Research Policy 49 (2020) 104052

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Even winners need to learn: How government entrepreneurship programs can support innovative ventures

Mickaël Buffart[a,*], Grégoire Croidieu[b], Phillip H. Kim[c], Ray Bowman[d]

[a] Aalto University School of Business, Ekonominaukio 1, 02150 Espoo, Finland
[b] EM Lyon Business School, 23 Avenue Guy de Collongue, 69130 Écully, France
[c] Babson College, Blank Center for Entrepreneurship, 231 Forest Avenue, Babson Park, MA 02457, USA
[d] Grenoble École de Management & Economic Development Collaborative, Ventura County, Small Business Development Center (EDC SBDC), 4001 Mission Oaks Blvd Suite A-1 Camarillo, CA 93012, USA

ABSTRACT

Given the investment of public resources for supporting entrepreneurial growth, it is important to know whether such programs truly benefit innovative ventures. While prior research has indicated some benefits for growth outcomes, there is no clear consensus about the conditions for program effectiveness. We attribute this to the complex set of selection and treatment mechanisms associated with how programs navigate interlocking tradeoffs to maximize outcomes with their limited resources. To circumvent these challenges, policymakers often default to a "picking winners" approach based on past performance indicators. We develop and implement a carefully designed empirical strategy to determine whether this approach leads innovative ventures to achieve growth milestones and properly accounts for various observed and unobserved selection issues. We analyze data from the Small Business Development Center (SBDC), a government-sponsored program in the United States. Using a potential outcomes framework to investigate over 1,700 ventures that enrolled in SBDC advisory services from 2011 to 2016, we observe that treatment design is more crucial than selection for innovative firms to achieve growth. We found that treatment time and a client's willingness to learn collaboratively from their advisors are vital indicators of growth. Since treatment effectiveness is driven by support allocation, programs that desire to boost innovation outcomes must at a minimum formally prioritize innovation criteria to ensure these businesses receive sufficient support to address their growth objectives. Beyond this, we demonstrate that support effectiveness additionally depends on a willingness of participants to learn collaboratively by socializing their growth objectives with their advisors. Since even winners need to learn, programs must wrestle with the selection tradeoffs more acutely early on to ensure that the most promising clients can receive lengthier learning opportunities for growth.

1. Introduction

Governments around the world value a thriving innovation sector, since these businesses are expected to drive economic growth. To build capacity for this, policymakers focus on ways to promote innovative entrepreneurship – ventures that bring new products, services, technologies to market through novel business models and practices. (Aldrich and Ruef, 2018; Rubera and Kirca, 2012). By definition, innovative ventures are not only expected to generate an above-average performance in terms of growth and job creation (Colombelli et al., 2016), but also introduce new technologies and business practices, all of which are essential for an economy's global competitiveness (Arrow, 1962; Minniti, 2008). Once established, innovative ventures are expected to exhibit higher survival chances, grow faster, and generate positive knowledge spillover into the local economy (Audretsch and Keilbach, 2007; Colombelli et al., 2013). These ventures are exemplars of productive entrepreneurship because of their economic contributions to society (Baumol, 1990).

Despite their advantages, innovative ventures are difficult to establish. They are more likely to fail early, especially when based on novel technologies, since they require longer lead times to develop and may not address clearly defined market needs (Audretsch, 1995; Coad and Rao, 2008). Innovative business models and practices often require time to test and validate in the marketplace. One way to overcome these challenges is for policymakers to sponsor venture growth advising programs. Given that innovative ventures seem to have a higher growth potential and capacity to learn than other ventures (Clarysse et al., 2009; Colombelli et al., 2013), programs transferring knowledge to innovative ventures can assist them in achieving their growth potential.

With this innovative-based growth potential in mind, policymakers aim to disburse enough resources to fund entrepreneurship-support programs whose administrators welcome all qualifying clients who seek their services. However, in practice, public support programs often lack

* Corresponding author.
E-mail addresses: mickael.buffart@aalto.fi (M. Buffart), croidieu@em-lyon.com (G. Croidieu), pkim1@babson.edu (P.H. Kim), ray.bowman@edcsbdc.org (R. Bowman).

https://doi.org/10.1016/j.respol.2020.104052
Received 15 June 2020; Accepted 15 June 2020
Available online 01 July 2020
0048-7333/

sufficient resources to satisfy even the most basic demands for their assistance. Despite the public mandate of these initiatives, government-sponsored programs cannot support everyone and must take steps to invest only in those determined to yield the best outcomes. They must "do more with less." To tackle this dilemma, policymakers can focus their efforts by supporting the most productive innovative ventures as way to exercise stewardship over the limited public resources (Baumol, 1990).

Although focusing on productive innovative ventures appears as a straightforward policy mandate, the actual mechanisms to select and treat them are more complex. This requires designing support programs to recruit clients selectively and deliver high-quality advising to them. To maximize their limited support resources, programs face several tradeoffs. They must determine how to select the most promising clients and to allocate advising hours to those who are most likely to convert the support into tangible growth outcomes. However, anticipating which clients deserve more advising is challenging since it depends on using the correct selection criteria and treatment design. As in any selection process, these programs encounter incomplete information evaluating prospective clients (Huang and Pearce, 2015). They risk committing Type I errors ("false positives": waste limited resources by unnecessarily supporting clients who will not be dedicated) and Type II errors ("false negatives": prematurely rejecting high-potential clients or not recruiting them aggressively enough). One simple approach to tackle these dilemmas is to pick "winners" based on common, widely accepted criteria to ensure program effectiveness because winners are likely to grow even without receiving support (Shane, 2009). On balance, it provides an efficient framework to design programs while making the best use of limited public resources to generate stronger economic development (OECD, 2019; Woolley, 2017).

Our study investigates whether and how innovative ventures achieve growth outcomes from government-sponsored entrepreneurship programs. To answer this question, we need to evaluate the selection and treatment mechanisms involved with program design, which is a multi-faceted set of interlocking tradeoffs. Although these mechanisms are generic, we argue that they can be specifically applied to support programs that back innovative firms. We investigate if programs can employ the "right" selection criteria and if every qualifying client achieves growth by receiving the program treatment. Given that innovative ventures vary greatly in quality (Mustar et al., 2006), a picking-winners approach may amplify these errors and the misallocation of public funds. For this investigation, we analyzed the program operations of a Small Business Development Center (SBDC) in Southern California (from 2011 to 2015) – a publicly funded entrepreneurship support program overseen by the US Small Business Administration (SBA). We took advantage of this program's design to evaluate effectiveness of their selection and treatment approaches for innovative firms to achieve growth. We crafted a comprehensive identification strategy to determine whether treated innovative ventures achieved more growth than comparable non-treated businesses, net of selection biases, and if so, how program support translates into growth. We analyzed both quantitative and qualitative data about these ventures, their owners, and their advisors to study the selection and treatment mechanisms. We also examined Southern California businesses from Your-economy Time Series (YTS), an independent sample from the SBDC to complement our main analysis.

Our rigorous examination revealed that treatment design is more crucial than selection for innovative firms to achieve growth. We found that advising effort and a client's willingness to learn collaboratively from their advisors are vital predicators of growth. While the importance of these two predicators seem intuitive, they are actually difficult to navigate in practice because programs can only determine the true willingness of their clients to learn once they begin advising them. We estimated that clients willing to learn collaboratively require 3.4 times fewer advising hours than their counterparts to achieve the same level of growth outcomes (at the average amount of treatment level). We also calculated that the return on investment for the clients receiving a high level of advising is three times higher than for the clients receiving a low level. We also identified that most resources allocated to clients receiving a low level of advising was ineffective, since they were unable to achieve any growth outcome from this support. We estimated the cost of this low-level advising being three times higher for the same outcome achieved for the high-level advising clients. These findings reveal the extent to which support resources are misallocated away from the clients who can benefit the most. We attribute this to the sheer difficulties of anticipating who will learn, take full advantage of the support, and commit to implementing the assistance in their businesses solely with selection criteria. This means that significant resources are deployed to clients who are unlikely to learn and convert advice into growth outcomes (Type I error). At the same time, other high-potential clients could learn and achieve more growth outcomes if they received more advising support (Type II error). To improve growth outcomes, programs must wrestle with the selection tradeoffs more acutely early on to ensure that the most promising clients can receive lengthier learning opportunities for growth. Realistically, programs must accept some minimum misallocation as a tradeoff for identifying the more effective clients.

For entrepreneurship support programs targeted at innovative firms, our study offers practical implications. An emphasis on innovation can occur in two ways: 1) a greater emphasis on recruiting innovative firms by prioritizing it as a selection criterion and 2) a greater emphasis on learning new ways to innovate during treatment. This emphasis can occur only if the client is willing to learn from their advisors and work together to find solutions to their business problems. Often, these solutions depend on innovative business models and practices that can only be developed when advisors work hand-in-hand with their clients. Programs also need to allocate more treatment time to the most motivated firms willing to invest in their growth since innovation is a multi-stage process that requires time to unfold. The dominant approach in entrepreneurial policymaking is to "pick winners," which are high-potential and high-quality ventures that most benefit the economy and are likely to grow without the support of government programs (Baum and Silverman, 2004; Baumol; 1990; Minniti, 2008; Shane, 2009). However, on its own, just picking winners is an expensive strategy because it does not rely on any treatment effects to generate outcomes. For policymakers to achieve their growth objectives, we argue that "picking the winners" is just the first step. Even winners need to learn so advisors must be able to informally assess a venture's desire to seize learning opportunities and work collaboratively to generate growth. This will yield more growth outcomes, spur more innovation, and optimize public resources for nurturing more productive entrepreneurship in the economy.

2. Theoretical background

2.1. Government-support programs to boost innovative ventures

Governments everywhere have implemented policies to help innovative ventures launch and grow. These policies range from direct subsidies, tax cuts, working capital grants, business training, and counselling services (Autio et al., 2007). Given their anticipated contributions to overall economic growth, policymakers eagerly direct resources to these ventures by providing financial resources such as loans, R&D subsidies, contracts through agencies, or government-run venture capitalist funds (e.g. Brown and Earle, 2017; Hottenrott and Richstein, 2020). Building on their potential synergies, innovation and entrepreneurship policies often combine economic and technology initiatives to achieve their goals (Minniti, 2008; Woolley and Rottner, 2008). Examples include national "start-up acts" such as in Chile and France, where programs encourage foreign innovative ventures to launch and grow within their countries. The Start-up Act in Italy provided a variety of labor and financial policy incentives to facilitate intellectual property protection (Colombelli et al., 2019). The main objective of these policies is to nurture the innovative sectors of the broader economy as seen by the proliferation of "-tech" sectors such as cleantech, fintech, and edtech efforts. Through these policy interventions (also referred to as "input-related" policies), governments provide tactical support aimed at narrowing the knowledge and resource gaps entrepreneurs face as they launch and run their businesses and to trigger an increase in performance outputs, such as sales and job creation. (Audretsch et al., 2007).

2.1.1. Measuring entrepreneurial growth as milestones

One way to evaluate program effectiveness is to examine whether the interventions lead to entrepreneurial growth in the supported ventures. We define *entrepreneurial growth* as the extent to which new or existing businesses become more established or expand in size (Penrose, 1959; Achtenhagen et al., 2010). It is typically measured by an increase in sales or employees, since these indicators correlate to the extent to which businesses contribute to the overall economy's aggregated growth and expansion (Davidsson et al., 2006).

Despite the relevance of these growth indicators, researchers face challenges when measuring entrepreneurial growth in these two ways. First, small and young businesses – including innovative ones – are less likely to grow and are always much more vulnerable to fail (Aldrich and Auster, 1986; Stinchcombe, 1965). As a result, any tangible growth impact resulting from policy interventions may not be tracked prior to failure. Second, regardless of the type of business, accurately measuring entrepreneurial growth is difficult because this outcome involves different kinds of activities, occurring at different rates, and passing through different stages (Coad et al., 2007; Delmar et al., 2003; McKelvie and Wiklund, 2010). To circumvent these issues, we focus on tracking *entrepreneurial growth milestones* – or the accomplishments of standard indicators representing different ways a business expands its economic activity (Kim et al., 2015). By tracking growth milestones, we can measure growth using discrete outcomes in ways that mirror how support programs set goals for participants and monitor their progress against individual benchmarks related to the specific interventions (Autio and Rannikko, 2016).

2.1.2. Evaluating policy effectiveness

Given the resources devoted to entrepreneurship-support programs, it is natural to ask whether these interventions ("inputs") are effective at producing growth milestones ("outputs") in the innovative ventures they support. Up to now, evaluations of advice-giving program effectiveness have reached different conclusions about this basic question (Amezcua et al., 2013; Autio and Rannikko, 2016; Chrisman et al., 2012; Widersted and Månsson, 2015). We offer several reasons for these mixed and inconsistent findings. First, it is difficult to predict entrepreneurial growth a priori, since growth itself is a complex outcome with a variety of inputs that contribute to it (Davidsson et al., 2006). Second, growth implies a pattern of sustained productivity, but most businesses are unable to maintain this pace consistently over the long-term (Autio and Rannikko, 2016; Coad et al., 2007). Third, program design, size, and participants differ considerably, which makes program impact difficult to compare (Colombo et al., 2016; Heckman et al., 1997, 1999).

Most importantly, we argue that policy effectiveness depends on how programs select and treat their participants.[1] We define *treatment* as the actual support intervention received by the program participants intended to promote growth milestones in their ventures. We define *selection* as the mechanism by which participants are either included or excluded from receiving the intervention. Because programs involve different selection processes, treatment assignment is usually not random, which can impact the treatment effect on the outcome (e.g., Heckman et al., 1997; Rosenbaum, 2002). To determine the causal effect, researchers must address the selection biases that threaten the validity of a study. These selection biases vary and have multiple origins depending on program design (Maddala, 1983; Shadish et al., 2002). For example, a program may select participants or allocate support using different selection criteria or select ventures in different stages (Eckhardt et al., 2006). At the same time, program effectiveness also depends on how participants translate the treatment they receive into tangible growth outcomes. Programs can offer different kinds of support (such as training, referrals, or advising) and participants may benefit from some options more than others (Amezcua et al., 2013; Armanios et al., 2017). We argue that by incorporating both selection and treatment considerations into the empirical design, it is possible to understand more precisely how advising-based entrepreneurship-support programs can assist innovative ventures to achieve more growth milestones than other ventures. In the following sections, we examine the selection and treatment mechanisms that need to be addressed. Although we focus on innovative firms, we argue that these mechanisms generalize to all participants and describe these issues in this manner.

2.2. Selection mechanisms

2.2.1. Selection criteria

One form of selection bias occurs when using criteria to differentiate how to allocate support among participants in the program. In principle, public programs select firms targeted by the policy objectives and then provide them with support, such as personalized advice or codified lessons, to help ventures grow. In practice, programs struggle to design programs to implement these simple principles for achieving growth. Given limited resources, support programs cannot accommodate all deserving participants equally and must rely on various selection criteria to determine how to allocate the support. Typically, these criteria are assumed to predict which participants are most likely to translate the support into achieving growth milestones. They are also officially stated in program guidelines so the criteria can be used to "pick winners" from an initial screening (Shane, 2009). One common approach for this is to assess future potential based on past performance (Lerner, 2002). In other words, picking winners relies on characteristics already affirmed by the market and this guidance is expected to be a reliable signal for future growth potential. Since no single indicator leads to consistent growth outcomes, relying on past performance still remains a reasonably cost-effective method for selecting potentially innovative and high-growth businesses to support and to recruit a baseline pool of high performing "winners" that would likely grow even without the additional support. These criteria are practical, easy to implement, and unambiguous. This approach can be a low-risk, cost-effective way to spend public resources.

Several performance indicators are commonly used for selecting high-growth participants. These include a preference for larger businesses, in terms of sales, employees, or assets, since these ventures already have a proven track record (Birch et al., 1997; Chabaud and Messeghem, 2014; Wren and Storey, 2002). Another set of criteria involves innovative or younger high-growth ventures ("gazelles") since these businesses have the best potential for continued growth and broader economic impact (Colombelli et al., 2013; Coad et al., 2014). Gazelles with innovative products or services are particularly attractive since they not only contribute revenue and jobs, they can also usher in new technologies to the marketplace (Colombelli, 2016; Colombo et al., 2016). Besides these performance indicators, scholars have identified other selection criteria including: growth intentions (as opposed to actual growth itself – Wennberg et al., 2016; Wiklund et al., 2003), opportunity-based ventures (as opposed to necessity-based – Acs, 2006), received an initial round of external funding or completed other external screening (Shane, 2009), or businesses in high-growth sectors (such as technology or manufacturing – Mason and

[1] When we refer to "program," we mean "entrepreneurship support program."

Brown, 2013). Since many different selection criteria are used, it is not surprising to find high variation in treatment effectiveness.

It is important to recognize that whatever selection criteria is established at the outset increases the likelihood of allocating support in a corresponding way. As such, this selection bias needs to be accounted for to isolate the causal effect (Shadish et al., 2002: 54–55). We define *formal criteria* as priority characteristics that qualify participants into the program and serve as legitimizing devices used to assign treatment in a way that can be validated by external stakeholders (Armanios et al., 2017; Colombelli et al., 2019; Rotger et al., 2012; Söderblom et al., 2015). Once formalized, these criteria become the official standards by which ventures are evaluated and selected, and determine the milestones to be tracked (Autio and Rannikko, 2016; Kim et al., 2019). When resources are limited, ventures with certain criteria, such as being innovative, have a higher likelihood of being granted sufficient support. From a policy perspective, these formal criteria represent the preferences policymakers wish to prioritize for allocating limited support. They can design formal criteria to match local or regional needs to spur innovation and guard against simply defaulting to idiosyncratic preferences such as an advisor's personal expertise. Using these criteria also increases fairness in the selection process since programs abide by eligibility rules (Perrow, 1986). By contrast, ventures with characteristics not linked to any formal criteria will not receive as much consideration.

Depending on how the selection criteria are determined, programs still run the risk of not selecting the right participants because it is especially difficult to evaluate ex ante the actual quality of high-potential ventures (Huang and Pearce, 2015; Nightingale and Coad, 2014). Innovative ventures are heterogeneous and vary in quality (Mustar et al., 2006). Programs cannot be fully aware of all these differences when selecting, which amplifies the likelihood that the selected pool may not completely match the program's original intentions. Programs may not sufficiently screen out low-potential participants (Type I errors) because they still match on the formal criteria or risk overlooking high-potential participants (Type II errors) if they lack the official qualifications.

2.2.2. *Self-selection*

In addition to using formal criteria, selection can also occur at different stages of the program recruitment process. While the selection criteria are intended for the program to differentiate based on certain characteristics, selection that occurs at different stages results from participants who self-select out of the program. The reasons for self-selection include practical considerations, lack of commitment or unwillingness to follow through on the program requirements, over-confidence and reluctance to implement the advice they received, or a mismatch between what the program can offer and what the participant desires or needs to achieve their growth objectives (Bertoni et al., 2019; Brixy et al., 2013). For instance, innovative ventures may withdraw if they are not satisfied with the program's quality of sector expertise.

Self-selection produces a selection bias, so this needs to be adequately addressed in any program evaluation. Since program participation is typically voluntary, self-selection at different stages of program involvement is driven by whether participants apply or desire to remain in the program. Highly qualified ventures may not even apply because of a lack of awareness or program visibility, which can affect the applicant pool's quality and increase the risk of Type II errors of overlooking qualified applicants. When participants drop out prematurely, they waste the limited resources allocated to them. Participants may not fully realize their willingness or aptitude for receiving support until the initial meeting. What makes this form of self-selection challenging to manage is that this level of commitment cannot be accurately assessed by programs during selection. As a result, self-selection is difficult to screen out even through formal criteria and complicates program evaluation.

2.3. *Treatment mechanisms*

In addition to selection mechanisms, understanding how treatment mechanisms operate is the other half of analyzing effectiveness and designing better entrepreneurship support programs. Program effectiveness depends in part on whether treatment matches the participant's needs and whether the participants are prepared to translate the given advice delivered through these mechanisms into tangible growth outcomes throughout the process. While treatment designs vary across programs, most depend on some form of learning to transfer knowledge to participants (Bulte et al., 2016; Chrisman et al., 2005; Rotger et al., 2012). We also argue that programs need to determine how best to allocate the support (typically in terms of advising hours) across all qualified participants. We reiterate the common challenge faced by programs when designing their treatment options: that although selection criteria, such as "picking winners," are intended to recruit high-potential participants, they still do not guarantee that all participants will benefit from treatment effectively. And because of limited resources, allocating more support to some participants comes at the expense of others.

2.3.1. *Types of learning*

Consistent with classical organizational learning theory, we define *learning* as the integration of an advisor's knowledge into the venture's routines and behaviors (Levitt and March 1988). Innovative ventures depend on various knowledge sources and vary in the ways they learn (Agarwal and Shah, 2014; Clarysse et al., 2009). In particular, innovative ventures focus on knowledge exploration and exploitation, which require a deeper commitment to internalize the expert advice they receive and apply it toward their growth efforts (Colombelli et al., 2013; Colombo et al., 2015). Entrepreneurs cannot tackle all of their growth imperatives on their own, since these efforts involve legal, financial, operational, and other kinds of expertise they may lack (Lerner, 2010). In some situations, entrepreneurs simply want a solution to a specific problem they have. Some may prefer to learn on their own through codified means (e.g. books, lectures, and other training programs), expect to find their answers in established sources, or ask specific questions to advisors.

In other situations, entrepreneurs know they need assistance but may not completely understand the key issues at hand. This may favor a relational experience with advisors and experts (Chrisman et al., 2005). As such, they are more willing to socialize their problem with their advisors to help them define it. Because entrepreneurship is a collaborative endeavor, those who devote effort to seeking and implementing high quality advice can increase their likelihood of sustaining growth (Kim et al., 2013; Ruef, 2010). When participants are genuinely seeking advice to tackle their growth challenges, they can advocate directly for what kinds of outcomes they seek. What matters most about this distinction is a participant's willingness to engage with the advisor to build a collaborative learning relationship around their growth problems. Experienced advisors can discern these different learning patterns early on in their client interactions and use them to develop productive advising pathways.

2.3.2. *Learning and growth*

In principle, learning is straightforward treatment mechanism; however, the findings regarding its impact on entrepreneurial growth are mixed. While some scholars have found benefits attributed to learning on firm growth (Autio et al., 2008; Autio and Rannikko, 2016; Clarysse et al., 2009), others reported no or limited effects on performance (Amezcua et al., 2013; Armanios et al., 2017). To gain clarity on this issue, we raise the possibility that learning effectiveness rests on treatment length. Ventures typically do not implement all at once the new knowledge they acquire, so the effect of treatment on growth typically follows a curvilinear U-shaped learning curve (Musaji et al., 2020; Toft-Kehler et al., 2014). This means that significant translation of new knowledge occurs further along the learning curve as additional treatment is delivered. Given the exploratory nature of this approach, learning takes time to be effective (Rotger et al., 2012), since it is intrinsically a social process between the advisor and client (Cohen et al., 2019; Mason and Brown, 2013). This type of learning also matches with the multiple stages of creating, testing, and introducing innovative products, services, and business models. Innovation does not happen overnight but requires in-depth discussions and ongoing explorations by both parties to conceive and experiment possible solutions.

The challenge for programs is how to assess this willingness to learn collaboratively early and accurately when treatment begins, in order to allocate the right amount of resources to these participants. In this model, participants need to work collaboratively over months with experienced advisors, who are often successful business owners themselves and provide sector-specific technical or managerial assistance. This forces programs to decide how best to allocate their limited advising hours to reach as many participants as possible and to ensure enough learning translates into growth. Trying to accomplish both will result in subpar outcomes. Both constraints require programs to tackle difficult tradeoffs early in the treatment process about how best to allocate their advising hours. Over allocating support to unprepared participants (Type I error) will take away limited resources from willing participants who wish to learn more through additional advising. Programs must also confront the complication of not having enough information about participants to make this allocation accurately. This may require them to depend on their first impressions in their initial meetings to anticipate who are genuinely willing to learn and progress furthest along the learning curve to achieve growth.

2.4. Summary

We have outlined the rationale for boosting innovative ventures and their entrepreneurial growth should be integrated into a comprehensive empirical design to evaluate program effectiveness for boosting entrepreneurial growth and innovative ventures. For each selection and treatment mechanism, we highlighted the tradeoffs and challenges faced by programs that seek to employ them effectively given limited resources. In Fig. 1, we summarize these different tradeoffs to visually represent the breadth and depth of these issues involved with program design and evaluating effectiveness. As this figure shows, there are multiple interdependent facets to designing, executing, and evaluating program effectiveness with compounding tradeoffs and challenges throughout. To analyze these issues comprehensively, we require a carefully designed empirical strategy to determine whether program treatment leads to achieving growth milestones and properly accounts for various observed and unobserved selection issues. In the following section, we describe these details.

3. Study context

To tackle our research questions, we desired an entrepreneurial-support program context that operated with the selection and treatment mechanisms we highlighted to qualify growth-oriented innovative ventures into the program for treatment. Specifically, we sought a context to test the "picking winners" approach that depends on formal criteria to select applicants – a simple but operationally complex design. To disentangle its competing mechanisms, we also desired a setting with fine-grained information on how the participants were selected, how they worked with advisors, and the types of advice they received. To track growth milestones, we also needed data on how the businesses performed after working with their advisors.

We analyzed data from a Small Business Development Center program (SBDC). This program started in 1979 as a part of the U.S. Small Business Administration. Today, it has grown to over 1000 service centers advising over 1.2 million businesses annually, becoming the main source of technical and managerial assistance for small businesses in the United States. The stated purpose of these centers is to help entrepreneurial ventures grow in terms of revenue, employees, or capital. The SBDC program has been successful, generating over $6.8 billion in sales and over 75,000 full-time equivalent jobs for their long-term clients (Rowe, 2016).

We focused our investigation on the SBDC of Ventura and Santa Barbara Counties in California, located north of Los Angeles. This SBDC was created in 2010 and has been repeatedly ranked among the most efficient centers in California by helping clients to generate over 3000 jobs and more than $250 M in sales. We studied this center because it fulfilled our study requirements of having fine-grained data on selection and treatment protocols, participants, advisors, and their resulting outcomes – similar to other research designs using SBDC data to analyze performance outcomes (Chrisman and McMullan, 2004; Chrisman et al., 2012). The Center advises a wide range of businesses, as long as they are located in the Ventura, Santa Barbara, or Los Angeles counties, have less than 500 employees, and are for profit. With broad eligibility rules, this program must allocate its limited resources to clients who are anticipated to benefit the most from the support.

	How?	Who?	Issues and tradeoffs
Selection	1. How to allocate support among participants? 2. How to select the right participants?	3. Which participants are most likely to translate the support into achieving growth milestones? 4. Which participants self-select themselves out of the program, either before or during the program?	1. Given limited resources, support programs cannot accommodate all deserving participants equally. 2. It is especially difficult to evaluate ex ante the actual quality of high-potential ventures. Formal criteria amplify type I errors, whereas informal criteria increase type II errors. 3. Formal criteria based on past performance to assess future potential, to allocate support, and to increase fairness yet no single indicator leads to consistent growth outcomes. 4. The level of commitment of the participants cannot be accurately assessed by programs during selection.
Treatment	5. How to best allocate the support across all qualified participants? 6. How to assess the willingness to learn collaboratively early and accurately when treatment begins in order to allocate the right amount of resources to these participants?	7. Does treatment match the participant's needs and are the participants prepared to translate the given advice into tangible growth outcomes throughout the process?	5. Although selection criteria such as "picking winners" are intended to recruit high-potential participants, they still do not guarantee that all participants will benefit from treatment effectively. And because of limited resources, allocating more support to some participants comes at the expense of others. 6. Over allocating support to unprepared participants (Type I error) will take away limited resources from more willing participants who wish to learn more through additional advising. Programs must also confront the complication of not having enough information about participants to make this allocation accurately. 7. Innovative ventures depend on various knowledge sources, vary in the ways they learn and have different needs.

Fig. 1. Program selection and treatment challenges and tradeoffs.

3.1. General SBDC sample information

In our study, we analyzed 1,712 businesses that enrolled for the first time with the SBDC between January 1st, 2011 and December 31st, 2015 (five years).[2] Overall, the data quality is high, with few missing values and minimal errors (since the clients were required to provide the information to receive services, and their advisors were expected to maintain accurate records to be paid by the SBDC). We followed the client ventures from their first consultations and tracked their growth milestones until October 5th, 2017. Zooming in on short-term growth achievements for new clients enabled us to assess the impact of their advisory effort on growth. At the end of the observation period, the ventures had been tracked for at least 21 months after enrollment and up to a maximum of 81 months (depending on their length of participation). In our analysis, we accounted for this advising time difference and possible right truncation.

During the study period, the SBDC delivered 8,487 sessions with clients, for a total of 16,488 h of advising work (consisting of 10,930 h of face-to-face contact time and 3005 h of preparation time; the remainder is virtual contact time). This averaged between 8 and 12 h per client (depending on the year).

3.2. SBDC selection process

We describe the SBDC selection process for allocating support to its clients before and during the program (visually summarized in Fig. 2, adapted from Maddala, 1983: 266). This process occurs across multiple stages and criteria. There are four different stages of selection, two related to self-selection by the participants (*application* and *participation*) and two related to the program for allocating support (*evaluation* and *activation*). Sequentially, the selection process starts with the application stage and proceeds to the evaluation, activation, and participation stages.

In the first stage (*application*), interested businesses self-select into the process of receiving assistance by applying to the SBDC. During our observation period, 1,762 businesses applied out of a population of approximately 925,000 eligible ventures. To explore the similarity of SBDC applicants with non-applicants, we conducted a difference in difference (DID) matching analysis on a separate, independent data source: Your-economy Time Series (YTS) from the Business Dynamics Research Consortium (http://bdrc.uwex.edu/our-databases.iegc). This dataset includes records for over four million businesses in California. We extracted information on businesses operating in the three-county jurisdiction between 2011 and 2017 and appearing in the dataset for at least one year. We further narrowed the data to those in the same industry, founded the same year, operating in the same county within a 300 km driving distance to the SBDC and reporting similar pre-treatment growth trends (between years N-2 and N-1), resulting in approximately 90,000 businesses. From this sample, we located 488 SBDC clients and formed an independent control group of similar businesses. Full details of this analysis appear in Appendix 1.

Once a business applies for SBDC support, it moves into the second stage (*evaluation*). During this stage, SBDC management evaluates whether and how much to deploy their advising hours for each applicant. This is necessary because the SBDC faces budget constraints that prevent unlimited advising to all applicants. For this evaluation, each applicant completes an initial Scope of Work questionnaire to estimate the type and amount of work required. The management team evaluates the responses with a scoring tool, which allocates an initial amount of advising hours based on four formal criteria: years in business, gross revenue, operating in the manufacturing industry, and whether the applicant was referred to the SBDC by its program partners. This scoring tool does not specifically prioritize innovativeness, so it allows us to assess to what extent these particular ventures receive support and how it translates into achieving growth. Based on the evaluation scoring tool, 1,712 business were assigned initial advising hours, while approximately 50 did not meet the SBDC eligibility criteria.[3] At this point, the staff assigns an advisor and an area of work from a standard list of project objectives (e.g. developing a business plan or raising capital).

After the evaluation, clients meet with their appointed advisor for an initial meeting to begin the advising treatment. At this third stage (*activation*), the advisor assesses the project, defines a methodology, assigns treatment, and authorizes additional hours, especially for high-growth potential clients. The SBDC is regularly audited to ensure the advisors have the freedom to exercise their expertise as they see fit. As a result, 1,068 clients were activated for treatment. 644 clients were deactivated and removed from the program by the advisor after the first meeting.[4]

In the fourth and final stage (*participation*), most clients work with their advisors and receive assistance (up to their maximum allocations), while some clients do not commit to the advising program (e.g. Brixy et al., 2013) or drop out for unknown reasons (774 clients participated up to the end of initially allocated hours, while 294 dropped out before the end of the treatment). Given the multiple facets of program selection and treatment, we provide more information in the Appendix that support the details presented in the following sections.

4. Methods

4.1. Variables

4.1.1. Outcome variable – growth milestones

We define growth milestones as the achievement of particular events indicating growth in the business (Kim et al., 2015). In the main analysis, we used SBDC records as a data source to measure these milestones. We created three separate milestone categories – (1) capital investment, (2) job increase, or (3) sales increase. (1) A capital investment milestone occurs when the business raises external funding, either with outside debt or equity, for expansion purposes. (2) A job increase milestone occurs when the business hires someone, and (3) a sales increase milestone is recorded when the client experiences a significant increase in their sales. The smallest sales increase amount in our sample is $35,000. We measured growth in two ways: with a dichotomous variable (1 = yes) when a firm achieved any milestone and a count variable of the number of milestones achieved: a firm may reach the same milestone multiple times (such as multiple capital increases over time) or reach different milestones (e.g., a capital increase and a recruitment). Table 1 shows the distribution of milestone events by category. 335 ventures reached at least one milestone, while the total number of milestones reached across all ventures is 754.

4.1.2. Selection variables

There are two kinds of selection variables related to the program design: (1) those related to the selection stages and (2) those related to the selection criteria (see Fig. 2). For the first selection stage

[2] Older clients that returned to the SBDC during the timeframe are not part of our sample. Five firms were dropped because of incomplete data. One firm was removed because no session notes were recorded (input error). Four additional firms were removed because the firms contained missing values for some of the relevant variables.

[3] These businesses were not accepted into the program and referred to other mentoring programs. Unfortunately, we have no records for these 50 businesses to conduct any additional analyses.

[4] Sometimes, the de-activation by the advisor occurs after the second meeting. In any case, it systematically occurs before receiving five hours of advising treatment (the threshold used in other SBDC studies - Chrisman et al., 2012).

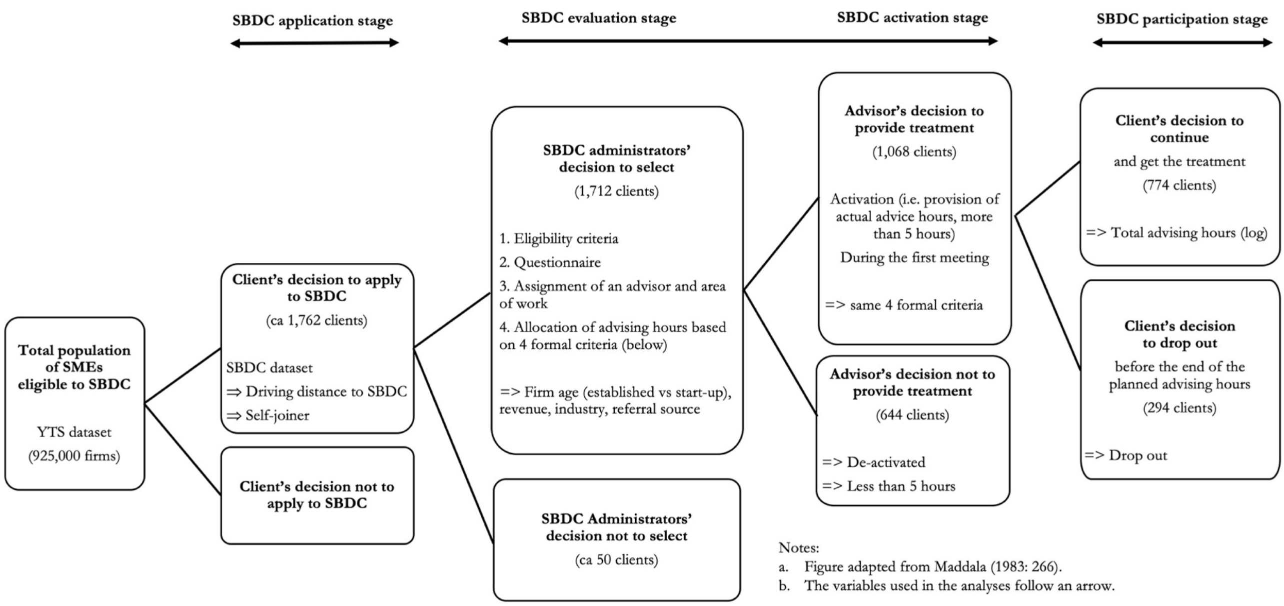

Fig. 2. SBDC selection process, stages and criteria.

Table 1
Distribution of growth milestone events among SBDC clients(N = 1712).

			N firms
No milestone has been achieved			**1377**
1 milestone has been achieved:			**156**
Capital increase	Job increase	Sales increase	
X			37
	X		46
		X	76
More than 1 milestone has been achieved:			**179**
Capital increase	Job increase	Sales increase	
X			10
	X		5
		X	22
X	X		18
	X	X	85
X		X	14
X	X	X	25

Notes: 335 (20%) firms achieved at least one milestone; among those, 156 firms achieved exactly 1 milestone, 179 firms achieved more than one milestone and 91 (5%) firms achieved more than 2 milestones; the maximum number of milestones achieved is 17; the total number of milestones achieved is 754. Among the firms that achieved multiple milestone, 10 achieved only capital increase (multiple times); 18 achieved capital increase and job increase but no sales increase; 25 achieved all the possible milestones.

(application), we used a dummy variable (1 = *applied to* SBDC) among the population of businesses in Ventura, Santa Barbara, and Los Angeles counties (described in Appendix 1). For the second selection stage (evaluation), we did not create a variable, given the lack of information: These approximately 50 ventures were denied a first meeting with an advisor because they did not meet the SBDC eligibility criteria for receiving support or their application file was incomplete.[5] They are thus grouped with the non-applicants and accounted for in the application stage analyses (see Appendix 1). For the third selection stage (activation), we noted when an advisor decided to collaborate with a client on an agreed-upon scope of work and to activate it for treatment (1 = *activated*). For the final selection stage (participation), we used a dummy variable for clients that left the program after being activated (1 = the *client dropped out* before the end of the initial allocation).

In addition to selection stage variables, we included the formal selection criteria used in the SBDC scoring tool. At the activation stage, four formal criteria are used to allocate advising hours: (1) *established firms* (1 = in business for more than one year) to reflect an emphasis on supporting established firms[6] (e.g., Coad et al., 2014; Lerner, 2010; Minniti, 2008; Shane, 2009). (2) *gross revenue* (in sales in USD declared by the client during the initial request for advice) to capture past performance (Birch et al., 1997; Wren and Storey, 2002). (3) *manufacturing industry* (1 = yes). Some scholars specifically focused on businesses in the manufacturing sector as potential high performers (Mason and Brown, 2013). (4) *referred client* (1 = yes) for ventures that joined the SBDC based on the recommendation of other program partners (such as the SBA program, the SCORE program, Women's Business Centers, and others), or a financial provider (such as the Ventura County Credit Union, Rabobank, Pacific West Bank, and others). A client could be more likely to grow if it has been screened by other program partners (Shane, 2009).

Besides these four selection criteria variables, we included an additional variable for *innovative* ventures (1 = yes). Although innovative ventures are not formally screened as a criterion in the scoring tool, we included it as a selection variable to determine how support was allocated to these ventures and compared this with the other four formal criteria for program effectiveness. To create this variable, we analyzed the text from the advisor's initial meeting notes with their clients for keywords consistent with definitions of innovative ventures (Aldrich and Ruef, 2018; Rubera and Kirca, 2012). By focusing on the first meeting notes, we ensured that an assessment of the venture's innovative qualities was taken prior to the actual treatment. Integrating

[5] The SBDC rejects a priori only firms that are not small businesses, not likely to become small businesses (for potential entrepreneurs), not in the administrative area of the SBDC, or submitted incomplete application file. Since they do not meet the eligibility criteria, the 50 firms rejected early in the process (out of ca. 1,750) are equivalent to non-applicants and, for this reason, we assigned them to this group.

[6] While firm age is an element of the SBDC scoring tool, established firms required a higher score to receive the same amount of advice. As such, startup firms were allocated more advice than recently established firms, even though they had a lower SBDC score, everything else made equal.

all the meetings would violate a key condition of causality (Holland, 1986) and unnecessarily inflate the likelihood one of these terms emerged in conversations between the advisors and clients. We assigned the value of 1 if at least one keyword appeared, using the innovation lexicon assembled by Godin (2014).[7] The results were verified by a seasoned SBDC administrator.

4.1.3. Treatment variables

We define *advising effort* as the number of hours spent with the client, including the advisor's preparation time (Chrisman et al., 2005; Chrisman and McMullan, 2004).[8] We constructed our continuous measure directly using each SBDC advisor's work logs, so this variable accurately reflects the time spent to advise the client. We also created dummy variables to account for different treatment levels used in the robustness checks. Following previous studies (Chrisman et al., 2005), we used the *five-hour threshold* to determine if a venture received the treatment. This threshold is also consistent with the practices of the SBDC, which attempts to provide a minimum of five hours of advice to all activated clients. As robustness checks, we also tested other thresholds in our analyses (e.g. 10 h, 15 h).

4.1.4. Additional variables

Besides our selection and treatment variables, we also controlled for alternative factors that might lead to either program selection, advising hours allocation, or growth milestone achievements. We distinguish three types of variables: (1) variables about the client, (2) variables about the main advisor and (3) variables about the initial meeting between the client and the advisor. We also provide extra information about the choice, creation, and rationale of these variables in Appendix 2.

Regarding the client, we divide the ones that are not referred to the program into two groups: *self-motivated to join* (1 = client joined on their own, while looking for an advising solution by themselves, either on the internet or through direct contact with the SBDC), and the other clients (0 = clients that decided to join the program after receiving targeted information, through universities, cities, forums, etc., for example) to reflect their personal interest in consulting SBDC to grow their business. We included other control variables in the different models: advising areas requested by the client (*increase in sales, capital and startup*); *female-owned firm* (1 = yes); location (1 = firm in *Santa Barbara*); different dichotomous variables to account for *industry* (using the categorization of the SBDC scoring tools).[9]

Regarding the main advisor, we created the following variables: advisor holds an *MBA or a Ph.D.* (1 = yes); the advisor's *number of prior business startups* and *years of management experience*. Prior studies have emphasized the role of education and experience in mentoring and counseling activities (Behncke et al., 2010; Heckman et al., 1999). Using higher education levels and the number of business startups was a way to identify the most highly educated and experienced advisors amongst the group.

Regarding the first meeting between the client and the main advisor, we included five variables based on prior research and our understanding of this particular SBDC context by analyzing written reports of the initial client meetings. We coded for *collaborative learning* (1 = client-advisor meeting involved open-ended discussions, guidance, and feedback about their growth objectives). We also formed a binary variable if the advice was related to *bridging* to others (1 = advisor recommends different people or resources to contact by providing their contact information). To code these last two variables, we again used the advisor's initial meeting notes since it recorded details about the client's willingness to learn new approaches, to what extent the advisor anticipated the collaboration to continue beyond the initial meeting, and any referrals made to assist the client. In Appendix 2, we provide coded meeting notes examples, coding procedures (including more fine-grained categories), and inter-rater reliability details.

Besides policy-based formal selection criteria, the SBDC advisors may also rely on experience or use heuristics to make decisions to form positive or negative perceptions of the client, which has been studied as "gut feel" in the literature (Huang, 2018; Huang and Pearce, 2015). We measured this advisor sentiment as expressed with *positive* or *negative* words (as a percentage of total words) (Tetlock, 2005). For this calculation, we used Loughran and McDonald's (2013) financial and business dictionary in their study of venture capitalists assessing IPOs.

4.2. Preliminary analysis and results

Table 1 and Table 2 present descriptive statistics, and Table 3 shows the pairwise correlations of the variables in our analyses. Consistent with prior research, growth is a rare event in our sample: only 20 percent of ventures reached a milestone, 10 percent achieved more than one, and five percent more than two. After applying the scoring model, only 45% received more than five hours of advice (including advisor's preparation time), which is the minimum threshold for treatment (Chrisman et al., 2005, 2012).

Despite not being prioritized in the SBDC scoring model, innovative ventures are well represented (33% of the total ventures in the sample). This ratio is similar to other businesses in the SBDC scoring model (the formal criteria): ventures with revenues (43%), established firms (41%), clients referred by program partners (27%), and manufacturing ventures (14%).

Looking at the correlation between our covariates (Table 3), we do not observe any major correlation issues. In general, most coefficients lie between -0.1 and 0.1. Industry groups are correlated (the groups do not overlap). Similarly, established firms are less likely to request advice in starting up. Finally, the education and experience of the counselors are highly correlated ($\rho = 0.71$). We decided not to include management experience in the reported models. As a check, we also reran the models with both variables and obtained stable results. Apart from these two variables, the moderate correlations are unlikely to cause any issues in the different models: the variance inflation factors remain around or below 2 for all covariates.

As a first step to determine whether selection bias exists in allocating hours, we ran an ordinary least squares (OLS) regression to estimate the treatment allocation (Table 2).[10] We used a log transformation of the treatment (advising hours) to approach normality (the non-logged hours are skewed to the left, since most ventures receive little treatment). From this model, we observe that innovative ventures, ventures with higher gross revenue, manufacturing ventures and startup ventures (as opposed to established firms) have a higher likelihood to receive a larger treatment dose.[11] Also, the client's growth orientation (i.e., advising area = increase in sales) and the client-advisor

[7] The list of keywords used is: *innovat, trademark, patent, R&D, prototyp, technolog, new prod, new service, new business model, new process, new market, new brand, novel, creativ, original, modern, diversif, chang, reorganiz, reform, introduc, diffus, invent, disrupt, revolut, radical, design, experiment.*

[8] We also ran our models without advisor's preparation time and obtained consistent findings.

[9] Industries included: arts; education; food services; professional, technical, research and development activities; retail, and other services. For reasons specified in the analyses, we also grouped industries in four categories (types) used in the SBDC scoring tool for hours allocation.

[10] Given the skewed distribution of treatment hours, we also conducted a quantile regression and report these results in Appendix 3.

[11] Contrary to the four criteria used by SBDC to decide hours allocation, the result we have in this model for innovative firms is unstable in the complementary models we report in Appendix 3. One explanation could be that some promising profiles of innovative firms (manufacturing firms, high revenue prior to joining the program, etc.) would receive even more hours of treatment, but this is not the case for all innovative firms.

Table 2
Descriptive Statistics and selection for treatment (OLS).
All SBDC Clients (N = 1712).

Variables	Mean	S.D.	Min	Max	OLS (2) DV: Advising effort	
Outcome - growth milestone:						
Achieved any milestone (binary)	0.196	0.397	0	1		
Number of milestones achieved (count)	0.440	1.296	0	17		
Treatment:						
Advising effort (total, in ln hours)	1.559	1.083	−1.386	6.066		
Five-hours threshold (binary treatment)	0.452	0.498	0	1		
Selection stages:						
Client activated	0.624	0.485	0	1		
Client dropped out	0.172	0.378	0	1		
Selection – formal criteria and innovativeness:						
Established firm	0.405	0.491	0	1	−0.126*	(0.058)
Gross revenue (ln)	−1.642	11.567	−11.513	17.622	0.086***	(0.010)
Firm industry: manufacturing (type 4)	0.143	0.350	0	1	0.353***	(0.087)
Referred firm	0.270	0.444	0	1	0.069	(0.058)
Innovative firm	0.326	0.469	0	1	0.143**	(0.052)
Variables about the client:						
Self-motivated to join	0.237	0.425	0	1	−0.118*	(0.060)
First advising area: increase in sales	0.152	0.360	0	1	0.263***	(0.073)
First advising area: capital	0.127	0.333	0	1	0.083	(0.077)
First advising area: start-up	0.251	0.433	0	1	−0.041	(0.064)
Female owned firm	0.276	0.447	0	1	0.147**	(0.054)
Firm in Santa Barbara	0.269	0.444	0	1	0.008	(0.058)
Firm industry: service (type 2) (1)	0.448	0.500	0	1	0.096	(0.067)
Firm industry: retail and arts (type 3)	0.189	0.391	0	1	0.195*	(0.079)
Variables about the main advisor:						
Has an MBA or a Ph.D.	0.457	0.498	0	1	−0.335***	(0.050)
Number of prior business startups	1.605	0.812	0	4	0.050	(0.032)
Years of management experience	23.362	14.433	0	40		
Variables about the initial meeting client-advisor:						
Positive gut feel	0.007	0.009	0.000	0.071	8.388**	(2.750)
Negative gut feel	0.010	0.010	0.000	0.091	−1.355	(2.494)
Collaborative learning	0.546	0.498	0	1	0.411***	(0.049)
Bridging	0.223	0.416	0	1	−0.079	(0.058)
Instrumental variable:						
Driving distance to SBDC (> 2 h)	0.038	0.191	0	1	−0.367**	(0.129)

Notes: (1) Industry types correspond to the industry groups used by SBDC in their scoring template to allocate an initial number of hours to the client. Type 2 = all service industries and professional and technical activities; Type 3 = retail and arts industries; Type 4 = manufacturing; Type 1 = no industry (i.e. for startups for instance).
(2) *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; Two-tailed test; Standard errors in parentheses; OLS model includes an intercept (1.018***); $R^2 = 0.194$.

interaction during their first meeting (i.e., positive gut feel, learning behaviors) are strong predictors of treatment dose. Overall, these patterns show that the advising hours are not equally shared among ventures and the assignment is nonrandom, which confirms that selection bias should be corrected in order to properly evaluate the causal effect of treatment on achieving growth milestones. This requires us to conduct additional procedures to distinguish whether ventures perform better because of treatment or because of their own characteristics.

4.3. Identification strategy

To understand whether innovative ventures benefited from the program, we first needed to assess whether the program is effective through an appropriate identification strategy. Following established principles in the program evaluation literature (Angrist and Pischke, 2009; Imbens and Wooldridge, 2009), we estimated the causal effect of advice treatment on achieving growth milestones. For this, we based our analysis on the potential outcomes framework (Rubin, 1974). Since the SBDC treatment is operationalized with a continuous indicator, we specified a continuous potential outcome. Following Hirano and Imbens (2004), we considered the potential outcome as an interval, which represents a "dose" of advice a client received. As a robustness check, we used binary treatment indicators (1 = activated, 5, and 10 h) to estimate the average treatment effect (ATT) on the treated clients (Appendix 4).

Since we do not have an experimental design typically associated with the potential outcomes framework, we implemented several steps to conduct a quasi-experimental design. To be interpreted as plausibly causal, the treatment must satisfy some form of exogeneity, an assumption known as the conditional independence assumption (CIA), or unconfoundedness (Lechner, 1999; Rosenbaum and Rubin, 1983b). In observational studies like ours, satisfying this assumption required us to construct control groups identical along both observable and unobservable characteristics before the treatment is given.[12] However, in our setting, we have a multifaceted selection process, with several stages and criteria that posed multiple threats to the internal validity of our quasi-experimental design (Shadish et al., 2002). This process created differences between the treated and control groups (see Table 2 in the result section) and would violate the CIA assumption. In the following section, we present the results from the three-stage estimation procedures we implemented to address this issue. We provide details about the estimation procedures at each step.

5. Results

We deployed three estimation procedures to implement our identification strategy and address different validity threats: (1) matching

[12] Another assumption called SUTVA is necessary for this causal interpretation (Holland, 1986). SUTVA stands for stable unit treatment value assumption, which means that the client's potential outcomes do not depend on the potential outcomes of other clients in the population.

Table 3
Pairwise correlations.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13
Outcome - growth milestone:													
1. Achieved any milestone	1												
2. # milestones achieved	.69	1											
Treatment:													
3. Advising effort (ln)	.55	.53	1										
4. Five-hours threshold	.45	.34	.82	1									
Selection stages:													
5. Client activated	.36	.26	.67	.71	1								
6. Client dropped out	−0.09	−0.09	−0.15	−0.18	.35	1							
Selection criteria:													
7. Established firm	.07	.08	.11	.08	.09	.05	1						
8. Gross revenue (ln)	.31	.27	.29	.27	.25	.04	.42	1					
9. Industry: manufacturing	.08	.10	.13	.10	.12	.12	.17	.10	1				
10. Referred firm	.07	.08	.08	.07	.06	−0.05	.09	.07	.07	1			
11. Innovative firm	.06	.04	.11	.08	.08	−0.01	.05	.04	.11	.05	1		
About the client:													
12. Self-motivated to join	−0.04	−0.05	−0.09	−0.09	−0.08	.04	−0.01	−0.05	−0.05	−0.34	−0.04	1	
13. Advising area: sales increase	.10	.13	.14	.11	.12	.01	.17	.16	−0.02	−0.01	.03	.00	1
14. Advising area: capital	.04	.03	.05	.04	.05	.04	.16	.13	.06	.15	.00	−0.04	−0.16
15. First advising area: start-up	−0.04	−0.07	−0.09	−0.06	−0.08	−0.07	−0.40	−0.23	−0.08	−0.04	−0.02	−0.01	−0.25
16. Female owned firm	.07	.04	.05	.02	.02	−0.05	.05	.01	−0.05	−0.05	.03	.02	.07
17. Firm in Santa Barbara	−0.04	−0.08	−0.02	.00	−0.02	−0.03	−0.18	−0.13	−0.04	−0.05	.05	−0.08	−0.03
18. Industry: type 2	−0.01	−0.02	−0.05	−0.03	−0.01	−0.02	.06	.05	−0.39	−0.01	−0.14	.00	.07
19. Industry: type 3	.04	.01	.04	.03	.00	−0.07	−0.02	.02	−0.20	.02	.11	−0.01	.01
About the main advisor:													
21. Has an MBA or a Ph.D.	−0.13	−0.06	−0.21	−0.23	−0.24	.00	−0.05	−0.07	−0.13	−0.07	.02	.08	−0.02
22. # of prior business startups	−0.10	−0.10	−0.02	−0.06	−0.06	−0.01	−0.06	−0.10	.02	.01	.04	.01	−0.23
23. Years of managt. experience	−0.13	−0.10	−0.08	−0.16	−0.18	−0.01	−0.02	−0.10	−0.01	.04	.04	.03	−0.24
About the initial meeting:													
24. Positive gut feel	.10	.07	.14	.11	.11	−0.01	.15	.10	.08	−0.01	.09	−0.01	.14
25. Negative gut feel	.01	.03	.01	.05	.07	.07	.02	.08	.04	.03	−0.09	−0.08	−0.03
26. Collaborative learning	.14	.11	.23	.22	.20	−0.01	−0.01	.07	.01	.01	.09	−0.03	.11
27. Bridging	−0.03	.03	−0.03	−0.05	−0.04	.01	.00	−0.04	.11	−0.04	.06	.00	−0.02
Instrumental variable:													
28. Driving distance (> 2 h)	−0.05	−0.04	−0.09	−0.08	−0.07	.01	−0.03	−0.05	.01	.03	−0.01	−0.02	−0.02

Variables	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Outcome - growth milestone:														
1. Achieved any milestone														
2. # milestones achieved														
Treatment:														
3. Advising effort (ln)														
4. Five-hours threshold														
Selection stages:														
5. Client activated														
6. Client dropped out														
Selection criteria:														
7. Established firm														
8. Gross revenue (ln)														
9. Industry: manufacturing														
10. Referred firm														
11. Innovative firm														

(*continued on next page*)

Table 3 (*continued*)

Variables	14	15	16	17	18	19	20	21	22	23	24	25	26	27
About the client:														
12. Self-motivated to join														
13. Advising area: sales increase														
14. Advising area: capital	1													
15. First advising area: start-up	−0.22	1												
16. Female owned firm	−0.06	−0.06	1											
17. Firm in Santa Barbara	−0.14	.05	−0.02	1										
18. Industry: type 2	.04	−0.15	.05	−0.02	1									
19. Industry: type 3	−0.03	.03	.10	.02	−0.46	1								
About the main advisor:														
21. Has an MBA or a Ph.D.	−0.06	.01	.05	.06	.08	.03	1							
22. # of prior business startups	.06	.02	−0.02	.19	−0.06	.02	.14	1						
23. Years of managt. experience	.10	.01	.00	.00	−0.01	−0.01	.48	.71	1					
About the initial meeting:														
24. Positive gut feel	−0.07	−0.07	.02	−0.02	.00	−0.01	−0.07	−0.01	−0.06	1				
25. Negative gut feel	.06	−0.01	−0.05	−0.04	.02	.00	−0.05	−0.03	−0.06	−0.13	1			
26. Collaborative learning	−0.06	.07	.01	.10	−0.05	.03	−0.09	−0.07	−0.12	.09	.02	1		
27. Bridging	−0.02	−0.01	.04	.03	−0.04	−0.05	.08	−0.01	.08	.02	−0.06	.06	1	
Instrumental variable:														
28. Driving distance (> 2 h)	−0.04	.03	−0.02	.19	−0.01	−0.05	.07	−0.11	−0.14	−0.02	−0.04	.00	−0.03	1

Notes: $N = 1712$; All correlations above $|0.048|$ are significant at $p < 0.05$.

analysis using generalized propensity scores (to address selection on observable criteria when allocating hours in the middle activation stage), (2) a Heckman-type corrected regression (to address unobserved self-selection biases when treatment is initially allocated in the evaluation and delivered in the final participation stages), and (3) the difference-in-difference analysis using the YTS dataset, an independent data source, to address omitted, or unobserved, time-varying and -invariant factors that might affect assignment to treatment.[13] The purpose for combining these procedures is to assess whether innovative ventures benefited from this program, net of the various selection biases we reviewed. By using multiple procedures, we disentangled competing mechanisms driving selection biases or treatment effects.

5.1. Generalized propensity score

We first conducted a matching analysis to address the selection bias generated from using formal observable criteria ("picking winners") to assign hours of treatment. Matching is an appropriate quasi-experimental method to address this bias because it balances the distribution of observable criteria between treated and non-treated participants in all the relevant pre-treatment characteristics (Caliendo and Kopeinig, 2008; Dehejia and Wahba, 1999; Rosenbaum and Rubin, 1983a). Since our treatment variable (advising effort) is a continuous measure, we employed the generalized propensity score (GPS) technique developed by Hirano and Imbens (2004). The GPS technique accommodates different doses (or amounts) of treatment while retaining the quasi-experimental properties of the matching methods. Compared to binary treatment matching estimators, continuous treatment analyses have the advantage of investigating different levels of support (Imbens and Wooldridge, 2009). This feature is well-suited to evaluate the effectiveness of distributing advising hours among participants who have been prioritized by the scoring model's formal criteria.

Similar to binary matching, the GPS technique accomplishes the following: (1) estimates and corrects for selection bias (balancing properties) and (2) estimates the dose response function for comparable observations. GPS aims to improve the balancing properties between the different treatment doses. The balancing properties can be satisfied only for observations with common support (i.e. similar ventures that received non-treatment and different treatment doses). Ventures that do not satisfy the common support requirement (i.e., dissimilar to all other ventures in the sample) are excluded from the analysis. We sought common support on all variables to be used for the GPS: the four scoring model criteria, whether the venture was innovative, client characteristics, the main advisor's characteristics, and the client-advisor interaction variables.

The next step is to confirm that the groups of firms that will receive the different dose of treatment were comparable. Since the matching method is sensitive to outliers, we excluded extreme values for our highly skewed variables: 2 ventures with more than 10 milestones (median is 0 milestones), 16 ventures with more than 100 h of treatment (median is 4.25 h), and 46 ventures with annual revenue over $3 million before receiving SBDC support (median is $0).[14] We retained 1540 observations that met the requirement of common support.[15] After removing the observations, we reran the OLS model from Table 2. We divided the range of time of advice into 8 treatment intervals. In Table 4, we present the balancing properties of the different intervals before and after the GPS-adjustments. Following the technique of

[13] To conserve space, full results about this estimation are shown in Appendix 1.

[14] We also ran models with these observations; results were consistent, although the balance properties were not optimal.

[15] Similar to an OLS, the multivariate outliers are the observations with the biggest standard errors.

Table 4
Balancing properties given the GPS – T-statistics for equality of means.

Treatment (total time)	Unadjusted								Adjusted for the GPS							
Upper bound (hours)*	3.5h	5h	7.5h	10h	15h	25h	40h	435h	3.5h	5h	7.5h	10h	15h	25h	40h	100h
Selection criteria																
Established firm	**3.48**	−0.67	−1.20	0.93	−0.46	−0.23	−1.88	**−4.45**	−0.15	−0.72	−0.88	1.04	0.26	0.17	−1.46	0.04
Gross revenue (ln)	**10.02**	−0.36	**−1.98**	0.39	**−3.89**	**−3.91**	**−3.08**	**−6.60**	1.00	−0.54	−0.13	1.49	−1.95	−1.63	−0.74	−0.89
Industry: manufacturing	**3.74**	0.72	−0.43	−0.85	−1.46	−0.11	−0.55	**−3.97**	−0.56	0.46	0.19	0.43	−0.27	−0.11	0.53	0.34
Referred firm	**2.25**	0.93	0.59	0.36	**−2.05**	−1.73	−0.66	**−2.08**	0.01	0.57	0.12	1.29	−0.66	0.24	0.46	−0.75
Innovative firm	**3.66**	−0.71	−0.99	−0.87	−0.07	−0.95	−1.35	**−2.42**	−0.74	−0.56	0.73	−0.98	1.46	−0.06	0.58	−0.72
About the client																
Self-motivated to join	**−3.90**	0.28	0.75	**2.15**	0.74	**3.92**	−0.79	1.70	−0.22	0.25	0.14	0.61	−0.02	1.09	−1.01	−0.13
Advising area: sales increase	**5.41**	−1.10	−1.26	−1.41	−2.17	−0.79	−0.01	**−2.38**	0.87	−0.72	−0.76	−1.44	0.33	0.05	1.19	−0.34
Advising area: capital	1.43	0.83	0.20	−0.91	0.69	−0.09	**−2.92**	−0.36	−0.47	0.57	0.42	0.17	0.54	0.96	**−2.94**	−0.54
Advising area: start-up	**−2.65**	−0.04	1.14	−0.97	1.87	−0.87	**2.32**	**4.94**	0.01	−0.28	1.76	−0.91	1.82	−0.11	1.46	1.46
Female owned firm	1.40	−1.09	1.62	−0.10	−0.82	**−2.16**	0.23	−0.72	−0.78	−0.97	1.57	1.54	0.10	−1.92	0.77	0.57
Firm in Santa Barbara	−0.15	0.66	−0.09	−1.23	−0.42	−0.39	1.62	1.24	0.40	0.83	0.35	−1.41	−1.62	0.01	1.06	0.10
Industry: type 2	−0.85	−0.59	−1.09	0.71	−0.81	1.86	−0.38	**3.40**	0.47	0.03	−0.36	−1.27	−1.76	0.82	−0.96	0.87
Industry: type 3	−0.85	−0.59	−1.09	0.71	−0.81	1.86	−0.38	**3.40**	−0.34	−0.28	0.47	0.38	0.64	−0.45	0.27	−0.15
About the main advisor																
A. has an MBA/Ph.D.	**−9.82**	0.12	0.58	**3.45**	**2.72**	**5.36**	**5.23**	**5.90**	−1.55	0.10	−0.79	0.89	0.08	1.72	**3.54**	0.49
# of prior business startups	**−2.64**	0.06	0.54	0.78	**3.55**	0.44	0.08	−1.45	−1.68	0.55	0.95	0.56	**2.52**	0.59	0.39	−1.63
About the initial meeting																
Positive gut feel	**5.10**	−1.19	−0.04	−0.57	**−2.50**	−1.54	−1.72	**−2.30**	−0.30	−0.80	0.10	1.46	−1.66	0.83	0.18	0.16
Negative gut feel	0.83	**2.52**	−1.32	−0.72	0.44	−1.30	0.11	−0.69	0.43	**2.28**	−0.18	−1.56	0.91	−0.30	−0.76	0.45
Collaborative learning	**10.13**	−1.62	**−2.16**	**−3.36**	**−3.86**	**−4.76**	−1.15	**−3.66**	0.67	−1.25	−1.06	−0.11	0.59	0.41	1.87	−0.02
Bridging	−1.20	−1.37	−0.62	1.73	0.56	**2.15**	**2.52**	−1.38	0.66	−1.53	−0.44	−0.52	−0.43	1.58	1.90	0.75
N	823	115	243	136	137	123	72	63	744	107	223	131	123	111	63	38

Notes: t-values reported in bold indicate a significance at 0.05 level. Overall, balancing property is satisfied at a level lower than 0.01.

* While the different groups might look uneven, the choice of division is not arbitrary. The lowest group (below 3.8 h) contains the clients that have not been activated. Then, for activated clients, the SBDC assigned them at least up to five hours of treatment. This is also the threshold for treatment used by Chrisman et al. (2005). The next bound, 7.5 h, represent the median of amount of treatment provided by SBDC. 10 h and 15 h are the other threshold of treatment we used in the YTS analysis reported in appendix. For the bigger dose of treatment, we made sure the groups remained large enough to avoid small sample bias.

Table 5
Parameter estimates of conditional distribution of the milestones given the treatment.

Dependent variable:	Model 1 Reaching a milestone	Model 2 # of milestones reached
Treatment	−0.032	−0.297**
	(0.042)	(0.103)
Treatment squared	0.057***	0.223***
	(0.011)	(0.027)
GPS	0.236	2.837**
	(0.438)	(1.067)
GPS squared	−0.975	−4.954**
	(0.741)	(1.805)
Treatment x GPS	0.191**	0.035
	(0.067)	(0.164)
Constant	−0.013	−0.290*
	(0.055)	(0.135)
Observations	1540	1540
Adj. R-squared	0.304	0.309

Notes: *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$. Two-tailed test. GPS stands for Generalized Propensity Scores used in the dose response matching analysis. SBDC dataset, activation stage. Standard errors in parentheses.

Hirano and Imbens (2004), each value is a t-statistics testing for the difference between a given dose of treatment (treatment interval) and the other doses, for each variable. The balance property improved after adjustment with the GPS. All the variables about the selection criteria (formal criteria and innovation) satisfy the balancing property at a level of 0.05 (standard two-sided *t*-test) after adjustment using the GPS. For the additional variables about the client, only one did not satisfy the balancing property for only one treatment interval: advising area (capital). However, this variable is not significantly correlated with the treatment dose received in the OLS model (first stage equation). All in all, nearly all the t-tests (148 out of 152) are statistically insignificant after adjustment using the GPS and the balancing properties are satisfied overall at a level lower than p-value < 0.01. As a final check to reduce any risk of misspecification from any remaining imbalance, we also used more restrictions for outliers and common support and found consistent results.

After correcting the selection bias in allocated treatment (advising hours), we then estimated the parameters of our outcome variable (growth milestones) as a conditional distribution for different doses of treatment and the GPS. Similar to Hirano and Imbens (2004), we used a dose-response function for these two variables, including a quadratic form and their interaction. Table 5 provides the estimates of the Dose Response Function. In Table 5, the treatment dose (in advising hours) squared is positively related with both the likelihood of reaching a milestone (Model 1) and the number of milestones reached (Model 2), while accounting for the differences between ventures using the GPS. In Fig. 3, we plot the relationship between the treatment dose and the number of milestones reached, where we reveal a curvilinear relationship. This means that a small amount of treatment does not bring any measurable benefit, but only after some length of treatment, the venture will start to increase the likelihood of achieving growth milestones. Moreover, as an implication of the U-curve relationship, allocating a small, but equal amount of treatment to all clients does not yield any positive effects on performance, but more treatment will eventually lead to a higher number of milestones reached. In Table 6, we calculated the cost of allocating a small amount of treatment on many clients (as opposed to a large amount on a few). In our sample, clients receiving up to 5 advising hours of treatment cost the SBDC three times more advising hours per milestone, on average (as most of them reach no milestones), than the clients receiving more than 10 advising hours. In our sample, the “low-treatment” group, 938 firms receiving 5 h or less of treatment, cost about 2220 h to the SBDC, and reached 42 milestones (53 h of work for 1 milestone, on average). The “high-treatment” group, 379 firms receiving above 10 h of treatment,

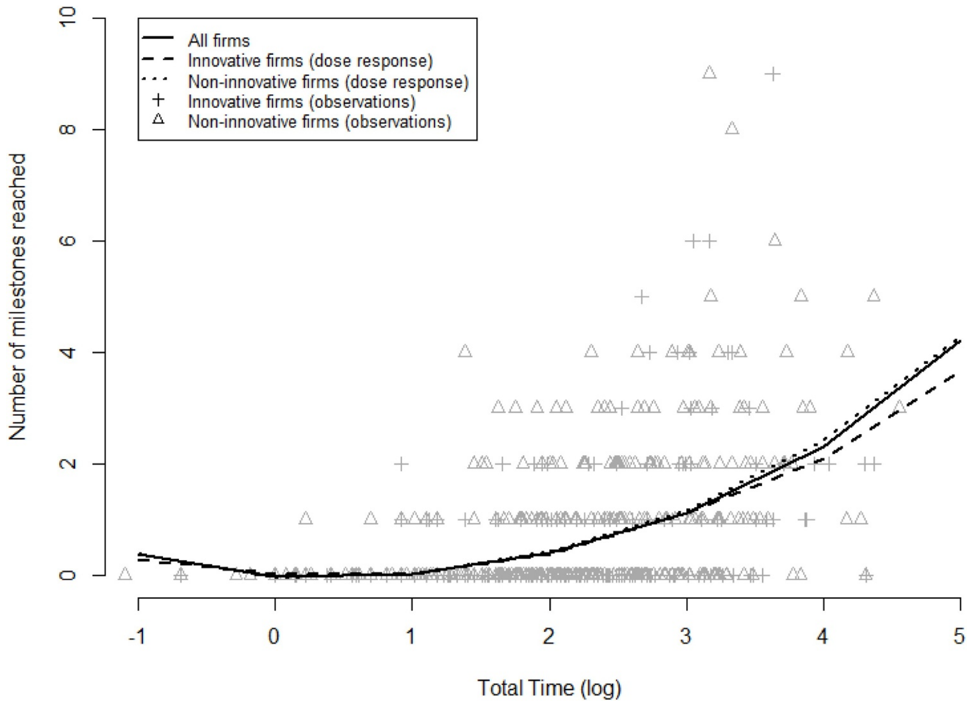

Fig. 3. Dose response function (outcome: number of milestones).

Table 6
Return on investment per level of treatment.

Amount of treatment received	**Low treatment group** 5 h or less	**Limited treatment group** 5 to 10 h	**High treatment group** Above 10 h
N firms	938	379	379
Total time (hours)	2220	2714	9144
N milestones	42	108	496
ROI for SBDC	1 milestone for **53 h** of work	1 milestone for **25 h** of work	1 milestone for **18 h** of work

cost 9144 h to the SBDC, and reached 496 milestones (18 h of work for 1 milestone, on average). In other words, the return on investment is higher when the investment per client is higher. This means that with limited resources, it is more beneficial to treat some ventures with more hours than simply distributing fewer hours across all ventures. This requires imposing formal selection criteria upfront to prioritize which ventures should receive the additional hours. We also note that the GPS positively and significantly interacts with the treatment in Model 1 (DV: reaching a milestone). This means that the treatment is increasingly effective when there is a higher propensity to receive a higher dose of treatment. However, this result is unstable when using multiple milestones. One possible explanation is that the mechanisms for reaching a single milestone are different than those necessary for repeated growth outcomes.

Additionally, we plot in Fig. 3 the dose response comparing the innovative and the non-innovative ventures (by subsamples).[16] With the average hours of treatment, being innovative makes no difference, meaning that they achieve a similar number of growth milestones as non-innovative ventures. Innovative ventures benefit less from an increase in the number of hours of treatment, but the difference is not significant due to a small number of cases. We conducted a counterfactual descriptive analysis to further investigate these findings.

5.1.1. *Counterfactual analysis: what if innovativeness is prioritized?*

In this sub-section, we investigate what would happen if innovativeness was a formal selection criterion in the SBDC setting. To that end, we exploit the fact that some innovative ventures meet one of the four formal criteria (very few meet several of them), whereas others do not meet them. The purpose is to draw a comparison between innovative ventures that were (formally) prioritized in receiving support versus the ones who were not. Table 7 presents the descriptives of this counterfactual analysis (N and %), following the process from selection to outcome through treatment. From top to bottom, we grouped the clients by criteria.

This analysis suggests that innovative ventures would benefit from being prioritized as a formal criterion in its scoring tool. In absolute terms (N), we see that ventures that do not meet at least one formal criterion, be it innovative or not, are far less numerous to be activated, to reach a first milestone, and to receive treatment after reaching a milestone. They are also far less numerous in reaching multiple milestones. Moreover, after reaching a first milestone, the number of innovative clients that do not meet a formal criterion declines sharply, well below the other groupings. These findings strengthen the plausible explanation that since innovativeness is not a formal criterion,

[16] We also looked for other heterogenous effects between treatment and different subsamples of the four formal criteria and growth motivations (Caliendo and Künn, 2011). The different treatment effects were insignificant.

Table 7
Counterfactual Analysis with Innovative Ventures.

Meets at least one formal criteria	Venture is innovative	Total clients N	Up to reaching a milestone: Activated N	Activated *Ratio*	Reach a milestone N	Reach a milestone *Ratio*	After reaching a milestone: Receive treatment after one milestone N	Receive treatment after one milestone *Ratio*	Reach more than one milestone N	Reach more than one milestone *Ratio*
Formal criteria	Innovative	487	338	*69%*	115	*34%*	95	*20%*	60	*12%*
	Not innovative	997	595	*60%*	185	*31%*	139	*14%*	100	*10%*
No formal criteria	Innovative	71	43	*61%*	12	*28%*	4	*6%*	5	*7%*
	Not innovative	157	92	*59%*	23	*25%*	21	*13%*	14	*9%*

Notes: The formal criteria from the SBDC scoring model are (not) being established, having high gross revenue (3rd quartile), being a manufacturing firm, or being refereed by another program.

innovative ventures receive less support throughout the SBDC process, especially as the advisors cannot leverage the formal criterion that would legitimize the allocation of additional hours after reaching a first milestone. In relative terms, we see that the percentage of ventures reaching one or several milestones are higher for ventures that meet at least one formal criterion compared with the ones which do not, regardless of whether being innovative or not. We also observe that innovative ventures with at least one formal criterion have the highest percentage of reaching one or several milestones of all the groupings, whereas innovative ventures without a formal criterion have the lowest percentage of reaching more than one milestone. Formally prioritizing ventures also matters in relative terms. Our curvilinear finding between treatment and growth reinforces the benefits of priority allocations of advising hours because public resources are always limited.

5.2. Heckman treatment effect

Although propensity-score based matching effectively addressed observable selection biases related to the formal criteria, it is not an efficient method. It forced us to remove all the observations outside of the common support zone and drop outliers to fulfill the balancing property for a comparable group of ventures. Moreover, it did not balance out *unobservable differences* between treated and non-treated groups. The SBDC multi-stage selection process affects the pool of applicants that actually received a treatment (from 1,762 applicants to 774 clients advised more than five hours – 44%). As a result, we need another set of analyses to account for this unobservable selection bias.

We used a Heckman-type correction to (1) address the unobservable issue, (2) correct for the effect of selection on growth, and (3) assess the effect of the formal selection criteria if the venture is selected for treatment (Huergo and Moreno, 2017). More specifically, we used the Heckman treatment effect model because the outcome (achieving a growth milestone) is accessible to both treated and non-treated groups (Guo and Fraser, 2010); receiving advice is not a necessary condition to perform better, a priori. In addition to addressing factors leading to unobserved sample selection processes, Heckman models provide another benefit. Similar to matching models, when appropriately implemented, the results themselves qualify as a causal interpretation since they account for confounding factors produced by the formal selection criteria (Antonakis et al., 2010).

The Heckman treatment effects models required two steps to execute: (1) we evaluated the selection bias, using an instrumental variable that is a good predictor of the *selection* and not a predictor of the *outcome*, which is known as the exclusion restriction condition (Bascles, 2008; Wolfolds et al., 2019). (2) We then employ the instrument to correct for the selection bias in our model predicting growth of the SBDC clients. We used *distance to SBDC* (1 = more than 2 h of driving) as an instrument variable for the Heckman correction. Distance to a program facility has been used in multiple previous studies as a valid instrument to correct for the selection bias in advising programs (e.g. Card, 1993) and as a key factor for entrepreneurs to receive support (e.g., Powell et al., 2002). Given that longer distances reduce the opportunities for client-advisor interactions, it should diminish the likelihood of being activated, but not the likelihood of reaching growth milestones. Our results in Model 1 and 2 in Table 8 show that this exclusion restriction condition is met.[17] When we compare the logit Models 1 and 2, we see that the instrument correlates negatively with the treatment (Model 1, effect on treatment, $\beta = -0.572$; p-value < 0.05), but is not correlated to the outcome (Model 2, effect on outcome, p-value > 0.1). The Heckman procedure is efficiently correcting for the identified selection bias using an instrument that predicts the selection but not the outcome: while distance to SBDC is correlated to the likelihood of being activated and receiving the treatment, ventures do not perform better or worse based on their relative distance to SBDC. As such, we confirmed that distance is a valid instrument to assess the Heckman correction (Guo and Fraser, 2010). Finally, in all the models, the rhô is significant. This shows the correction is necessary for the relevant level of selection (Huergo and Moreno, 2017).

After the two models to assess the quality of the instrument (Models 1 and 2), we present eight regression models using a Heckman-type correction to assess for selection to treatment. Models 3 to 10 are Heckman treatment effect models, to correct for selection bias of the activated ventures (see Fig. 2). Two of the four formal criteria (gross revenue and manufacturing firm) are positively and significantly correlated to the growth outcomes in all the models (3 to 10). However, an established firm is negatively correlated with the growth outcomes (also consistently in all models) and being referred shows unstable results (positive and significant on the number of milestones reached and not significant otherwise). As such, the SBDC formal selection criteria are not strong indicators of growth. Also, while being innovative is positively and significantly correlated with the likelihood of being activated (Model 1, selection model) and with the amount of treatment received (see Table 2, OLS model), it does not have an effect on growth in any model, which means that innovative ventures achieve growth similarly to other treated ventures (consistent with the results of the GPS analysis). These findings also show that the "picking winners" approach is not necessarily associated with growth.

Across all models (3 to 10), the effect of treatment on achieving one or multiple milestones is highly significant. These consistent findings mean that the treatment provided by the SBDC advisors translates into growth, net of selection bias and reinforce the GPS results (reported in Section 5.1) regarding program effectiveness.[18] Finally, we observed some statistically significant relationships regarding the client-advisor

[17] The distance to SBDC is the driving distance between the address of the client and the address of the SBDC, accounting for usual traffic, calculated using the OpenStreetMap API. In Appendix 5, we also present the distribution of travel time between clients and SBDC.

[18] We also tested models with both the linear (main) and curvilinear (squared) measure of advising effort and found consistent results (Toft-Kehler et al., 2014).

Table 8
Heckman Treatment Effect models.

	1	2	3	4	5	6	7	8	9	10
Model	Probit	Probit	Heckman Treatment Effect Models							
Dependent variable	Activated	Achiev. a milestone	Achiev. a milest.		# milestones		Achiev. a milest.		# milestones	
Advising effort (ln hours)		1.392***		0.104***		0.237***		0.103***		0.236***
		(0.110)		(0.009)		(0.022)		(0.009)		(0.022)
Client activated		1.172**	−0.331***	−0.490***	−0.759***	−1.192***	−0.333***	−0.491***	−0.760***	−1.193***
		(0.406)	(0.015)	(0.019)	(0.035)	(0.058)	(0.015)	(0.019)	(0.035)	(0.058)
Client dropped out			−0.100***		−0.196***		−0.101***		−0.199***	
			(0.019)		(0.054)		(0.019)		(0.054)	
Formal criteria:										
Established firm	−0.301*	−0.313	−0.083**	−0.077**	−0.259**	−0.250**	−0.083**	−0.078**	−0.259**	−0.251**
	(0.136)	(0.189)	(0.026)	(0.025)	(0.079)	(0.077)	(0.026)	(0.025)	(0.079)	(0.077)
Gross revenue (ln)	0.164***	0.152***	0.047***	0.042***	0.138***	0.129***	0.047***	0.042***	0.138***	0.129***
	(0.024)	(0.030)	(0.004)	(0.004)	(0.013)	(0.013)	(0.004)	(0.004)	(0.013)	(0.013)
Industry: manufacturing	0.807***	0.460	0.157***	0.134***	0.475***	0.436***	0.158***	0.135***	0.476***	0.437***
	(0.211)	(0.296)	(0.039)	(0.038)	(0.119)	(0.115)	(0.039)	(0.038)	(0.119)	(0.115)
Referred firm	0.100	0.240	0.044	0.044	0.174*	0.174*	0.044	0.045	0.175*	0.175*
	(0.135)	(0.183)	(0.026)	(0.025)	(0.078)	(0.076)	(0.026)	(0.025)	(0.078)	(0.076)
Innovative firm	0.299*	−0.060	0.037	0.032	0.050	0.043	0.036	0.031	0.049	0.042
	(0.122)	(0.167)	(0.023)	(0.022)	(0.071)	(0.068)	(0.023)	(0.022)	(0.071)	(0.068)
Self-motivation to join	−0.138	0.185	0.001	0.005	−0.036	−0.028	0.001	0.005	−0.036	−0.028
	(0.135)	(0.208)	(0.027)	(0.026)	(0.082)	(0.079)	(0.027)	(0.026)	(0.082)	(0.079)
Advising area: sales incr.	0.527**	0.097	0.089**	0.076*	0.414***	0.391***	0.087**	0.074*	0.411***	0.389***
	(0.179)	(0.220)	(0.033)	(0.031)	(0.099)	(0.096)	(0.033)	(0.031)	(0.099)	(0.096)
Advising area: capital	0.177	0.222	0.051	0.045	0.054	0.046	0.050	0.045	0.053	0.045
	(0.181)	(0.245)	(0.035)	(0.033)	(0.105)	(0.101)	(0.035)	(0.033)	(0.105)	(0.101)
Advising area: start-up	−0.123	0.425	0.012	0.017	−0.034	−0.025	0.009	0.014	−0.037	−0.029
	(0.143)	(0.220)	(0.029)	(0.027)	(0.087)	(0.084)	(0.029)	(0.027)	(0.087)	(0.084)
Female owned firm	0.165	0.362*	0.082***	0.075**	0.162*	0.147*	0.082***	0.075**	0.161*	0.146*
	(0.125)	(0.172)	(0.024)	(0.023)	(0.073)	(0.071)	(0.024)	(0.023)	(0.073)	(0.071)
Firm in Santa Barbara	0.066	0.029	−0.003	0.002	−0.155*	−0.143	−0.006	−0.0003	−0.157*	−0.146
	(0.134)	(0.194)	(0.025)	(0.024)	(0.077)	(0.074)	(0.025)	(0.024)	(0.077)	(0.074)
Industry: type 2	0.297*	0.494*	0.065*	0.067*	0.113	0.123	0.063*	0.066*	0.111	0.121
	(0.152)	(0.250)	(0.030)	(0.029)	(0.091)	(0.088)	(0.030)	(0.029)	(0.091)	(0.088)
Industry: type 3	0.231	0.530	0.088*	0.081*	0.146	0.132	0.087*	0.080*	0.144	0.130
	(0.179)	(0.278)	(0.035)	(0.034)	(0.107)	(0.104)	(0.035)	(0.034)	(0.107)	(0.104)
Adv. has MBA or Ph.D.	−0.876***	0.104	−0.126***	−0.119***	−0.181**	−0.178**	−0.124***	−0.117***	−0.178**	−0.175**
	(0.113)	(0.171)	(0.022)	(0.021)	(0.067)	(0.065)	(0.022)	(0.021)	(0.067)	(0.065)
Adv. # of prior bus. startups	0.022	−0.314**	−0.029*	−0.035**	−0.073	−0.086*	−0.030*	−0.036**	−0.074	−0.088*
	(0.074)	(0.096)	(0.014)	(0.013)	(0.042)	(0.041)	(0.014)	(0.013)	(0.042)	(0.041)
Positive gut feel	15.898*	10.711	3.613**	3.296**	6.993	6.455	3.573**	3.265**	6.932	6.406
	(6.656)	(8.583)	(1.236)	(1.184)	(3.737)	(3.610)	(1.234)	(1.184)	(3.737)	(3.611)
Negative gut feel	11.478*	−5.215	0.671	0.936	3.325	4.170	0.696	0.956	3.359	4.196
	(5.837)	(8.460)	(1.120)	(1.073)	(3.388)	(3.274)	(1.119)	(1.073)	(3.388)	(3.275)
Collaborative learning	0.728***	0.144	0.130***	0.112***	0.319***	0.288***	0.151***	0.131***	0.347***	0.315***
	(0.112)	(0.168)	(0.022)	(0.021)	(0.067)	(0.065)	(0.023)	(0.022)	(0.069)	(0.067)
Bridging activity	−0.232	−0.121	−0.041	−0.041	0.072	0.070	0.012	0.007	0.144	0.138
	(0.132)	(0.203)	(0.026)	(0.025)	(0.079)	(0.076)	(0.031)	(0.030)	(0.089)	(0.087)
Collab. learning X bridging							−0.092**	−0.082**	−0.124	−0.116
							(0.029)	(0.029)	(0.071)	(0.072)
Instrument: driving distance (> 2 h)	−0.572*	−0.625								
	(0.285)	(0.554)								
Constant	−0.317	−5.901***	0.226***	0.176***	0.424**	0.346**	0.221***	0.172***	0.415**	0.339*
	(0.228)	(0.511)	(0.045)	(0.044)	(0.137)	(0.132)	(0.045)	(0.044)	(0.137)	(0.132)
Log Likelihood	−965		−1270	−1217	−2999	−2946	−1265	−1213	−2998	−2945
Rhô			.956***	.949***	.990***	.989***	.956***	.950***	.990***	.989***
			(0.003)	(0.004)	(0.001)	(0.001)	(0.003)	(0.004)	(0.001)	(0.001)

Notes: *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; Two-tailed test; For all the models, the treatment is being activated for receiving advice; the first step equation includes the distance to SBDC as an instrumental variable to predict the selection. Since *advising effort* and *client dropped out* variables are correlated, we included them in different models. Standard errors in parentheses.

relational variables. In all our models, we found that when collaborative learning occurs, it is positively correlated with reaching a milestone and to the number of milestones reached. In fact, when clients and advisors have thorough discussions about the client's growth objectives, the client is more likely to be assigned more advising hours, to remain in the program longer, and to accept the treatment provided by their advisor. As a consequence, clients with a collaborative learning behavior get a bonus of milestones (β = 0.319 extra milestones on average in Model 5, β = 0.288 in Model 6) for the same amount of treatment. In other words, our Model 6 (Table 8) predicts that a collaborative client that receives 15 h of treatment will reach as many milestones, on average, than a non-collaborative client who would receive 51 h, everything else made equal. From these results, we see a clear difference in achieving growth milestones driven by the client's willingness to collaborate with their advisors, a feature that is difficult to assess only with selection. The client's learning behavior and subsequent solution proposed by the advisor during the meeting are strong and consistent predictors of the client remaining in the program.[19] Thus, advisors play a central role in evaluating and then sustaining the treatment efforts with their clients.[20]

Taken together with the GPS findings, these results show that if entrepreneurship support programs have a strong incentive to allocate resources to the most promising ventures, relying solely on observable characteristics (including innovativeness) may not be sufficient for achieving growth. Simply relying on a “picking winners” strategy (or any selection criteria) does not ensure that participants will achieve growth resulting from their treatment. Instead, the potential success of the treatment depends on whether the client can collaborate with their advisors to problem solve innovative solutions and eventually gain the knowledge to achieve their growth objectives.

We also controlled for the effect of bridging advice on reaching a milestone. When advisors connect their clients to other support providers, the clients gain access to needed resources, financiers, and experts who can help their ventures grow (Amezcua et al., 2013; Kim and Longest, 2014; Storey, 1994). As an alternative explanation of the mechanism driving growth in the counseling activity, however, bridging is not significant in any of the models (Table 8). The lack of a relationship indicates that providing referral advice is not enough for participants to achieve their growth outcomes and a deeper advising relationship is required. Another explanation might be that entrepreneurs have many other ways to reach relevant contacts in southern California, and the added value of this activity compared to collaborative learning is not as valuable. Additionally, we also controlled for the moderating effect of bridging on collaborative learning (Models 7 to 10). Bridging advice negatively moderates the relationship between the collaborative learning and reaching a milestone (but not for reaching multiple milestones). One interpretation of this result is that collaborative learning clients are less likely to benefit from the treatment if they are referred out to different resources, since they prefer more in-depth consultations with their advisors.

5.3. Difference-in-difference exact matching

Our GPS analysis potentially suffers from a second type of issue related to other unobserved (omitted) factors that might affect the assignment to treatment. Unobserved differences might explain the treatment and the outcome and hence violate the CIA assumption. These unobserved differences could be time-varying or time-invariant and related to other factors associated with applying to SBDC. To solve for this issue, we conducted a third estimation procedure that compares the SBDC applicants with firms in the southern California region using the YTS data. Given its longitudinal feature, the YTS dataset allows us to include a pre-treatment growth variable in our model, an important factor in support program analyses (Caliendo et al., 2017), and to use a difference-in-difference estimator, which cancels out time-invariant omitted variables (Angrist and Pischke, 2008).[21] Given its large size, the YTS also enables us to use an exact matching procedure that significantly enhances the quality of the balancing of differences (Iacus et al., 2012). Also, relying on a different dataset with a different measurement of time, growth milestones, and innovativeness has the advantage to address potential concerns about measurement bias (Shadish et al., 2002).

In this third analysis, we account for the firms’ growth trajectories two years prior to joining SBDC, as well as observable characteristics of the firms, such as age, size, industry, and location. We show that treated ventures (10 and 15 h of treatment) show a significant difference in growth to similar firms in the region that did not apply to the SBDC. However, innovative firms still perform similarly than non-innovative ones, confirming our prior GPS results. To conserve space, we report complete details about our empirical choices and results for this additional analysis in Appendix 1.

5.4. Results summary

We conducted three sets of analyses addressing distinctive biases from unobserved factors and using different measures of growth from difference samples and for different years. Regardless of the technique and model, we observed that treatment had a positive and significant relationship to achieving growth if the treatment occurred with sufficient quantity (net of selection biases generated from prioritizing formal criteria and from self-selection). In fact, we showed that the effect of treatment on growth is (1) curvilinear and (2) time-lagged, which we infer as evidence for a learning curve mechanism (Musaji et al., 2020; Toft-Kehler et al., 2014). At the same time, while we show that the formal selection criteria may have an impact on performance (especially venture age and size), innovative ventures achieve growth similar to other ventures as a result of the same amount of treatment. Our counterfactual analysis showed that most innovative ventures are excluded from treatment after the first milestone; since growth is more likely when sufficient treatment is provided, we conclude that innovative ventures need to be prioritized in the original scoring model to ensure they receive enough advising hours to be able to achieve their growth milestones. Our Heckman analysis revealed that clients who engage collaboratively with their advisors are more likely to receive longer treatment and achieve more milestones. Since collaborative learners cannot be assessed prior to the treatment commencing, it indicates that relying on formal selection criteria alone is not enough to ensure growth when allocating support initially.

6. Discussion

In our study, we tackled the question of whether innovative ventures achieve more growth by having received business advice in a publicly sponsored entrepreneurship support program. To answer this question, we analyzed over 1,700 businesses that sought advice from the Small Business Development Center of Ventura and Santa Barbara Counties in Southern California. By employing a series of analytical procedures, we determined that for innovative ventures to achieve more growth milestones, they need sufficient advising support. This depends first on having the formal criteria used to allocate support; thus, innovativeness must be included as a formal criterion to ensure these ventures receive sufficient support. However, we also concluded that once *granted* support, ventures that openly socialized their growth objectives with their advisors were best positioned to learn from them and translate their advice into achieving more growth milestones than other ventures. In this section, we highlight how our study advances our understanding of how entrepreneurship-support programs can promote the growth of innovative ventures.

6.1. Formalizing innovativeness as a selection criterion – getting in formally

Our empirical analyses consistently revealed that innovative ventures were likely to achieve growth but in ways similar to other ventures. This outcome is surprising given the high expectations for innovative ventures: to grow more than others, produce new jobs, inject

[19] Of all clients who receive more than five hours of treatment, 58% engage in collaborative learning with their advisors, compared to 38% who opt for ready-made solutions.

[20] We checked whether innovativeness and the four formal selection criteria strengthen this collaborative learning effect. We did not find any significant finding regarding these interactions. In our results, the picking winners approach is independent from collaborative learning, which confirms that both approaches are needed for program effectiveness.

[21] In Appendix 4, we report a sensitivity analysis that addresses time-varying omitted variables (Rosenbaum, 2002).

new knowledge and technologies into the local ecosystems, and provide an outsized stimulus into their economies (Audrestch, 1995; Coad and Rao, 2008; Colombelli et al., 2013). This is especially relevant for regions that wish to turnaround the loss of innovative firms and jobs (Atkinson et al., 2019). Since the SBDC study context did not specifically focus on innovative ventures, one conclusion we draw from our analysis is that the way in which programs prioritize certain types of participants matters for how they allocate support. Our findings show that innovative ventures are not guaranteed to grow more than other ventures, unless support is targeted directly for them. This provides one explanation for why past research has not shown more consistent findings regarding the positive impact of program support for innovative ventures. If additional support is explicitly targeted at innovative ventures, they have more opportunities to receive in-depth advice, learn from experts, and translate this into productive growth outcomes.

While this insight seems straightforward, we underscore the importance of these priorities and the ways they are codified through policies and subsequently implemented through programs. Our findings reinforce the role that organizations such as the SBDC, and their personnel, play in shaping these policies (Dobbin, 2009). Given limited resources, these organizations depend on formal criteria to evaluate and accept program participants. Our research highlights that support organizations can play a role in offering preferential treatment to certain participants, such as innovative ventures, which represent a regional or national priority, and grant additional resources to promote their growth. This is in line with other research on organizational sponsorship or institutional intermediaries that play a similar role within regional ecosystems (e.g. Amezcua et al., 2013; Armanios et al., 2017; Cohen et al., 2019). To take it one step further, if policymakers wish to target specific sectors, we argue it is necessary for these sectors to be explicitly included in any formal criteria used to allocate support, to ensure these participants are actively recruited, matched with qualified advisors, and allowed sufficient time to learn from them and translate their knowledge into growth. According to the organization and social evaluation literature, formal criteria are legitimizing devices that shape the distribution of attention and rewards in organizations (Lamont, 2009; Rivera and Tilcsik, 2019). Once formalized, these criteria become objective standards by which to evaluate and select ventures to treat (Kim et al., 2019). They also safeguard programs against selecting solely on idiosyncratic preferences driven by program administrator experiences and backgrounds.

6.2. Treating the learners – getting through collaboratively

Given the importance of formal criteria in allocating resources, one would expect that a carefully designed allocation process would ensure growth among program participants in a productive manner. However, our findings show that formal criteria simply based on "picking winners" are not enough to guarantee this outcome. Once support has been allocated to participants, achieving growth requires a collaborative learning approach between client and advisor. While it is important to carefully design policies for allocating resources, formal criteria should not be the only levers policymakers use to support innovative ventures. It is also critical to design programs that encourage meaningful client-advisor interactions to enable socialized learning and generate innovative solutions.

Given our focus on advice and counseling efforts, we align with research that depicts entrepreneurship as a collective process dependent on multiple stakeholders (Aldrich and Kim, 2007). Advice can come from a variety of sources (Ruef, 2010), but most entrepreneurs depend on family, friends, or peers for informal advice and support (Kim et al., 2013; Kim and Longest, 2014; Chatterji et al., 2019). Relying on others for advice doesn't guarantee that all entrepreneurs can convert this advice into actual, internalized learning that yields demonstrable growth. Rather than a task-oriented approach that looks for a single answer, we argue that clients who enter the advising relationship with a learning mindset are more likely to achieve growth milestones.

An important challenge for policymakers is how best to design programs that enroll participants who are willing to engage in high-quality interactions with their advisors. Although research points to innovative ventures having a greater willingness to learn because of their knowledge-based foundations (Colombelli et al., 2016), they still have considerable heterogeneity among them (Mustar et al., 2006). They operate in different sectors and start with different origins, and consequently differ substantially in how they learn (Colombelli et al., 2014; Colombo et al., 2015; Greve, 2007). Some innovative ventures might be more open to seizing learning opportunities, whereas others might be constrained by the intense efforts innovation might require and be less attentive to external support. It is notoriously difficult to evaluate ex ante the actual learning preferences of high-potential ventures (Huang and Pearce, 2015; Nightingale and Coad, 2014). This represents a puzzle for policymakers and entrepreneurship-support program designers: while innovative ventures need to be formally prioritized in order to receive support, it is also not guaranteed these same ventures will translate their advising into actual growth. Our analysis revealed considerable misallocation of advising effort that would be more productively used if the true willingness of participants to commit and learn collaboratively could be discerned prior to treatment. To minimize this loss, programs are increasingly reliant on their advisors to quickly assess which clients are willing learners who intend to work closely with them to tackle their growth problems. Our work establishes this two-step evaluation, both conceptually and empirically, and provides a more holistic perspective on entrepreneurship-support program evaluation that has historically overlooked this separation between selecting the right ventures (and commit Type I errors) and selecting those who will most effectively apply their learnings (and commit Type II errors). More generally, this selection-and-treatment puzzle and corresponding solution applies to other program admissions situations whenever limited opportunities or resources must be distributed to a pool of deserving applicants and the selected applicants receive treatment aimed at their growth. While the initial selection criteria may seem to be an efficient method for allocating support, programs still need to rely on the actual treatment providers to informally determine which participants are best suited to convert the treatment into tangible outcomes. Since "doing more with less" is a constant challenge for public policymakers, our findings can apply beyond efforts just for innovative ventures.

6.3. Avenues for future research

Although we conducted our study as comprehensively as possible, there are some limitations. Our study depends on selection and treatment mechanisms for one program (SBDC), location (Southern California), and treatment type (advising hours). While our primary goal has been to establish internal validity for the treatment mechanisms by adequately accounting for various selection biases, future research can replicate our study design in other contexts and treatment types to determine if similar treatment effects and mechanisms exist. Although matching methods are the most robust methods today to assess a causal link when selection is done on observable factors, we acknowledge that the compared ventures are never identical nor fully representative of the overall population. Although we tried to match ventures on a wide range of venture-, entrepreneur- and behavioral-characteristics, minor differences may appear in some models. However, we believe the consistency of results along the different estimations procedures, and with different data sources in the articles and in the Appendix, are in favor of the robustness of our results. Future research can replicate our study design in other contexts to determine if similar treatment effects and mechanisms exist. For our outcome, we measured growth based on milestone achievements. Although we argued that this measure is appropriate for our setting given the heterogeneity of expectations and growth paths of the ventures in the SBDC program, we realize that policymakers are also interested in other outcomes for innovative ventures. Future research can also examine advising effectiveness in terms of actual growth rates in sales or employees or economic impact generated to their local economies. Finally, our analysis focused solely on how advising time led to growth via learning as a primary treatment mechanism. Future research can probe more deeply into alternative treatment mechanisms, especially in terms of the content and nature of the client-advisor interactions, and also into different kinds of treatment designs and their combination.

7. Conclusion

Policymakers play a vital role in allocating resources to promote innovative entrepreneurship as a means of spurring economic growth in their regions. Our study demonstrated that entrepreneurship-support programs can be most effective when the formal selection criteria are aligned with a particular objective (such as boosting innovative ventures) and even "winning" ventures themselves are also willing to collaboratively learn with their advisors concerning their growth objectives. Since resources are always limited and using public resources requires careful stewardship, policymakers will always face tradeoffs for how to pick the right winners who will generate growth. Our analytical approach offers new insights that addresses empirical shortcomings in past research and informs policymakers about how best to design entrepreneurial-support programs for innovative ventures.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We appreciate the advice from the Special Issue editors, especially, Tommaso Minola, our reviewers, Marco Caliendo, participants at the Baugh-Ratio Symposium and the research support from Nikki Bown.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.respol.2020.104052.

References

Achtenhagen, L., Naldi, L., Melin, L., 2010. "Business growth" – do practitioners and scholars really talk about the same thing? Entrepreneurship 34 (2), 289–316.

Acs, Z.J., 2006. How is entrepreneurship good for economic growth? Innovations 1 (1), 97–107.

Agarwal, R., Shah, S.K., 2014. Knowledge sources of entrepreneurship: firm formation by academic, user and employee innovators. Res. Policy 43 (7), 1109–1133.

Aldrich, H.E., Ruef, M., 2018. Unicorns, gazelles, and other distractions on the way to understanding real entrepreneurship in the United States. Acad. Manag. Perspect. 32 (4), 458–472.

Aldrich, H.E., Kim, P.H., 2007. Small worlds, infinite possibilities? How social networks affect entrepreneurial team formation and search. Strategic Entrep. J. 1 (1-2), 147–165.

Aldrich, H.E., Auster, E.R., 1986. Even dwarfs started small: liabilities of age and size and their strategic implications. Res. Organ. Behav.

Amezcua, A.S., Grimes, M.G., Bradley, S.W., Wiklund, J., 2013. Organizational sponsorship and founding environments: a contingency view on the survival of business-incubated firms, 1994-2007. Acad. Manag. J. 56 (6), 1628–1654.

Angrist, J., Pischke, J., 2008. Mostly Harmless econometrics: An Empiricist's Companion. Princeton university press.

Angrist, J., Pischke, J., 2009. A Note on Bias in Just Identified IV With Weak Instruments. London School of Economics, pp. 28.

Antonakis, J., Bendahan, S., Jacquart, P., Lalive, R., 2010. On making causal claims: a review and recommendations. Leadersh. Quart. 21 (6), 1086–1120.

Armanios, D.E., Eesley, C.E., Li, J., Eisenhardt, K.M., 2017. How entrepreneurs leverage institutional intermediaries in emerging economies to acquire public resources. Strategic Manag. J. 38 (7), 1373–1390.

Arrow, K., 1962. Economic Welfare and The Allocation of Resources for Invention. The Rate and Direction of Inventive Activity. Princeton University Press and NBER.

Atkinson, R.D., Muro, M., Whiton, J., 2019. The Case for Growth Centers: *How* to Spread Tech Innovation across America, Brookings Institution and ITIF. Report available at. https://www.brookings.edu/wp-content/uploads/2019/12/Full-Report-Growth-Centers_PDF_BrookingsMetro-BassCenter-ITIF.pdf.

Audretsch, D.B., 1995. Innovation, growth and survival. Int. J. Ind Organ. 13 (4), 441–457.

Audretsch, D.B., Keilbach, M., 2007. The theory of knowledge spillover entrepreneurship. J. Manag. Stud. 44 (7), 1242–1254.

Audretsch, D.B., Grilo, I., Thurik, A.R., 2007. Handbook of Research On Entrepreneurship Policy. Edward Elgar Publishing.

Autio, E., Rannikko, H., 2016. Retaining winners: can policy boost high-growth entrepreneurship? Res. Policy 45 (1), 42–55.

Autio, E., Kanninen, S., Gustafsson, R., 2008. First- and second-order additionality and learning outcomes in collaborative R&D programs. Res. Policy 37 (1), 59–76.

Autio, E., Kronlund, M., Kovalainen, A., 2007. High-*Growth* SME *Support Initia*tives in nine *Countries: Analysis, Categoriza*tion, and *Recommendations: Report Pre*pared for the Finnish Ministry of Trade and Industry. Ministry of Trade and Industry.

Bascle, G., 2008. Controlling for endogeneity with instrumental variables in strategic management research. Strategic Organ. 6 (3), 285–327.

Baumol, W.J., 1990. The problem of development: introduction. J. Polit. Econ. 98, 1–11.

Baum, J.A., Silverman, B.S., 2004. Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups. J. Bus. Vent. 19 (3), 411–436.

Behncke, S., Frölich, M., Lechner, M., 2010. A caseworker like me–does the similarity between the unemployed and their caseworkers increase job placements? Econ. J. 120 (549), 1430–1459.

Bertoni, F., Martí, J., Reverte, C., 2019. The impact of government-supported participative loans on the growth of entrepreneurial ventures. Res. Policy 48 (1), 371–384.

Birch, D., Haggerty, A., Parsons, W., 1997. Who's creating jobs? Cognetics. . Cambridge, MA.

Brixy, U., Sternberg, R., Stüber, H., 2013. Why some nascent entrepreneurs do not seek professional assistance. Appl. Econ. Lett. 20 (2), 157–161.

Brown, J.D., Earle, J.S., 2017. Finance and growth at the firm level: evidence from SBA loans. J. Finance 72 (3), 1039–1080.

Bulte, E., Lensink, R., Vu, N., 2016. Do gender and business trainings affect business outcomes? Experimental evidence from Vietnam. Manage. Sci. 63 (9), 2885–2902.

Caliendo, M., Kopeinig, S., 2008. Some practical guidance for the implementation of propensity score matching. J. Econ. Surv. 22 (1), 31–72.

Caliendo, M., Künn, S., 2011. Start-up subsidies for the unemployed: long-term evidence and effect heterogeneity. J. Public Econ. 95 (3–4), 311–331.

Caliendo, M., Mahlstedt, R., Mitnik, O.A., 2017. Unobservable, but unimportant? The relevance of usually unobserved variables for the evaluation of labor market policies. Labour Econ. 46, 14–25.

Card, D., 1993. In: Using Geographic Variation in College Proximity to Estimate the Returns to Schooling. National Bureau of Economic Research Working Paper 4483.

Chabaud, D., Messeghem, K., 2014. Les ETI. Rev. Franç. Gest. 7, 95–110.

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

Chrisman, J.J., McMullan, W.E., 2004. Outsider assistance as a knowledge resource for new venture survival. J. Small Bus. Manag. 42 (3), 229–244.

Chrisman, J.J., McMullan, E., Hall, J., 2005. The influence of guided preparation on the long-term performance of new ventures. J. Bus. Vent. 20 (6), 769–791.

Chrisman, J.J., McMullan, W.E., Kirk Ring, J., Holt, D.T., 2012. Counseling assistance, entrepreneurship education, and new venture performance. J. Entrep. Public Policy 1 (1), 63–83.

Clarysse, B., Wright, M., Mustar, P., 2009. Behavioural additionality of R&D subsidies: a learning perspective. Res. Policy 38 (10), 1517–1533.

Coad, A., Rao, R., 2008. Innovation and firm growth in high-tech sectors: a quantile regression approach. Res. Policy 37 (4), 633–648.

Coad, A., Cowling, M., Siepel, J., 2007. Growth processes of high-growth firms as a four-dimensional chicken and egg. Ind. Corp. Change 26 (4), 537–554.

Coad, A., Daunfeldt, S.O., Hölzl, W., Johansson, D., Nightingale, P., 2014. High-growth firms: introduction to the special section. Ind. Corp. Change 23 (1), 91–112.

Cohen, S.L., Bingham, C.B., Hallen, B.L., 2019. The role of accelerator designs in mitigating bounded rationality in new ventures. Adm. Sci. Q. 64 (4), 810–854.

Colombelli, A., 2016. The impact of local knowledge bases on the creation of innovative start-ups in Italy. Small Bus. Econ. 47 (2), 383–396.

Colombelli, A., Grilli, L., Minola, T., Mrkajic, B., 2019. To what extent do young innovative companies take advantage of policy support to enact innovation appropriation mechanisms. Res. Policy.

Colombelli, A., Krafft, J., Quatraro, F., 2014. The emergence of new technology-based sectors in European regions: a proximity-based analysis of nanotechnology. Res. Policy 43 (10), 1681–1696.

Colombelli, A., Krafft, J., Vivarelli, M., 2016. To be born is not enough: the key role of innovative start-ups. Small Bus. Econ. 47 (2), 277–291.

Colombelli, A., Krafft, J., Quatraro, F., 2013. High-growth firms and technological knowledge: do gazelles follow exploration or exploitation strategies? Ind. Corp. Change 23 (1), 261–290.

Colombo, M.G., Doganova, L., Piva, E., D'Adda, D., Mustar, P., 2015. Hybrid alliances and radical innovation: the performance implications of integrating exploration and exploitation. J. Technol. Transf. 40 (4), 696–722.

Colombo, M.G., Cumming, D.J., Vismara, S., 2016. Governmental venture capital for

innovative young firms. J. Technol. Transf. 41 (1), 10–24.
Davidsson, P., Delmar, F., Wiklund, J., 2006. Entrepreneurship and the Growth of Firms. Edward Elgar Publishing.
Dehejia, R.H., Wahba, S., 1999. Causal effects in nonexperimental studies: reevaluating the evaluation of training programs. J. Am. Stat. Assoc. 94 (448), 1053–1062.
Delmar, F., Davidsson, P., Gartner, W.B., 2003. Arriving at the high-growth firm. J. Bus. Vent. 18 (2), 189–216.
Dobbin, F., 2009. Inventing Equal Opportunity. Princeton University Press.
Eckhardt, J.T., Shane, S., Delmar, F., 2006. Multistage selection and the financing of new ventures. Manage. Sci. 52 (2), 220–232.
Godin, B., 2014. Project on the Intellectual History of Innovation. The vocabulary of innovation: a lexicon 20.
Greve, H.R., 2007. Exploration and exploitation in product innovation. Ind. Corp. Change 16 (5), 945–975.
Guo, S., Fraser, M.W., 2010. Propensity Score Analysis. Sage.
Heckman, J.J., LaLonde, R.J., Smith, J.A., 1999. The economics and econometrics of active labor market programs. Handbook of *Labor Econo*mics, vol. 3. pp. 1865–2097.
Heckman, J.J., Smith, J., Clements, N., 1997. Making the most out of programme evaluations and social experiments: accounting for heterogeneity in programme impacts. Rev. Econ. Stud. 64 (4), 487–535.
Hirano, K., Imbens, G.W., 2004. The propensity score with continuous treatments. Applied Bayesian *Modeling* and *Causal Infer*ence from *Incomplete-Data Perspecti*ves 226164. pp. 73–84.
Holland, P.W., 1986. Statistics and causal inference. J. Am. Stat. Assoc. 81 (396), 945–960.
Hottenrott, H., Richstein, R., 2020. Start-up subsidies: does the policy instrument matter? Res. Policy 49 (1), 103888.
Huang, L., 2018. The role of investor gut feel in managing complexity and extreme risk. Acad. Manag. J. 61 (5), 1821–1847.
Huang, L., Pearce, J.L., 2015. Managing the unknowable: the effectiveness of early-stage investor gut feel in entrepreneurial investment decisions. Adm. Sci. Q. 60 (4), 634–670.
Huergo, E., Moreno, L., 2017. Subsidies or loans? Evaluating the impact of R&D support programmes. Res. Policy 46 (7), 1198–1214.
Iacus, S.M., King, G., Porro, G., 2012. Causal inference without balance checking: coarsened exact matching. Polit. Anal. 20 (1), 1–24.
Imbens, G.W., Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. J. Econ. Lit. 47 (1), 5–86.
Kim, P.H., Longest, K.C., 2014. You can't leave your work behind: employment experience and founding collaborations. J. Bus. Vent. 29 (6), 785–806.
Kim, P.H., Kotha, R., Fourné, S.P., Coussement, K., 2019. Taking leaps of faith: evaluation criteria and resource commitments for early-stage inventions. Res. Policy 48 (6), 1429–1444.
Kim, P.H., Longest, K.C., Aldrich, H.E., 2013. Can You Lend Me a Hand? Task-Role Alignment of Social Support for Aspiring Business Owners. Work Occup 40. pp. 211–247. https://doi.org/10.1177/0730888413481365.
Kim, P.H., Longest, K.C., Lippmann, S., 2015. The tortoise versus the hare: progress and business viability differences between conventional and leisure-based founders. J. Bus. Vent. 30 (2), 185–204.
Lamont, M., 2009. How Professors Think. Harvard University Press.
Lechner, M., 1999. Earnings and employment effects of continuous gff-the-job training in east germany after unification. J. Bus. Econ. Stat. 17 (1), 74–90.
Lerner, J., 2002. When bureaucrats meet entrepreneurs: the design of effective public venture capital programmes. Econ. J. 112 (477), F73–F84.
Lerner, J., 2010. The future of public efforts to boost entrepreneurship and venture capital. Small Bus. Econ. 35 (3), 255–264.
Levitt, B., March, J.G., 1988. Organizational learning. Annu. Rev. Sociol. 14 (1), 319–338.
Loughran, T., McDonald, B., 2013. IPO first-day returns, offer price revisions, volatility, and form S-1 language. J. Financ. Econ. 109 (2), 307–326.
Maddala, G.S., 1983. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge University Press.
Mason, C., Brown, R., 2013. Creating good public policy to support high-growth firms. Small Bus. Econ. 40 (2), 211–225.
McKelvie, A., Wiklund, J., 2010. A focus on of growth mode instead of growth Rate. Entrepreneurship 34 (2), 261–288.
Minniti, M., 2008. The role of government policy on entrepreneurial activity: productive, unproductive, or destructive? Entrepreneurship 32 (5), 779–790.
Musaji, S., Schulze, W.S., De Castro, J.O., 2020. How long does it take to “get to” the learning curve? Acad. Manag. J. 63 (1), 205–233.
Mustar, P., Renault, M., Colombo, Moray, N., 2006. Conceptualising the heterogeneity of research-based spin-offs: a multi-dimensional taxonomy. Res. Policy 35 (2), 289–308.
Nightingale, P., Coad, A., 2014. Muppets and gazelles: political and methodological biases in entrepreneurship research. Ind. Corp. Change 23 (1), 113–143.
OECD, 2019. OECD SME and Entrepreneurship Outlook 2019. OECD Publishing, Paris.
Penrose, E., 1959. The Theory of the Growth of the Firm. John Wiley & Sons, New York.
Perrow, C., 1986. Complex Organizations: A Critical Essay. Glenview, IL.
Powell, W.W., Koput, K.W., Bowie, J.I., Smith-Doerr, L., 2002. The spatial clustering of science and capital: accounting for biotech firm-venture capital relationships. Reg. Stud. 36 (3), 291–305.
Rivera, L.A., Tilcsik, A., 2019. Scaling Down Inequality: rating Scales, Gender Bias, and the Architecture of Evaluation. Am. Sociol. Rev.
Rosenbaum, P.R., 2002. Overt bias in observational studies. Observational Studies. Springer, New York, NY, pp. 71–104.
Rosenbaum, P.R., Rubin, D.B., 1983a. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. J. R. Stat. Soc. 45 (2), 212–218.
Rosenbaum, P.R., Rubin, D.B., 1983b. The central role of the propensity score in observational studies for causal effects. Biometrika 70 (1), 41–55.
Rotger, G.P., Gørtz, M., Storey, D.J., 2012. Assessing the effectiveness of guided preparation for new venture creation and performance: theory and practice. J. Bus. Vent. 27 (4), 506–521.
Rowe, T., 2016. A Brief History of America's SBDC Network-America's SBDC. Association of Small Business Development Centers.
Rubera, G., Kirca, A.H., 2012. Firm innovativeness and its performance outcomes: a meta-analytic review and theoretical integration. J. Mark. 76 (3), 130–147.
Rubin, D.B., 1974. Estimating causal effects of treatments in randomized and non-randomized studies. J. Educ. Psychol. 66 (5), 688.
Ruef, M., 2010. The Entrepreneurial Group: Social Identities, Relations, and Collective Action. Princeton University Press.
Shadish, W., Cook, T., Campbell, D., 2002. Experimental and Quasi-Experimental Designs for Generalised Causal Inference. Houghton Mifflin Company, New York.
Shane, S., 2009. Why encouraging more people to become entrepreneurs is bad public policy. Small Bus. Econ. 33 (2), 141–149.
Söderblom, A., Samuelsson, M., Wiklund, J., Sandberg, R., 2015. Inside the black box of outcome additionality: effects of early-stage government subsidies on resource accumulation and new venture performance. Res. Policy 44 (8), 1501–1512.
Stinchcombe, A.L., March, J.G., 1965. Handbook of *Organizations*.
Storey, D.J., 1994. New firm growth and bank financing. Small Bus. Econ. 6 (2), 139–150.
Tetlock, P.R., 2005. Expert Political Judgment: How Good is It? How can We Know? Princeton University Press.
Toft-Kehler, R., Wennberg, K., Kim, P.H., 2014. Practice makes perfect: entrepreneurial-experience curves and venture performance. J. Bus. Vent. 29 (4), 453–470.
Wennberg, K., Delmar, F., Mckelvie, A., 2016. Variable risk preferences in new firm growth and survival. J. Bus. Vent. 31 (4), 408–427.
Widerstedt, B., Månsson, J., 2015. Can business counselling help SMEs grow? Evidence from the Swedish business development grant programme. J. Small Bus. Enterpr. Dev. 22 (4), 652–665.
Wiklund, J., Davidsson, P., Delmar, F., 2003. What do they think and feel about growth? An expectancy-value approach to small business managers’ attitudes toward growth. Entrepreneurship 27 (3), 247–270.
Wolfolds, S.E., Siegel, J., 2019. Misaccounting for endogeneity: the peril of relying on the Heckman two-step method without a valid instrument. Strategic Manag. J. 40 (3), 432–462.
Woolley, J., 2017. Infrastructure for entrepreneurship. Oxford Research Encyclopedia of Business and Management.
Woolley, J.L., Rottner, R.M., 2008. Innovation policy and nanotechnology entrepreneurship. Entrepreneurship 32 (5), 791–811.
Wren, C., Storey, D.J., 2002. Evaluating the effect of soft business support upon small firm performance. Oxf. Econ. Pap. 54 (2), 334–365.