some descriptive statistics about these variables by treatment group.

Table 1: Descriptive Statistics

	Scientific		Control		Total	
	Mean	SD	Mean	SD	Mean	SD
Revenue Growth (Stay = 1)	1.70	3.486	1.05	2.687	1.33	3.069
Stay (Dummy)	0.54	0.499	0.69	0.465	0.61	0.488
Probability of Termination (Baseline)	0.17	0.196	0.21	0.209	0,19	0,20
Probability of Termination (Week 8)	0.17	0.228	0.16	0.192	0,16	0.21
Probability of Termination (Last)	0.26	0.288	0,29	0.287	0.27	0.29
Estimated Value (Baseline - log)	4.16	0.285	4,12	0.299	4.14	0.29
Estimated Value (Week 8 - log)	4.06	0.403	4.11	0.302	4,09	0.36
Estimated Value (Last - log)	3.99	0.478	3.99	0.400	3.99	0.44
Startup Experience (Years)	1.28	3.082	1.21	2.323	1.24	2.72
Team Size (Baseline)	2.32	1.441	2.24	1.365	2.277	1.40
Education	1.91	0.794	1.99	0.906	1.950	0,85
Age	31.19	8.541	31.10	7.635	31.145	8.08
Hours Worked (Baseline)	12.93	18.851	12.92	19.385	12.922	19.09

By assuming a cumulative normal distribution, we can estimate the value of z_j by simply calculating the inverse of the latter given the predicted probabilities of termination (p_j). Mathematically, since $p_j = \Phi(\frac{w_j - c - c'_0 - (c_{jT} + \theta)T}{\sigma_j}) = \Phi(z_j)$, we can retrieve z_j as:

$$z_j = \Phi^{-1}(p_j) \tag{15}$$

We cannot know the z estimate at the exact time in which the decision has been taken (what we labelled with F). We, therefore, employ a selection model, where we include as our selection variable the estimate z_L, and we rely on a latent estimation for such probability.

If we were to only estimate the first two equations of the structural model described above, this could be done with a standard Heckman model where the exclusion restriction would be satisfied by the inclusion in the selection equation of the estimate z_L from Eq. 15 evaluated in the *late* period. This would allow us to estimate the *learning effect* θ conditional on the decision to stay in the market. However, relying solely on such two equations does not allow us to estimate the *debiasing effect*.

The opportunity to leverage data on the entrepreneurs' estimation of the potential value of their ideas enables us to retrieve all the structural parameters of interest and be able to separate the *debiasing* effect from the *learning effect*. Particularly, we leverage on the first two post-training data points (E and L) and consider such predicted values for two additional equations, that we label with *. Indeed, it is the availability of the own estimations by entrepreneurs that allow to estimate empirically both Eq. 3* and Eq. 4* and ultimately retrieve the two variances σ_E and σ_L that allow us to estimate the variance σ_F from Eq. 8. This additional step is what allows us to identify the *debiasing effect* in the three different data points we are considering. Indeed, by estimating the three variances, we are able to subtract θ from the estimated coefficients on T in Eq. 14 and Eq. 12 and finally compute the debiasing effect for Eq. 13.

We thus end up with the following structural model to be estimated, made up of six equations:

$$v = a + \theta T + \sigma\epsilon \tag{1}$$

$$Pr(Stay) = \Phi(\frac{-\sigma_L z_L - b_F + (c_F - c_L) + (c_{FT} - c_{LT})T)}{\sigma_F}) \tag{14}$$

$$z_L = \frac{\sigma_E z_E + b_L - (c_L - c_E) - (c_{LT} - c_{ET})T}{\sigma_L} \tag{13}$$

$$z_E = \frac{\sigma_0 z_0 + b_E - (c_0' - c_0) - c_E - (c_{ET} + \theta)T}{\sigma_E} \tag{12}$$

$$\hat{v_L^*} = c + c_0' + c_L + (c_{LT} + \theta)T + \sigma_L\epsilon \tag{4*}$$

$$\hat{v_E^*} = c + c_0' + c_E + (c_{ET} + \theta)T + \sigma_E\epsilon \tag{3*}$$

We estimate these equations through a multi-equation conditional mixed-process estimator using the `cmp` user-written command in STATA 16 (Roodman, 2011). The fitted algorithm is a modified version of a seemingly unrelated regressions estimator. In other words, it employs a maximum likelihood (ML) estimator with the assumption that the errors from the different, independent, equations are distributed according to a joint normal distribution. The `cmp` estimator allows us to model a simultaneous equation framework where endogenous variables in a multi-staged process appear both on the right and left end sides of six empirical equations representing the structural model described in the previous subsection. We estimate the following set of empirical equations, linked to the structural equations above:

Eq. $1: v^* = \alpha_v + \beta_v X + \theta T + D + \epsilon_v$

Eq. $14: \Phi(\alpha_F + \gamma_F z_L + \beta_F T + D)$

$$\text{Eq. 13}: z_L = \alpha_L + \gamma_L z_E + \beta_L T + +\epsilon_L$$

$$\text{Eq. 12}: z_E = \alpha_E + \gamma_E z_0 + \beta_E T + \lambda_E X + D + \epsilon_E$$

$$\text{Eq. 4*}: v_L^* = \alpha_{v_L} + \beta_{v_L} T + \lambda_{v_L} X + D + \epsilon_{v_L}$$

$$\text{Eq. 3*}: v_E^* = \alpha_{v_E} + \beta_{v_E} T + \lambda_{v_E} X + D + \epsilon_{v_E}$$

Where D is a set of dummies for RCT and class instructors, X is a set of controls recorded at the baseline period as described above and the α represent constant terms of each equation. All the equations are linearly estimated, but the selection one (Eq. 14) which follows a probit model. Again, equations are estimated simultaneously assuming a joint normal distribution of the error terms. Standard errors are clustered by classroom.

From the estimated coefficients of the above regressions, we can thus retrieve all the parameters of interest that belong to our theoretical structural model. Specifically, the *learning effect* is straightforwardly estimated from the first equation, and it is the coefficient θ on the treatment dummy computed from the first model. All the other structural coefficient have instead to be computed leveraging on the estimated variances and coefficients from the econometric models. Particularly, the computation entails a non-linear combination of different estimated parameters. We conduct such computation using the `nlcom` routine on STATA.

Retrieving the OLS variances from Eq. 3* and Eq. 4*, we can estimate the variance of the model related to the decision (selection equation) from Eq. 14 and we compute all the structural coefficient related to the *debiasing effect* at different points in time from the other equations. Recall that in all equations but the value one, the estimated coefficient on the treatment dummy captures both the hypothesized effects. Thanks to the estimation of variances, we can subtract the estimated *learning effect* (θ) from such coefficients and finally retrieve the correct estimation for the *debiasing effect*. Table 2 details the calculations.

Table 2: Structural Parameter Computation

Parameter	Computation	Equations
θ	θ	1
σ_E	*OLS variance*	3*
σ_L	*OLS variance*	4*
σ_F	$-\frac{\sigma_L}{\lambda F}$	14, 4*
c_{ET}	$\beta_{v_E} - \theta$	3*, 1
c_{LT}	$\beta_{v_L} - \theta$	4*, 1
c_{FT}	$\beta_F \sigma_F + c_{LT}$	14, 4*

This computation strategy leverages on the straightforward calculations from Eq, 3* and Eq. 4*. An alternative computation strategy is shown in the Appendix.

5 Results

5.1 Structural Estimation Results

Table 3 reports the results of the structural estimation.

Table 3: Estimated Structural Parameters

	Parameter	Std. Err	z-score
θ	0.81	0.390	2.06
σ_E	0.35	0.032	10.70
σ_L	0.43	0.040	10.88
σ_F	1.81	1.580	1.15
c_{ET}	-0.86	0.387	-2.21
c_{LT}	-0.82	0.374	-2.20
c_{FT}	-1.68	0.593	-2.83

Structural parameters retrieved after ML estimation of the six equations described in Section 4.3. Parameters and their standard errors are retrieved using the `nlcom` routine in Stata 16. All estimated equations include dummies for RCT and instructor, with standard errors clustered at the classroom level.

Our estimation results show a positive and significant θ coefficient, that represents what we called the *learning effect*. On average, "scientific entrepreneurs" experience an increase in revenues of around 80 percentage points compared to traditional ones, conditional on the decision to stay on the market.

The three variance estimates, σ_E, σ_L, σ_F grow in magnitude as the final decision is taken. Particularly, the variance related to the decision equation is around six times the variance experienced in earlier stages (1.81 vs 0.35).

Parameters c_{ET}, c_{LT} and c_{FT} are those identifying the *debiasing effect* in the three decision periods that we consider in our structural framework. Results show that "scientific entrepreneurs" are more likely to perceive a lower potential future value of their business ideas at the time of the decision (c_{FT} =-1.68). The debiasing effect, however, materializes already at an early stage, i.e. eight weeks after the beginning of the training, and persists over time.

In the Appendix, we also report the results of the six equations estimated via `cmp` and an alternative computation of structural parameters. To test the robustness of the following results, we also run a number of checks. We first include additional controls in all empirical equations to take into account of some imbalances between groups prior to the training. Then, we employ an alternative measurement for z, where - instead of considering the last available data point - we considered the previous one. Results are consistent with our main analysis and are available upon request.

Overall, these results suggest that "scientific entrepreneurs" perform better, on average, compared to traditional entrepreneurs, even when taking into account the effect of selection (the decision to not terminate the project). They also show that they tend to make a downward adjustment to their estimation of their business' ideas values. This downward adjustment on the potential future value of their business idea, is what we believe is the mechanism behind the higher rate of projects' termination by "scientific entrepreneurs" shown in Camuffo et al., 2020. Thus, despite the positive *learning effect* that would allow "scientific entrepreneurs" to perform better on average, this mechanism of reduction in the value of their expectations is what drives more firms towards market exit.

To provide additional evidence in support of this mechanism, we look at the estimates made by entrepreneurs on the potential value of their ideas at different points in time. We compare the averages of entrepreneurs estimation across four groups defined by two dimensions: whether the entrepreneur belonged to the treatment versus control group and whether they terminated the project within the observation window. We compute a difference-in-difference calculation to understand the magnitude of the debiasing effect. Table 4 reports results for the baseline (i.e. pre-training) period, 8 weeks after the beginning of the training and at the last available data point.

Table 4: Entrepreneurs' evaluations

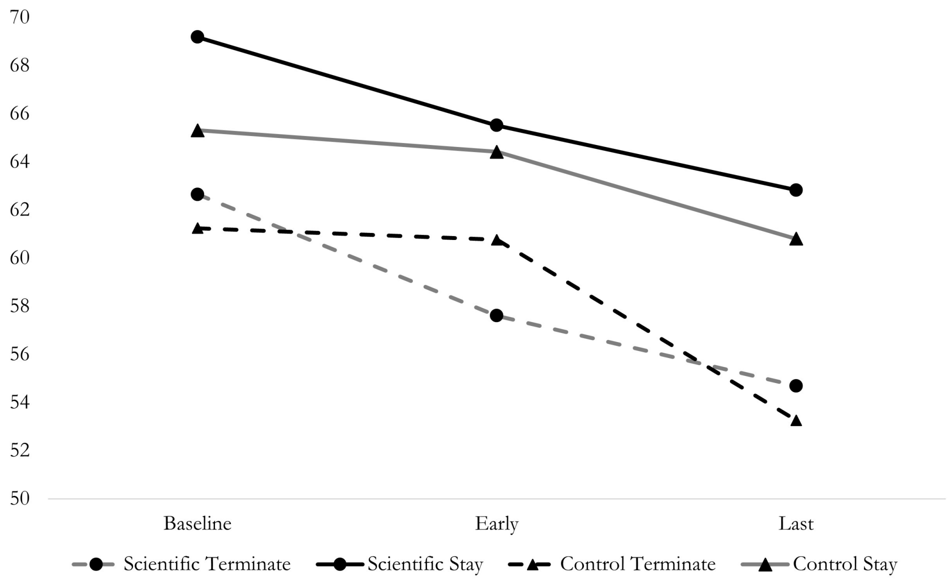
	Baseline Period				*Early Evaluation (8 weeks)*		
	Terminate	Stay	Difference		Terminate	Stay	Difference
Control	61.25	65.32	-4.074	Control	60.77	64.62	-3.855
Scientific	62.65	69.19	-6.539	Scientific	57.61	65.52	-7.910
Difference	-1.401	-3.866	2.464	Difference	3.156	-0.898	4.054

Last Available Evaluation

	Terminate	Stay	Difference
Control	53.27	60,81	-7.542
Scientific	54.70	62,83	-8.132
Difference	-1.431	-2.021	0.590

Table 4 provide some relevant insights that are in line with the estimates presented in Table 3. First, projects that were not terminated (Stay) show higher estimation value than those that were terminated (Terminate), in line with the idea that entrepreneurs, on average, terminate the projects that they assess to have a lower value. Second, for all groups, estimates are progressively lower over time. Third, the path of reduction is different for "scientific" and "control" entrepreneurs. The difference-in-difference calculation increases from the baseline to the early evaluation period, showing how Scientific entrepreneurs significantly reduced their own estimation at the very early stages[1], regardless on whether their final decision was to stay in or exit from the market. There is also a further reduction when looking at the last period available, which changes from firm to firm according to their period of exit. When looking at "control" entrepreneurs, we notice, instead, that the bigger drop happens at a later stage. This pattern clearly shows up graphically (Figure 1).

Figure 1: Entrepreneurs' Evaluations



[1]To check the statistical significance of the DiD coefficients, we run three models regressing the three self-evaluations on the treatment dummy, the terminate dummy and the interaction between the two (plus controls for RCT and instructors, clustering standard errors by classroom). While the DiD coefficient (interaction) is the highest in the Early period, we find no statistical significance at standard levels. Despite this result, that could be due to several reasons, we still acknowledge the fact that the difference between treatment groups in the "terminate" condition becomes positive only in the Early evaluation period, signalling the effect of different patterns of updates.

Overall, these results are in line with the idea that "scientific" entrepreneurs tend to be more cautious when estimating the potential success and value of their business ideas. This happens gradually over time, but, interestingly, it happens earlier for treated than for control entrepreneurs. This seems in line with the idea that scientific entrepreneurs reach more quickly a more realistic assessment of their ideas, and potentially terminate bad projects earlier, thus saving time, money and resources.

5.2 The Trade-Off between Retaining and Discarding Ideas for Scientific Entrepreneurs

In the previous section, we have shown that treated entrepreneurs are more cautious in evaluating their ideas and that they reach a more cautious evaluation more quickly than control. This result is in line with the idea that scientific entrepreneurs reach more quickly a more realistic evaluation of the idea, and more quickly identify "falsely positive ideas".

We cannot exclude the possibility that the treatment rather reduces the confidence of entrepreneurs, leading them to discard truly good ideas, increasing - in other words - the number of falsely negative ideas that they terminate. This is a very important question, but one that is, nevertheless, not trivial to answer. Answering this question would require knowing what could have been the "true" realized value of terminated projects, were they not terminated; this is clearly not possible.

However, we provide some suggestive evidence that might at least partially address this concern. First, whereas objective measures regarding the "true" value of ideas are not available, we look at whether firms have received financing from external investors over time, interpreting is as a more "objective" measure of the value of ideas. We, therefore, create a dummy taking value 1 if the firm has received financing within the observation period, and 0 otherwise. In Table 5 we report, for each cell, the share of firms having received external financing, separated by intervention (treatment vs. control) and final decision (termination vs. stay in the market).

Table 5: Share of Firms Having Received External Financing

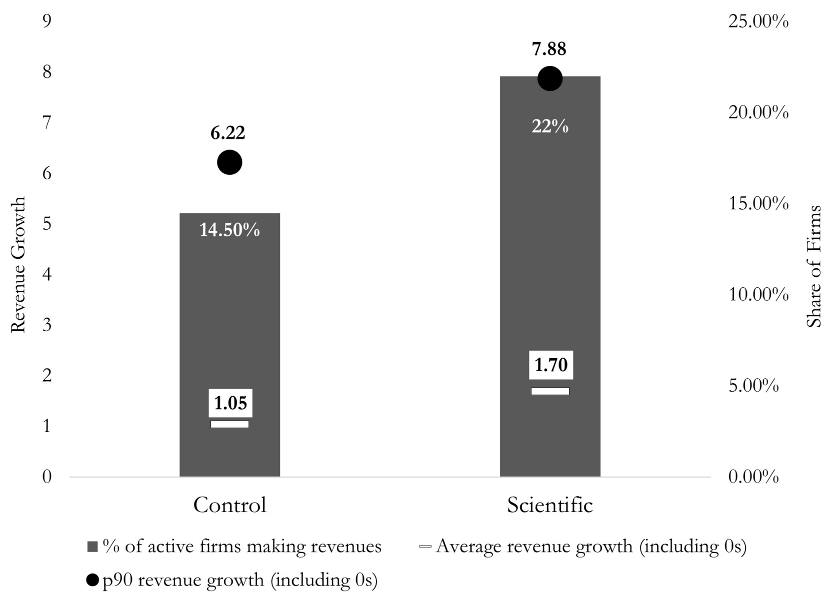
	Terminate	Stay	Difference
Control	2%	10%	-8%
Scientific	1%	19%	-18%
Difference	1%	-9%	10%

Looking at results for the treatment group, only 1% (1 firm out of 85) of firms that terminated the project collected external finance before their decision to terminate. This share corresponds, instead, to 2% (1 firm out of 59) for those in the control group. This goes in the direction of suggesting that ideas terminated by treated entrepreneurs are not better than those terminated by control entrepreneurs. Conversely, looking at entrepreneurs that decided to stay in the market, we see that 19% (19 firms out of 100) of treated entrepreneurs received external financing, which corresponds to almost double the 10% (13 firms out of 131) recorded in the "control" group. [2]

These numbers are consistent with the intuition that projects retained by scientific entrepreneurs tend to be of higher quality, with a selection of false positive projects taking place. Projects that were retained (i.e. stayed in the market) were on average more appreciated by external investors, who we can safely assume were blind with respect to the decision-making approach adopted by entrepreneurs. This is a strong signal towards our theory that the "scientific approach" helps selecting the best projects ex-ante.

Second, we provide some additional evidence on the distribution of revenues across treatment groups and exit decision, since it is also at the backbone of our structural estimation.

Figure 2: Additional Evidence on Revenue Growth



The columns indicate the share of firms with positive revenue growth, conditional on their decision to stay operational (right axis). The white bar and the black dot indicate, respectively, the 90th percentile and the average of the distribution of revenue growth (including 0s), conditional on the decision to stay operational (left axis).

It is worth noting that in both RCTs all firms started with no revenues, thus explaining the sizeable magnitude of the *learning effect*. Again, we compute the growth in revenues as the difference in logged revenues between the last available data point and the baseline, adding 1 to the latter value as to compute logs for the 0s.

The distribution of revenues and of revenue growth (especially at the end of the observation window) is indeed very skewed, with few firms having positive values. To further

[2]We also run a simple linear probability model, regressing the financing dummy on the interaction between the dummy for staying in the market and the intervention. We add as controls RCT and mentor dummies, clustering the standard errors by classroom. The coefficient on the interaction (i.e. the Difference-in-Differences coefficient) is significant at the 10% level.

explore this phenomenon, we created a dummy variable for firms still operational on the market at the end of the observation period, taking value 1 whether a firm shows a positive revenue growth. Figure 2 summarizes the share of firms making revenues together with two key moments of the revenue growth's distribution.

More precisely, only 14.5% of those in the control group (namely, 19/131) made revenues, versus the 22% (22/100) of those in the scientific group. [3]. Looking at the average revenue growth of all operational firms, including those with no revenue growth, Figure 2 shows an higher average revenue growth for scientific entrepreneurs. Whereas the medians for both groups are set to 0, the Figure shows that the value of 90th percentile is higher for scientific entrepreneurs.

This evidence brings further support to the results of our structural estimation, reinforcing the idea that scientific entrepreneurs make less false positives. Ideas that have been selected by scientific entrepreneurs, despite being fewer, have not only average higher revenues but also a higher chance of reaching the revenue stage

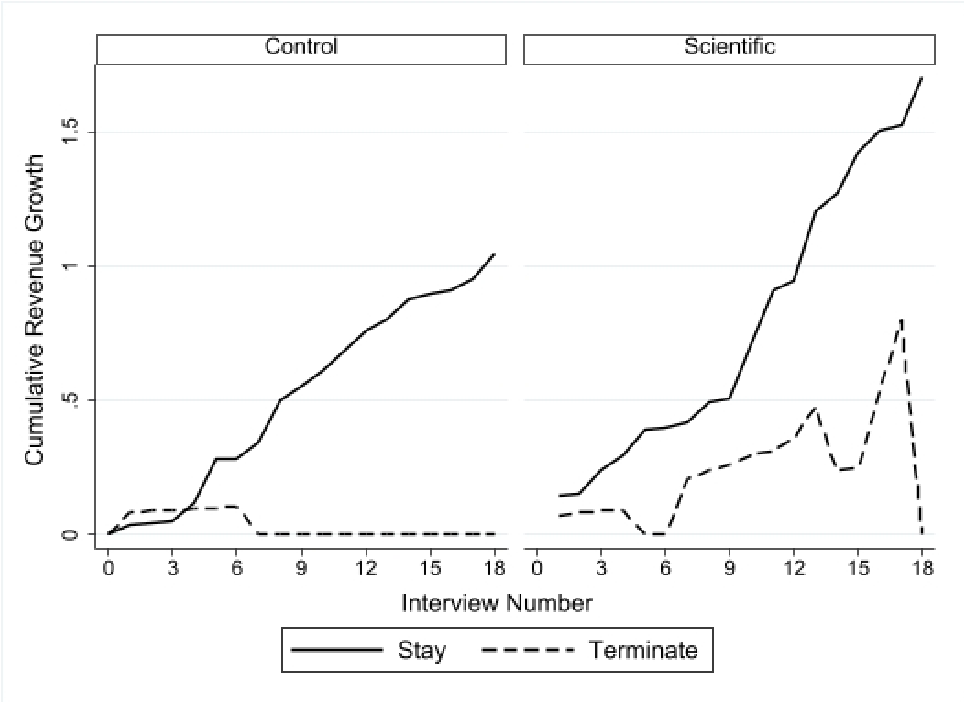
However, we are also interested in what happens on the false-negative side. To gather additional evidence, we follow the logic from Elfenbein and Knott, 2015 and classify firms into two types (*good*, or *bad*, based on whether they make revenues or not.

We thus leverage on the panel structure of our database and look at data on the revenue growth over time. Specifically, for each firm in the sample, we computed its cumulative revenue growth from the baseline to each observation in our panel. For firms that decide to remain active in the market, we expect a growing trend. For firms that terminate, we expect a more noisy pattern, as their revenue growth naturally goes to zero after their decision to terminate the project (and we conservatively set them as missing values in our database). We then compute the average by treatment group and the final termination decision. Figure 3 shows the results of these computations. The figure shows that, looking at firms that remain active, firms in the scientific group perform better, in line with previous findings.

Instead, what is more interesting is that firms in the scientific group that terminated their projects did made some revenues, although these revenues were lower than the ones of firms that stayed at the very same point in time. This is a first signal that, on average, ideas that were discarded performed less well than those that were not discarded, at least up to their termination decision.

[3] We also run a Probit with a Heckman selection model (`heckprobit`) using as a dependent variable the dummy recording positive revenues. The fitted model mimics the one run in the last two steps of the full structural model, using z_L as the selection variable. Results (not reported for the sake of brevity) are in line with the intuition that the probability of making revenues conditional on the decision to stay in the market is significantly higher for scientific entrepreneurs. We also run simpler tests (probit, t-test and chi2 test) on the subsample of operational firms, thus not accounting for selection. While the t-test shows a significant difference in the expected direction (one-tailed, $p = 0.07$, the other two tests do not show significant results.

Figure 3: Panel Data on Revenue Growth



The graph show the pattern of average revenue growth by treatment and final decision of staying or not in the market. For firms that exited, the value of revenue growth is set as missing after their decision to terminate, explaining the noisier pattern.).

When looking at the control group, our model suggests that the share of firms remaining operational in the market is likely to include false positives, but it could also includes projects that the treatment group would have discarded as false negatives. However, the facts that 1) the revenue growth of scientific firms that stay operational in the market is higher than that of control firms, and 2) that the revenue growth for scientific firms that terminated is on average always lower than the one of control firms that stayed, suggests that overall the reduction in false positives compensates the potential increase in false negatives experienced by the treatment group. To sum up, this suggestive evidence is in line with the idea that the scientific approach leads to a reduction in the rate of *false positives*, but to a less than proportional increase in the rate of *false negatives*.

Finally, to further support these insights, we go back to our structural model and focus on the first two equations, estimated with linear regression for the *value equation* (Eq. 1) and with a probit model for the selection equation (Eq. 14). We retrieve the correlation coefficient ρ between the two equations and the Mills' ratios from the selection equation for firms that stayed and terminate their projects.

Using the previously computed variances of the *value equation* (σ) and of the selection equation (σ_F), we can thus compute, for each entrepreneur in our sample, the expected value of the correction in the value equation by treatment condition and by decision to terminate the project or not. Mathematically, for firms that stayed in the market, this corresponds to:

$$correction = \rho \times \frac{\sigma}{\sigma_F} \times \frac{\phi(x\beta)}{\Phi(x\beta)} \tag{17}$$

Instead, for entrepreneurs that terminated, this corresponds to:

$$correction = -\rho \times \frac{\sigma}{\sigma_F} \times \frac{\phi(x\beta)}{(1 - \Phi(x\beta))} \tag{18}$$

The intuition behind this analysis is that the correction provides us with a measure of the extent to which the value of ideas needs to be adjusted due to the selection. A positive value of the correction suggests that entrepreneurs using a scientific approach underestimated the value of the project; a negative sign suggests that scientific entrepreneurs overestimate the value of the project. A positive difference in the correction between those who terminate and those who stay suggests that the underestimation of those who terminate is higher than the overestimation of those who stay. We are interested in the difference between control and treated entrepreneurs.

We compute these differences in Table 6, were we make the conservative assumption that the value model for entrepreneurs who remained active and those who terminate their projects is identical for entrepreneurs who terminate and those who stay. We call this the *lower bound* condition.

Table 6: Same Value Model for Terminate and Stay

	Terminate	Stay	Difference
Control	0.38 (*0.08*)	-0.17 (*0.07*)	0.557
Scientific	0.29 (*0.08*)	-0.25 (*0.07*)	0.543
Difference	0.092	0.078	0.014

Standard Deviation in Parentheses

The negative ρ coefficient estimated through the structural model leads to a negative correction for firms that stayed in the market. While it can be challenging to interpret such coefficient in the light of the Heckman selection model, the negative direction could be the signal of an insufficient reduction of the bias in the average entrepreneur of our sample. Such effect could be mostly driven by the weakest bias reduction provided by the "control" group, given the results from our structural estimation for the "scientific" entrepreneurs. Importantly, the difference-in-difference calculation leads to a number close to zero and not statistically significant (0.014). This suggests that there isn't a significant difference between treated and control entrepreneurs when it comes to the balance between overestimated and underestimated projects.

Our theory and empirics also suggest that "scientific" entrepreneurs perform better on average due to what we called "the *learning effect*". But under the stated assumption that the value model for entrepreneurs that terminated vs. did not terminate their project remains the same, the difference-in-difference estimation does not change.

We next relax the assumption that the value model does not change depending on whether projects were discarded or not and rather assume that the value model is different according to the decision taken. This assumption will lead us to what we call our *upper-bound* condition. Under this assumption, we subtract the *learning effect* $\theta = 0.81$ to value of the correction for the projects of scientific entrepreneurs who terminate their projects, which now becomes -0.51. We subtract the estimated learning effect since the value model we estimated already considers the treatment effect for scientific entrepreneurs. The negative correction signals the existence of a bias reduction also for entrepreneurs that terminated their projects. We report these results in Table 7. The difference-in-difference estimation becomes 0.82, suggesting that the selection results in a lower reduction of value for treated (vs. control) entrepreneurs. Bad ideas are effectively ruled out, without a substantial increase in the *false negative* rate.

Table 7: Different Value for Terminate and Stay

	Terminate	Stay	Difference
Control	0.38 (*0,08*)	-0.17 (*0.07*)	0.557
Scientific	-0.51 (*0.08*)	-0.25 (*0.07*)	-0.263
Difference	0.092	-0.727	0.819

Standard Deviation in Parentheses

These two cases represent two extreme cases, one where the selection induced by the "scientific approach" is particularly positive (the *upper-bound*) and one where the approach leads to some adverse selection processes (the *lower-bound*), but close to zero. Despite these results should be interpreted with caution as they are based on assumptions, we believe that they provide encouraging suggestive evidence of a well-balanced trade-off between the extent to which scientific entrepreneurs discard bad projects at the expense of good projects: in the worst case scenario (*lower bound*) these two effects essentially compensate each other, in the best case scenario (*upper bound*) the positive effect dominates the negative one.

6 Discussion and Conclusion

In this paper we have explored the implications of encouraging entrepreneurs to employ a "scientific-approach to decision making". This approach, based on developing a theory of the problem faced, a set of hypotheses logically flowing from it, a series of tests to validate those hypotheses and a disciplined evaluation of results, is expected to induce entrepreneurs to reason in more probabilistic terms. Our structural model predicts that entrepreneurs following this approach are more likely to terminate their projects, as a result of a *debiasing effect* that leads entrepreneurs to develop a more conservative estimation of the value of their ideas. It also predicts that treated entrepreneurs perform better because the scientific approach leads them to a better understanding of the problem and the solution space, an effect that we have called *learning effect*.

We estimated the structural model using data from two randomized control trial that involved 377 startups. The results validate the model and support the intuition that the method leads entrepreneurs to a being more conservative in selecting project, reducing the rate of "false positive", but also to enhance the value of any project they focus on.

To better understand the potential value of this finding for scholars and practitioners, we reflect upon the extent to which the conservative attitude of scientific entrepreneurs might actually lead them to increase their rate of "false negative", that is, of good projects that they discard. We provide suggestive evidence that supports the idea that this possible effect is more than compensated by the beneficial effect of the reduction in false positives.

Overall, we believe that these findings might inform existing research on innovation and entrepreneurship as well as policy and practice. Our development of a structural model, estimated with data from two randomized control trials, give us the opportunity to overcome some of the intrinsic limitations faced by studies in the area and enables us to disentangle and identify two separate effects that the use of probabilistic reasoning might have for entrepreneurs.

References

Ashraf, Nava, Abhijit Banerjee, and Vesall Nourani (2021). "Learning to Teach by Learning to Learn". en. In: *Working paper*, p. 115.

Bennett, Victor M. and Aaron K. Chatterji (2019). "The Entrepreneurial Process: Evidence from a Nationally Representative Survey". en. In: *Strategic Management Journal* n/a.n/a. ISSN: 1097-0266. DOI: 10.1002/smj.3077.

Bingham, Christopher B. and Kathleen M. Eisenhardt (2011). "Rational Heuristics: The 'Simple Rules' That Strategists Learn from Process Experience". en. In: *Strategic Management Journal* 32.13, pp. 1437–1464. ISSN: 1097-0266. DOI: 10.1002/smj.965.

Bingham, Christopher B. and Jerayr Haleblian (2012). "How Firms Learn Heuristics : Uncovering Missing Components of Organizational Learning". eng. In: *Strategic entrepreneurship journal : SEJ*. Strategic Entrepreneurship Journal : SEJ. - Chichester : Wiley & Sons, ISSN 1932-4391, ZDB-ID 2393106-1. - Vol. 6.2012, 2, p. 152-177 6.2.

Bloom, Nicholas and John Van Reenen (Nov. 2007). "Measuring and Explaining Management Practices Across Firms and Countries*". In: *The Quarterly Journal of Economics* 122.4, pp. 1351–1408. ISSN: 0033-5533. DOI: 10.1162/qjec.2007.122.4.1351.

Camuffo, Arnaldo et al. (Aug. 2020). "A Scientific Approach to Entrepreneurial Decision Making: Evidence from a Randomized Control Trial". In: *Management Science* 66.2, pp. 564–586. ISSN: 0025-1909. DOI: 10.1287/mnsc.2018.3249.

Eisenhardt, Kathleen M. and Christopher B. Bingham (Dec. 2017). "Superior Strategy in Entrepreneurial Settings: Thinking, Doing, and the Logic of Opportunity". en. In: *Strategy Science* 2.4, pp. 246–257. ISSN: 2333-2050, 2333-2077. DOI: 10.1287/stsc.2017.0045.

Elfenbein, Daniel W. and Anne Marie Knott (2015). "Time to exit: Rational, behavioral, and organizational delays". In: *Strategic Management Journal* 27952.7, pp. 957–975.

Felin, Teppo, Stuart Kauffman, and Todd Zenger (Feb. 2020). *Microfoundations of Resources: A Theory*. en. SSRN Scholarly Paper ID 3549865. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.3549865.

Felin, Teppo et al. (Aug. 2020). "Lean Startup and the Business Model: Experimentation Revisited". en. In: *Long Range Planning* 53.4, p. 101889. ISSN: 0024-6301. DOI: 10.1016/j.lrp.2019.06.002.

Folta, Timothy B. (1998). "Governance and Uncertainty: The Trade-off between Administrative Control and Commitment". In: *Strategic Management Journal* 19.11, pp. 1007–1028. ISSN: 0143-2095.

Gans, Joshua S., Scott Stern, and Jane Wu (2019). "Foundations of Entrepreneurial Strategy". en. In: *Strategic Management Journal* 40.5, pp. 736–756. ISSN: 1097-0266. DOI: 10.1002/smj.3010.

Kirtley, Jacqueline and Siobhan O'Mahony (2020). "What Is a Pivot? Explaining When and How Entrepreneurial Firms Decide to Make Strategic Change and Pivot". en. In: *Strategic Management Journal* n/a.n/a. ISSN: 1097-0266. DOI: 10.1002/smj.3131.

Lieberman, Marvin B., Gwendolyn K. Lee, and Timothy B. Folta (2017). "Entry, Exit, and the Potential for Resource Redeployment". en. In: *Strategic Management Journal* 38.3, pp. 526–544. ISSN: 1097-0266. DOI: 10.1002/smj.2501.

McGrath, Rita Gunther (1997). "A Real Options Logic for Initiating Technology Positioning Investments". In: *The Academy of Management Review* 22.4, pp. 974–996. ISSN: 0363-7425. DOI: 10.2307/259251.

NationalBusinessCapital (Jan. 2020). *2019 Small Business Failure Rate: Startup Statistics by Industry.*

Ries, Eric (Sept. 2011). *The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses.* en. Crown. ISBN: 978-0-307-88791-7.

Roodman, D (2011). "Fitting fully observed recursive mixed-process models with cmp". In: *Stata Journal* 1, pp. 159–206.

Sarasvathy, Saras D. (2001). "Causation and Effectuation: Toward a Theoretical Shift from Economic Inevitability to Entrepreneurial Contingency". In: *The Academy of Management Review* 26.2, pp. 243–263. ISSN: 0363-7425. DOI: 10.2307/259121.

Thomke, Stefan H. (June 1998). "Managing Experimentation in the Design of New Products". In: *Management Science* 44.6, pp. 743–762. ISSN: 0025-1909. DOI: 10.1287/mnsc.44.6.743.

Tversky, Amos and Daniel Kahneman (Sept. 1974). "Judgment under Uncertainty: Heuristics and Biases". In: *Science* 185.4157, pp. 1124–1131. DOI: 10.1126/science.185.4157.1124.

von Hippel, Eric and Marcie J. Tyre (Jan. 1995). "How Learning by Doing Is Done: Problem Identification in Novel Process Equipment". en. In: *Research Policy* 24.1, pp. 1–12. ISSN: 0048-7333. DOI: 10.1016/0048-7333(93)00747-H.

Yang, Mu-Jeung et al. (2020). "How Do CEOs Make Strategy?" In: *NBER Working Paper* 36.

Appendix

Balance Checks

Table A1: Balance Checks Milan RCT

Variable Name	Description	Treatment Mean	Treatment SD	Control Mean	Control SD	Difference b	Difference p
Age	Age (Team Average)	31.47	8.18	31.41	7.90	-0.06	(0.950)
Analytic Thinking	Agreement on a 1-10 scale with the following statements (Team Average): "Analyzing the situation and looking at the evidence is critical to our company's decision-making", "We carefully assess all the possible alternatives before making a choice for our company", "We prefer to gather all the relevant information before making a decision for our company", "Multiple elements are taken into account when making a decision for our company, pros and cons are carefully evaluated in every situation"	8.38	3.68	8.07	3.28	-0.32	(0.475)
Background: Economics	Team members with an economics background (%)	0.41	0.42	0.31	0.37	-0.10**	(0.046)
Background: Other	Team members with no economics backgrounds (%)	0.22	0.36	0.20	0.33	-0.02	(0.696)
Background: STEM	Team members with a STEM (Science Technology Engineering Mathematics) backgrounds (%)	0.38	0.40	0.49	0.41	0.11**	(0.032)
Certainty	Agreement on a 1-10 scale with the following statements (Team Average): "We are sure about our business model", "We are sure about our strategy"	5.93	1.94	5.61	1.91	-0.32	(0.191)
Consensus	Answer on a 1-10 scale to the following questions (Team Average): "To what extent do you and your team members have consensus on the long term objectives of the firm?", "To what extent do you and your team members have consensus on the short term objectives of the firm?", "To what extent do you and your team members have consensus on the survival strategy of the firm?"	8.85	1.67	8.86	1.66	0.00	(0.990)
Education	Highest educational level attained by team members (5=PhD, 4=MBA, 3=MSc, 2=BA, 1=high school, 0=otherwise; Team Average)	1.94	0.74	1.95	0.80	0.00	(0.969)
Experience: Entrepreneurial	Number of years of entrepreneurial experience (Team Average)	1.09	2.19	0.93	1.44	-0.17	(0.480)
Experience: Industry	Number of years of experience in industry (Team Average)	2.84	3.82	2.33	3.62	-0.51	(0.280)
Experience: Managerial	Number of years of managerial experience (Team Average)	2.29	3.69	2.27	4.18	-0.02	(0.971)
Experience: Work	Number of years of work experience (Team Average)	8.73	7.75	9.02	8.85	0.28	(0.788)
Full Time	Percentage of team members working full-time	0.57	0.43	0.62	0.42	0.05	(0.390)
Gender (Female)	Proportion of women in the team	0.27	0.37	0.25	0.36	-0.03	(0.541)
Hours: Total Weekly	Weekly hours dedicated to the company (Team Average)	10.17	9.65	10.96	11.45	0.78	(0.560)
Idea Potential	Independent assessment of the value of the idea	47.22	21.22	47.31	23.25	0.09	(0.975)
Idea Value: Max	Maximum estimated value of the project (0 to 100)	85.08	16.29	85.67	16.16	0.59	(0.773)
Idea Value: Mean	Estimated value of the project (mean, 0 to 100)	65.40	15.53	64.52	16.69	-0.88	(0.668)
Idea Value: Min	Minimum estimated value of the project (0 to 100)	45.71	19.86	43.21	22.93	-2.50	(0.357)
Idea Value: Range	Estimated value of the project (range, 0 to 100)	39.37	18.85	42.46	20.99	3.10	(0.221)
Intuitive Thinking	Agreement on a 1-10 scale with the following statements (Team Average): "We are prone to following our intuitions when making company-related decisions", "We consider feelings and intuitions rather than analysis in our startup decisions", "First impressions are important when making decisions", "It is important to rely on gut feelings and intuition when making decisions"	4.09	1.70	3.83	1.74	-0.25	(0.244)
Lombardy	Dummy variable taking value of 1 when the majority of team members comes from the Italian region of Lombardy, 0 otherwise	0.56	0.47	0.57	0.46	0.01	(0.883)
Months to Revenue	Number of months to revenue	11.52	5.80	11.51	5.85	-0.01	(0.987)
Part Time	Percentage of team members working part-time	0.08	0.18	0.08	0.17	0.00	(0.941)
Probability Termination	Probability of terminating the project	31.64	32.53	32.35	31.60	0.70	(0.863)
Team Size	Number of team members	2.25	1.46	2.28	1.37	0.03	(0.858)
Observations		125		125		250	

Table A2: Balance Checks Turin RCT

Variable Name	Description	Treatment Mean	Treatment SD	Control Mean	Control SD	Difference b	Difference p
Age	Age (Team Average)	30.60	9.29	30.53	7.14	-0.07	(0.963)
Analytic Thinking	Agreement on a 1-5 scale with the following statements (Team Average): "Analyzing the situation and looking at the evidence is critical to our companyÕs decision-making", "We carefully assess all the possible alternatives before making a choice for our company", "We prefer to gather all the relevant information before making a decision for our company" and "Multiple elements are taken into account when making a decision for our company, pros and cons are carefully evaluated in every situation"	4.30	0.63	4.40	0.56	0.11	(0.318)
Background: Economics	Team members with Economics backgrounds (%)	0.18	0.31	0.20	0.36	0.02	(0.701)
Background: Other	Team members with no Economics/STEM backgrounds (%)	0.56	0.43	0.44	0.46	-0.11	(0.152)
Background: STEM	Team members with a STEM (Science Technology Engineering Mathematics) backgrounds (%)	0.26	0.38	0.36	0.45	0.09	(0.223)
Confidence	Agreement on a 1-5 scale with the following statements (Team Average): "We are confident in our entrepreneurial skills", "We are sure we are deploying the best strategy for our business", "We are confident in our ability to manage our business", "We master the competences necessary for our venture" and "We are sure there is no better business model for our idea"	3.41	0.52	3.32	0.65	-0.09	(0.397)
Currently Studying	Number of team members enrolled in an education program at the time of training	0.26	0.30	0.21	0.30	-0.04	(0.426)
Education	Highest educational level attained by team members (5=PhD, 4=MBA, 3=MSc, 2=BA, 1=high school, 0=otherwise; Team Average)	1.85	0.89	2.06	1.09	0.21	(0.240)
Experience: Business Plan	Dummy taking value of 1 if the team had years of experience in business plan design, 0 otherwise	0.26	0.36	0.35	0.43	0.09	(0.228)
Experience: Entrepreneurial	Number of years of entrepreneurial experience (Team Average)	1.65	4.38	1.73	3.37	0.08	(0.908)
Experience: Industry	Number of years of experience in industry (Team Average)	2.77	5.72	3.03	5.04	0.25	(0.792)
Experience: Managerial	Number of years of managerial experience (Team Average)	1.54	2.78	1.76	3.76	0.22	(0.705)
Gender (Female)	Proportion of women in the team	0.31	0.38	0.25	0.36	-0.06	(0.356)
Hours: Total Weekly	Weekly hours dedicated to the company (Team Average)	11.39	10.06	11.76	12.36	0.37	(0.853)
Idea Maturity	Maturity of the idea (in months)	9.32	9.43	11.98	11.63	2.66	(0.158)
Idea Potential	Independent assessment of the value of the idea (two evaluators, average) based on five criteria: innovation, feasibility, sustainability, team competence, market size	49.22	11.99	49.16	12.86	-0.06	(0.978)
Idea Value: Mean	Estimated value of the project (mean)	65.82	18.53	63.30	16.05	-2.52	(0.415)
Intuitive Thinking	Agreement on a 1-5 scale with the following statements (Team Average): "We are prone to following our intuitions when making company-related decisions" and "We consider feelings and intuitions rather than analysis in our startup decisions"	2.74	0.83	2.70	0.99	-0.03	(0.838)
Later Stage	Dummy variable taking value of 1 if the firm is at a more advanced stage than others, 0 otherwise	0.13	0.34	0.11	0.31	-0.03	(0.666)
Locus of Control	Agreement on a 1-7 scale with the following statements (Team Average): "In most jobs you need a lot of luck to excel", "One typically earns what they are worth", "To make money you just need to know the right people", "To get a good position you need luck", "Income is mainly the result of hard work", "There is a direct relationship between a personÕs abilities and the position he/she holds", "Many of the difficulties encountered at work concerns senior colleagues", "Generally, people who work well get rewarded", "Promotions are awarded to people who work well", "To find a good job, having a good network is more important than actual skills", "A well-trained person always finds a satisfycing job" and "To get a really good job you have to have high-level acquaintances"	3.84	0.67	3.79	0.70	-0.05	(0.707)
Months to Revenue	Number of months to revenue	12.69	11.37	14.68	10.58	1.99	(0.310)
Piedmont	Dummy variable taking value of 1 when the majority of team members comes from the Italian region of Piedmont and 0 otherwise	0.55	0.45	0.52	0.48	-0.03	(0.748)
Probability Pivot Idea	Probability of changing the business idea	31.89	22.96	32.53	26.75	0.65	(0.884)
Probability Pivot Other	Probability of changing other components of the business model	52.20	22.97	52.92	26.17	0.73	(0.868)
Probability Pivot Problem	Probability of changing the problem and customer segment	34.57	22.49	34.48	25.20	-0.09	(0.983)
Probability Termination	Probability of terminating the project	13.64	16.53	17.42	21.66	3.78	(0.268)
Risk-averse	Agreement on a 1-7 scale with the following statements (Team Average): "In important matters I never take unnecessary risks, which can be avoided", "In important situations I never deliberately chose to take risks I could have avoided", "I always try to avoid situations that put me at risk of getting into trouble with other people", "I am always very careful and I put safety first" and "I prefer to avoid doing things that expose me to criticism and liability"	4.23	1.03	3.96	1.04	-0.27	(0.151)
Risk-taker	Agreement on a 1-7 scale with the following statements (Team Average): "I can be pretty reckless and take some big risks", "I think I often act boldly and courageously", "I am a brave and daring person and I like to tempt fate in various situations", "There is a direct relationship between a personÕs abilities and the position he/she holds" and "I think I am often less cautious than other people"	4.04	1.13	3.98	0.91	-0.05	(0.766)
Scientific intensity: 1 Theory	Theory development score	2.92	1.32	3.05	1.20	0.13	(0.559)
Scientific intensity: 2 Hypothesis	Hypothesis development score	2.14	1.63	1.98	1.51	-0.16	(0.571)
Scientific intensity: 3 Test	Test score	1.32	1.73	1.29	1.69	-0.03	(0.919)
Scientific intensity: 4 Valuation	Valuation score	0.84	1.49	0.94	1.63	0.09	(0.742)
Self-efficacy	Agreement on a 1-7 scale with the following statements (Team Average): "I think I will always be able to achieve a goal even if I have to perform a difficult task", "Faced with new tasks and challenges, I am always confident that I will be able to complete them", "I am sure I will succeed", "When I have a goal, I almost always get better results than others", "When I take a test or an exam I am sure I can pass it successfully", "I am confident that my results will be recognized and appreciated by others", "I am not worried about difficult situations, because so far I have always managed to get by with my skills", "I never had any problem understanding and facing even the most complicated situations" and "I think I get the crux of the matter first"	5.46	1.07	5.57	0.96	0.11	(0.557)
Self-regulation	Agreement on a 1-7 scale with the following statements (Team Average): "People can count on me to meet the set and planned deadlines", "I can hardly say no", "I change my mind quite often", "Others would describe me as an impulsive person", "I wish I had more self-discipline", "I get carried away by my feelings", "I am not easily discouraged", "Sometimes I canÕt stop but do something, even though I know it is wrong", "I often act without thinking about all the alternatives", "I often do things that seem right in the present, even at the expense of future goals" and "When I pursue a goal I follow the original plan, even when I realize that it is not the best"	4.99	0.82	5.25	0.85	0.25*	(0.090)
Startcup	Dummy variable taking value of 1 if the firm takes part to a local competition, 0 otherwise	0.11	0.32	0.18	0.39	0.07	(0.290)
Team Size	Number of team members	2.51	1.48	2.14	1.36	-0.37	(0.144)
Observations		61		66		127	

Full Estimation Results

	Value	Selection	z_L	z_E	v_L^*	v_E^*
	(Eq 1)	(Eq 14)	(Eq 13)	(Eq 12)	(Eq 4*)	(Eq 3*)
Model	*OLS*	*Probit*	*OLS*	*OLS*	*OLS*	*OLS*
Treatment	0.805**	-0.471**	-0.163**	0.0512	-0.0168	-0.0502*
	(0.390)	(0.129)	(0.0769)	(0.0842)	(0.0501)	(0.0272)
Startup Experience	0.165**			0.00354	0.0103	0.00666
	(0.0664)			(0.0181)	(0.00714)	(0.00512)
Team Size (*Baseline*)	0.256			-0.139***	0.0199	0.00880
	(0.159)			(0.0435)	(0.0149)	(0.0139)
Education	0.255			0.119*	-0.0444*	-0.0330
	(0.223)			(0.0649)	(0.0239)	(0.0211)
Age	-0.0687***			-0.0236***	0.00447	0.00177
	(0.0223)			(0.00740)	(0.00311)	(0.00231)
Hours Worked (*Baseline*)	0.0103			-0.000326	0.000479	0.00226***
	(0.0124)			(0.00353)	(0.00143)	(0.000843)
z_L		-0.237				
		(0.201)				
z_E			0.992***			
			(0.168)			
z_0				0.316***		
				(0.0411)		
Constant	1.614*	0.313	0.382	-0.370	3.942***	4.161***
	(0.826)	(0.317)	(0.339)	(0.281)	(0.148)	(0.0633)
Equation σ (*OLS only*)	1.080***		0.165**	-0.00740	-0.844***	-1.058***
	(0.0778)		(0.0669)	(0.0373)	(0.0919)	(0.0935)
ρ (*Selection*)	-0.228*					
	(0.133)					
Observations	377					

All equations contain dummies for RCT and instructor, with standard errors clustered at the classroom level (in parentheses). *** p<0.01, ** p<0.05, * p<0.1

Alternative Computation of Structural Coefficients

Table A3: Structural Parameter Computation

Alternative computation from Z equations

Parameter	Computation	Equations
θ	θ	1
σ_E	*OLS variance*	3*
σ_L	*OLS variance*	4*
σ_F	$-\frac{\sigma_L}{\lambda F}$	14, 4*
c_{ET}	$beta_E \sigma_E - \theta$	3*, 12
c_{LT}	$-\beta_L \sigma_L + c_{ET}$	4*, 13, 12
c_{FT}	$\beta_F \sigma_F + c_{LT}$	14, 4*, 13, 12

Table A4: Estimated Structural Parameters

Alternative estimation from Z equations

	Parameter	Std. Err	z-score
θ	0.81	0.390	2.06
σ_F	0.35	0.032	10.70
σ_L	0.43	0.040	10.88
σ_F	1.81	1.580	1.15
c_{ET}	-0.82	0.372	-2.21
c_{LT}	-0.75	0.381	-1.98
c_{FT}	-1.61	0.596	-2.70

Structural parameters retrieved after ML estimation of the six equations described in Section 4.3. This alternative computation retrieves the parameters c_{ET} and c_{LT} from Eq. 12 and 13 rather than from Eq. 3* and Eq. 4*. Parameters and their standard errors are retrieved using the `nlcom` routine in Stata 16. All estimated equations include dummies for RCT and instructor, with standard errors clustered at the classroom level.