Silicon Valley Everywhere? Place-Based Innovation Policy and the Direction of Entrepreneurial Entry

Freda Fang* Mohaddeseh Heydari-Nejad†

October 1, 2026

Abstract

Place-based innovation policies increasingly target particular industries. We study whether such policies redirect entrepreneurial entry toward targeted industries by making them more salient to potential founders. We exploit accelerator launches by the Creative Destruction Lab, a global mentorship program, each targeting one science-based industry in one city, and link program records to startup and founder data. Among founders who never apply to or join the program, entry into the targeted industry rises by 12 percent relative to other local industries, with no detectable increase in city-wide entrepreneurship. This reallocation is concentrated among science-trained first-time founders and is strongest when the targeted industry has little local presence. We develop and estimate a model of entrepreneurial industry choice with limited consideration that separates the consideration and expected payoff effects of a launch. For a science-trained first-time founder, the estimated consideration effect increases the odds of choosing the targeted industry by a factor of 2.7 at the average launch and 4.3 when the industry is locally emerging. Counterfactual estimates attribute 30 percent of science-trained first-time founders' entry into targeted industries to this consideration channel. Place-based innovation policy can therefore shape the direction of innovation not only through the resources it provides, but also by influencing the industries to which entrepreneurs direct their attention.

*Rotman School of Management, University of Toronto. Email: freda.fang@rotman.utoronto.ca

†Kelley School of Business, Indiana University. Email: mheydari@iu.edu. Acknowledgements: We thank Victor Aguirregabiria, Jessica Bai, Claire Celerier, Gilles Chemla, Michael Ewens, Avi Goldfarb, Juanita González-Uribe, Camille Hebert, and Ting Xu for helpful comments. We also appreciate valuable feedback from participants at the Kelley School of Business seminar, Rotman School of Management seminar, the RCFS Winter Conference, EARIE 2026 conference, and other conferences and workshops where earlier versions of this work were presented. We thank the Creative Destruction Lab (CDL) and Crunchbase for providing data access. *The authors did not receive any external funding for this research and declare no financial relationships or other potential conflicts of interest.*

JEL: L26, O38, O31, R11, R58, L53, G24
Keywords: Innovation policy; Accelerators; Entrepreneurship; Consideration sets; Information frictions; Regional economic development

1 Introduction

The allocation of entrepreneurial talent across industries is important for shaping the direction of innovation (Murphy et al., 1991; Ewens et al., 2018). Existing work emphasizes the economic fundamentals that guide this allocation, including expected returns, access to finance, and the availability of local skills and networks (Murphy et al., 1991; Saxenian, 1994; Nanda and Rhodes-Kropf, 2013). Yet favorable fundamentals cannot draw entrepreneurial talent into a specific sector unless they recognize it as something they could pursue. For example, a scientist planning to start a company may pursue a familiar application without considering an emerging industry in which the same expertise would be valuable. This gap between having relevant capabilities and recognizing where to apply them suggests an additional channel through which policy can shape innovation: by changing where potential founders direct their attention. In this paper, we shed light on how place-based innovation policies affect the allocation of startups across industries through the attention channel.

Separating the attention channel is empirically challenging because policies that direct founders' attention toward an industry can also change the expected payoff to entering it. Place-based innovation policies, including technology hubs, science parks, and accelerators, typically target specific sectors and provide additional resources such as capital, mentorship, and networks (Lerner, 2009; Bloom et al., 2019; Kline and Moretti, 2014). Because these changes occur together, it is hard to disentangle whether an increase in observed entrepreneurial activity within a target sector reflects more attractive opportunities or greater attention to those opportunities.

We exploit 23 city–industry launches by the Creative Destruction Lab (CDL), a global mentorship accelerator organized around science-based industry streams. Our setting offers a key empirical advantage for distinguishing these channels as a first step: we can distinguish ventures that seek program-related benefits from those that do not. We link detailed applicant and admission records from CDL to Crunchbase, identifying both admitted and failed applicants. Under the assumption that founders who enter an industry in anticipation of CDL's benefits all apply to the program, excluding the entire applicant pool removes entry motivated by ex-ante perception of increased payoffs coming from the program benefits. The remaining ventures are those that neither seek nor receive direct support, which serve as the major group we study. This approach provides an empirical basis for investigating the broader reach of policy: a targeted program may influence founders' industry choices through the attention it draws to a sector, extending its effects beyond the ventures it directly supports. While this removes the most direct payoff mechanism, it does not rule out

indirect payoff spillovers, which we account for in our structural model.

An additional empirical advantage of this program has been launched across multiple industries and cities over the course of 10 years, creating enough variation for us to study the allocation of new entrants across industries. We combine that variation with broad coverage of newly founded startups recorded in Crunchbase in the host cities, linked to founder backgrounds, financing activities, and further supplement founder connections using the LinkedIn data provided by Revelio lab.

We start our empirical analysis by examining how CDL launches shape entrepreneurs' entry into targeted sectors versus the rest. By deploying a Difference-in-Differences (DiD) framework with a rich set of fixed effects that control for city-level expansion, wide industry trends, and local industry composition, we first document a reallocation of entry towards the target sector. Specifically, the number of newly founded firms increased by 16% within the targeted sector. The effect remains statistically and economically meaningful at 12% when excluding all firms that are CDL applicants, which serves as supporting evidence of an attention channel. Furthermore, the attention channel works only when non-applicant entrants are exposed to the information of these launches. We examine one channel through which the information is spread: connections between non-applicant founding teams and CDL participants. We show that non-applicant entrants exhibit 25% greater schoolmate overlap with cohorts in newly launched streams than with cohorts in already-operating streams, even after accounting for similarities in academic credentials and professional experience. This finding is consistent with participants' personal networks transmitting information and directing other founders' attention toward targeted industries. Lastly, consistent with the fact that the CDL is a not-for-profit, mentoring-based program, we find no aggregate increase in total capital raised in targeted sectors. This evidence weighs against the view that launches induce entry through expectations of greater access to financing in targeted sectors. In fact, new entrants raise less total capital despite completing a similar number of financing rounds, suggesting greater competition for financing rather than an expansion in available capital.

Next, we focus on non-applicant entrants and examine which founders change their industry choices following a launch. We find that launches redirect science-trained, first-time founders toward targeted industries, with stronger responses in emerging industries than in established ones. These patterns are consistent with programs directing founders' attention toward applications of their expertise that they might otherwise give little consideration. We find no statistically significant increase in total entrepreneurial entry, suggesting that the principal response is a change in where entrepreneurs found companies.

Following the literature on limited consideration (Manzini and Mariotti, 2014; Brady

and Rehbeck, 2016), we propose a model of entrepreneurial industry choice that separates the payoff and consideration channels of accelerator launches. A launch can raise the expected payoff from entering its targeted industry, but it can also increase the weight that industry receives in a founder's consideration. To distinguish these channels, we use founders with prior founding experience as an informed benchmark. We use serial founders as an informed benchmark: their response captures changes in expected payoffs, while the additional response among comparable first-time founders is interpreted as a change in consideration. The model imposes two related identifying restrictions. First, a launch does not change serial founders' consideration weights. The motivation is that prior entrepreneurial experience gives serial founders greater knowledge of relevant industries and opportunities, making their choice set less responsive to the attention channel. Second, conditional on founder characteristics and the target industry's pre-launch local presence, the payoff channel spillover effect does not vary with founding experience. The response of serial founders therefore identifies the payoff channel, while the additional response of otherwise comparable first-time founders identifies the consideration channel. Our approach is close to identifying information frictions from informed agents (Bronnenberg et al., 2015; Handel and Kolstad, 2015) and of distinguishing informative from prestige effects of advertising by consumer experience (**?**). We allow the average industry choices to differ flexibly with founding experience, and the restriction applies only to responses to the launch conditional on observable characteristics.

The estimates show that the consideration channel redirects science-trained first-time founders toward targeted industries, particularly when those industries are locally emerging. The consideration effect declines as the local industry matures, which is reasonable as attention mechanism is expected to be weaker where an industry is already salient. The payoff channel, is small and statistically indistinguishable from zero, consistent with the fact that our sample is restricted to firms who do not directly benefit from the program, and the average payoff spillover to an average founder is reasonable to small. It is also consistent with our reduced-form findings on absence of any local financing increasing to a launch. In our model, we also abstract away from entry into entrepreneurship and expansion, which is supported by our reduced-form findings, and focus on the reallocation of the entrepreneurs across industries. This allows us to further control for the payoff spillover channel as our founders have the minimum resources to start a business. By conditioning on founding, we alleviate the concern that serial founders might be more likely to start another business that others. We also find that the consideration channel rises with science training and falls with PhD share and migrant share, while the payoff response shows the mirror image. This pattern is consistent with information channel redirecting founders whose primary constraint is limited information rather than limited resources, whereas payoff spillovers affecting founders who

are relatively well informed but resource-constrained

In our counterfactual analysis, we quantify how the allocation of entrepreneurial entry would differ in the absence of the consideration channel, holding the payoff channel and the population of founders fixed. This counterfactual measures the contribution of consideration to the associated reallocation of entrepreneurial activity across industries. The counterfactual estimates imply that consideration accounts for 22.8 percent of science-trained entry, and 30 percent of science-trained first-time entry into targeted industries. Note that total founding across all industries cannot change in the counterfactual where we fix the pool of potential founders. Our reduced-form city-level result that total founding does not rise after a launch support this design in our model.

Our paper contributes to four related literatures. First, we contribute to research on place-based innovation policy and the geography of entrepreneurship. Studies of entrepreneurial clusters emphasize how local networks and institutions shape entrepreneurial activity (Saxenian, 1994; Chatterji et al., 2014), while research on place-based policies examines whether interventions generate additional economic activity or redistribute it across locations (Ellison and Glaeser, 1997; Ellison et al., 2010; Glaeser et al., 2015; Kline and Moretti, 2014; Acs et al., 2016). We study the mechanisms through which sector-targeted interventions reshape the distribution of entrepreneurial entry across industries. Our findings identify an attention mechanism through which these policies redirect founders toward the targeted industry.

Second, we contribute to the literature on the allocation of talent and entrepreneurial entry. This literature studies how individuals respond to economic incentives, including local opportunities, financing conditions, and the costs of experimentation (Murphy et al., 1991; Adelino et al., 2017; Nanda and Rhodes-Kropf, 2013; Ewens et al., 2018; Balsmeier et al., 2025; Lee and Glennon, 2023; Khanna and Morales, 2025). A growing empirical literature also examines which industries entrepreneurs enter and how talent moves across industries (Foster et al., 2001; Decker et al., 2014). Prior findings show that entrepreneurs' sectoral choices are shaped by prior corporate experience and entrepreneurial spawning (Gompers et al., 2005; Chatterji, 2009; Muendler et al., 2012; Savage and Ziedonis, 2024), and that financial distress and industry decline can prompt skilled workers to pursue opportunities elsewhere (Babina, 2020). Other research highlights the role of early exposure and social connections: for example, childhood exposure to innovation in a particular technological field predicts later invention in that field (Bell et al., 2019). Together, these studies suggest that the allocation of talent depends both on the returns to different activities and on individuals' awareness of how they might participate in them. Our findings extend this literature by

suggesting that policy-induced attention to a specific sector can shape the allocation of entrepreneurial talent across industries.

We also contribute to research on startup accelerators by documenting their effects beyond participating ventures. Existing studies examine how accelerator services, learning, and advice affect venture performance (Gonzalez-Uribe and Leatherbee, 2018; Hallen et al., 2020; Heydari-Nejad, 2024; Sariri, 2025; Lyons and Zhang, 2018), while Fehder and Hochberg (2014) studies regional financing effects extending beyond accelerator portfolios. We show that accelerators can function as signaling tools that reshape the industry choices of founders outside the program, inducing entrepreneurial entry beyond their direct participants.

Finally, our work also relates to the literature on limited consideration and information frictions in choice. Manzini and Mariotti (2014) and Brady and Rehbeck (2016) distinguish the alternatives available to an individual from those that enter the individual's consideration set. Empirical studies show how advertising and information influence consideration and choice (Goeree, 2008; Honka et al., 2017). We apply this logic to entrepreneurs choosing an industry. Our structural model allows launches to change both industry consideration and expected payoffs. To distinguish these channels, we draw on the use of experience and expertise to study information effects (**?**Ackerberg, 2003; Bronnenberg et al., 2015), using differences between first-time and experienced founders. This brings the consideration-set approach to the allocation of entrepreneurial entry and provides a framework for quantifying the attention component of targeted innovation policy.

2 Data and Sample

2.1 Institutional Setting: Creative Destruction Lab

Creative Destruction Lab (CDL) is a global startup accelerator program founded at the University of Toronto's Rotman School of Management in 2012. CDL operates as a non-profit mentoring program specifically designed for early-stage, science-based ventures. The program's structure makes it particularly well suited for studying how a targeted program shapes local entrepreneurial activity outside its own portfolio. Between 2012 and 2021, CDL expanded from its original Toronto location to 10 sites across Canada, the United States, Europe, and Australia, launching 27 specialized technology streams. CDL launches programs that target a specific industry. A full description of the CDL dataset is provided in Sariri et al. (2025)

More specifically, the program launches sector-specific "streams", such as artificial intelligence, health, quantum computing, or fintech, in different cities at different times. Each stream is hosted by a business school in the target city and focuses on startups in a particular technological domain. This staggered rollout across both geographic and industry dimensions provides the quasi-experimental variation necessary to identify the effect of a targeted launch on the composition of local entry.

A *launch* in our analysis is therefore a city-industry-year event, for example, CDL opens a quantum stream in Toronto in 2017, an oceans stream in Halifax in 2019. Launched industries span the full range of local maturity, from industries with essentially very few local firms at the time of the launch to industries with several hundred existing local firms.

2.2 Data Sources

CDL Program Data

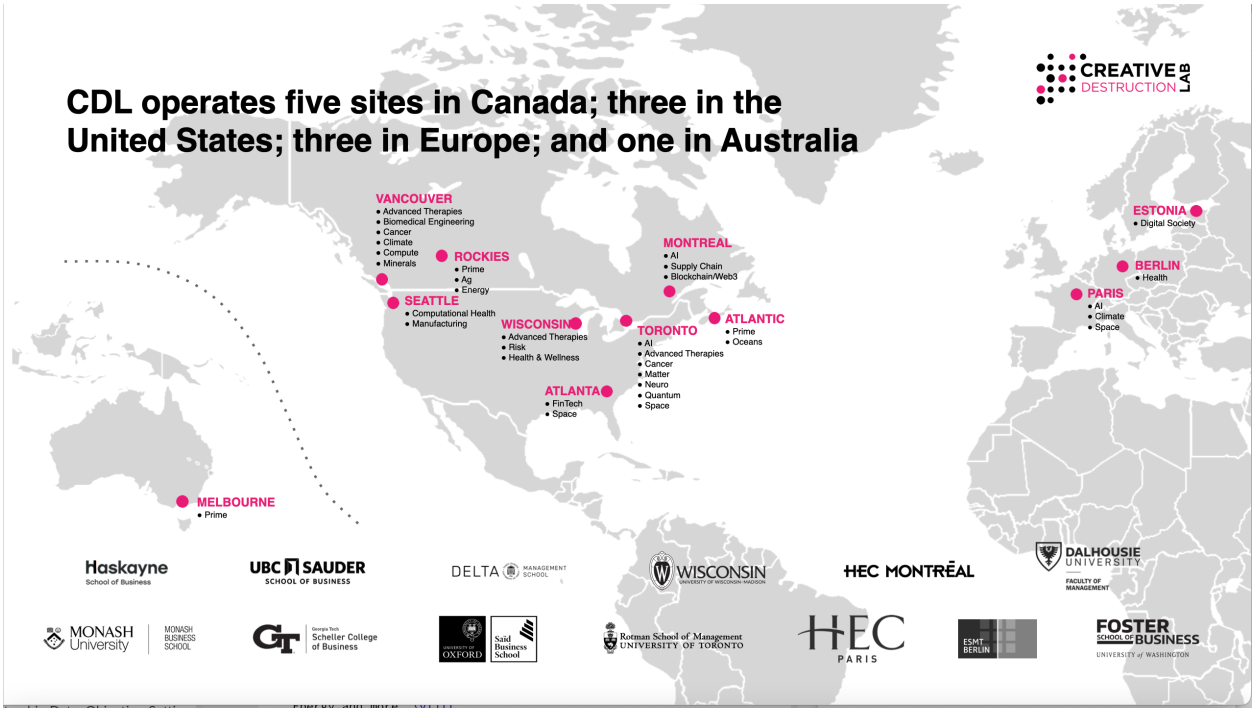
We obtain detailed information on CDL stream launches from the program's administrative records. This dataset includes the timing of each stream launch, the host city and business school, the industry focus of each stream, applicant ventures and also accepted one. For each stream launch we record: launch timing (the first year a stream began accepting ventures in a focal city), geographic information (host city, country, and associated business school) and industry classification (the focus of each stream).

To ensure accurate estimation of the effect of a targeted launch, we exclude two launches that lasted only one year before being discontinued: the *Recovery* stream, aimed at COVID-related recovery solutions, and the *Prime* stream in New York City, which ran for one year. We also exclude the general-purpose *Prime* streams, which target technology ventures without an industry focus, because our object of study is the effect of *industry-specific* targeting. These restrictions leave 23 industry-stream launches across ten cities between 2012 and 2022.

Figure 1 shows the geographic distribution of CDL sites and partner institutions.[1] Figure A.1 illustrates the staggered introduction of CDL programs across cities and sectors. Figure A.2 displays the average cohort size by location, revealing a steady expansion from roughly 20 admitted startups per cohort to more than 60 in later years. Panel (a) of Table 1 reports summary statistics for the CDL program.

[1]The map reflects the network as of 2025, including newly established sites in Australia, Berlin, and Estonia, which are not part of our sample, which ends in 2022.

Figure 1: **CDL Map.** This map shows the locations of CDL sites and the industry streams operating at each site in 2025.



Crunchbase Data

We complement the CDL data with Crunchbase, a comprehensive database of startup companies and their funding. From Crunchbase we extract firm-level data including founding date, geographic location, industry classification, and financing round details, and we map Crunchbase industry categories to standardized classifications that align with CDL's stream categories. This mapping identifies which startups would be eligible for specific CDL streams based on their technological focus. We restrict the sample to firm founding and financing activity from 2000 to 2022 to ensure sufficient pre- and post-treatment periods. Table 1, Panel (b), provides summary statistics for ventures based in the focal cities.

We observe the full list of CDL applicants and admitted ventures and match them to Crunchbase. We then aggregate the firm-level sample to a city-industry-year panel of entry counts, balanced over the ten cities, seventeen industries, and 2000-2022, with zeros retained.[2] Table 1, Panels (c) and (d), provides summary statistics for all cities and the focal cities, respectively. Both the number of new firms and aggregate financing are extremely right-skewed, with a few cities generating a large number of entries and attracting capital

[2]Results are not materially affected by using an unbalanced panel that drops city-industry-year cells with no activity.

inflows; this skew is one reason we estimate entry responses by Poisson pseudo-maximum likelihood (Santos Silva and Tenreyro, 2006; Chen and Roth, 2024).

Table 1: **Summary Statistics.** Panel (a) reports summary statistics for the CDL program at the city–stream and city levels. Panel (b) reports city–industry–year panel statistics for all cities and for the focal cities.

	count	mean	sd	p5	p50	p95
Number of Accepted Ventures (Cohort-Site-Stream)	94	24.766	10.264	18.000	24.000	44.000
Number of Accepted Ventures (Cohort-Site)	41	56.780	46.300	19.000	47.000	170.000
Number of Applicants (Cohort-Site)	42	175.905	197.756	19.000	114.500	422.000
Share of Applicants from Local (Cohort-Site)	41	0.380	0.232	0.083	0.333	0.778

	count	mean	sd	p5	p50	p95
Num_New_Firm	2842	19.979	37.790	1	6	94
Founding_Year	2842	2011.309	6.461	2001	2012	2021
Agg_Financing_MilUSD	1974	152.834	433.869	0.286	21.410	741.450
Announced_Year	1974	2013.926	5.971	2002	2015	2022

2.3 Founder Characteristics and Types

We examine two main founder characteristics to proxy for informedness and relevance of accelerator launches for each founding team. Prior founding experience proxies *informedness*, and experienced founders serve as a full-information benchmark, in the spirit of expert-consumer comparisons (Bronnenberg et al., 2015) and of separating informative from prestige effects of advertising by consumer experience (**?**). Scientific training proxies for *relevance* as CDL targets science-based industries and ventures. The motivation is that a team whose members have founded before has already surveyed the local industry space, has raised money, met investors, and seen which opportunities exist nearby. One would expect that a shift in the payoff to entering the targeted industry moves every team for whom the improved fundamentals matter, depending on their observable characteristics. A shift in information can move only teams that did not already know, and among those, only the ones for whom the news is relevant. The payoff channel therefore predicts a broad response depending on observables, while the consideration channel predicts a response concentrated among science-trained first-time founders, that also declines with the pre-launch salience of the industry in that region.

To measure these proxies, we link each firm to its founders' biographies in Crunchbase, degrees and fields of study, prior founding roles, prior employment, and location histories, and construct team-level characteristics. Appendix A.6 documents the construction in detail.

Two main characteristics define the founder types that we later use:

- *Science-trained*: the founding team includes at least one founder holding a degree in science or engineering, where degree subjects are classified by keyword matching into science and engineering fields (natural sciences, all engineering disciplines, computing and information sciences, mathematics and statistics, and medical and health sciences) versus business fields (including economics) and other.
- *First-time*: no member of the founding team has prioir founding experience, as identified from founders' job histories before the focal firm's founding date.

Table 2 reports the distribution.

Table 2: **Founder Types.** Distribution of non-applicant founding teams across the four founder types.

	Reduced form		Structural model	
	Teams	Share	Teams	Share
Science-trained, first-time	1,947	33.0%	1,841	33.0%
Non-science, first-time	2,701	45.8%	2,539	45.6%
Science-trained, serial	559	9.5%	540	9.7%
Non-science, serial	685	11.6%	653	11.7%
Total	5,892	100.0%	5,573	100.0%

Notes. Non-applicant founding teams in the ten CDL cities, 2000–2022, in the seventeen panel industries. A team is *science-trained* if at least one founder holds a degree in science or engineering, and *first-time* if no member has held a prior founding role. Experience proxies informedness and science training proxies relevance.

3 Reduced-form evidence

We exploit the staggered rollout of the CDL program across multiple cities and industries, and study the impact of a new stream launch on the composition of entrepreneurial entry into different industries. The unit of analysis in our main panel is city-industry-year.

3.1 Reallocation of entry toward newly launched streams

We first compare entry in treated city-industries with entry in not-yet-treated and never-treated industries in the host cities, deploying a Difference-in-Differences (DiD) framework with a rich set of fixed effects that control for city-level expansion (city × year), wide industry trends (industry × year), and local industry composition (city × industry). Specifically, we estimate the following specification at the city-industry-year level:

$$Y_{cit} = \beta \cdot \text{TargetedIndustry}_i \times \text{Post}_{cit} + \gamma_{ci} + \tau_{ct} + \eta_{it} + \varepsilon_{cit} \tag{1}$$

where Y_{cit} represents our outcome variables for city c, industry i, in year t. I restrict the estimation window to two years before and two years after each new launch. The coefficient β captures the change in outcomes for treated sectors relative to untreated sectors within a city after the launch of the CDL stream.

Table 3 reports the Poisson pseudo-maximum likelihood estimators. The dependent variable is the count of newly founded firms. Across all specifications, β is positive and statistically significant, indicating that a stream launch is associated with a meaningful increase in entry in the targeted industry relative to the rest.

One empirical evidence that motivates the attention channel comes from the fact that main results survive the sample restrictions. Columns (1) to (3) progressively exclude accepted CDL participants and then all CDL applicants, retaining only firms whose founders had no direct involvement with the program. The coefficient remains stable across these restrictions while gradually decreasing in magnitude, from 30% in the full sample to 23% when excluding all CDL applicants. This suggests that the reallocation of entry is not a mechanical artifact of program participation but extends to founders who never sought CDL's benefits, consistent with an attention and information channel operating beyond the program's direct reach.

Columns (4) to (6) further include city × industry fixed effects, which absorb any time-invariant differences in local industry composition, such as economic fundamentals that support the entrepreneurial entry of certain industries. The estimated effects halved while estimators remain statistically significant across all three sample restrictions. Columns (7) to (9) address the concern that treated cities may be systematically different from the rest of the country, for instance, if CDL deliberately targets cities with already-growing technology sectors, the estimated reallocation could partly reflect a pre-existing trajectory rather than the causal effect of the launch. To address this, we further include the top matched never-treated city for each treated city as a control sample, selected on country and prior-year

Table 3: **Reallocation of New Entrants Towards the CDL Newly Launched Stream.** This table reports the estimated effects of CDL stream launches on new firm entry at the newly launched stream as compared to the untreated ones among treated cities in columns (1) through (6), and adds the top matched never-treated cities as control groups in columns (7) through (9). We retain the top one matched city for each treated city inside the sample, based on country and prior-year industry breadth. The dependent variable across all models is the count of newly founded firms, and we employ a Poisson regression model with multiple levels of fixed effects. Columns report estimates for all entrants, the sample excluding accepted CDL participants, and the sample excluding CDL applicants. All specifications include city × year fixed effects to control for city-level expansion and isolate the reallocation effect, and industry × year fixed effects to control for wide industry trends, with columns (4) to (9) further adding city × industry fixed effects to control for local industry composition. Standard errors are clustered by city × industry.

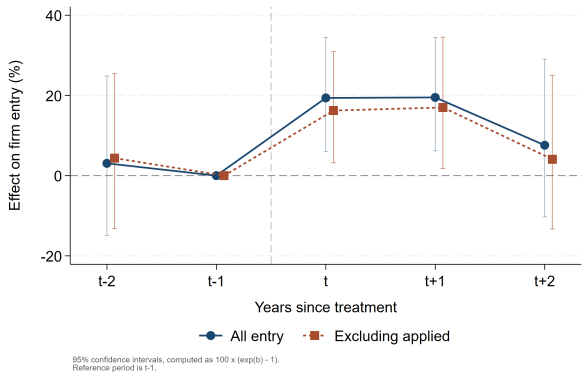
	Outcome Variable: Number of Newly Founded Firms								
	Treated cities						Matched nearest never treated cities		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated × Post	0.262***	0.233**	0.206**	0.133**	0.114**	0.100*	0.146**	0.127**	0.113*
	(0.100)	(0.100)	(0.100)	(0.058)	(0.056)	(0.059)	(0.065)	(0.064)	(0.068)
Implied % effect	29.9	26.3	22.9	14.3	12.0	10.5	15.7	13.5	11.9
Sample	Full	Excl. Acc.	Excl. App.	Full	Excl. Acc.	Excl. App.	Full	Excl. Acc.	Excl. App.
City × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
City × Industry FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2442	2442	2442	2439	2439	2439	4743	4743	4743
Pseudo R-squared	0.892	0.892	0.892	0.903	0.903	0.902	0.880	0.880	0.880

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

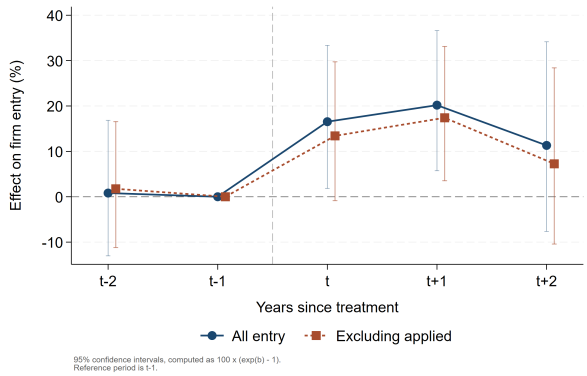
industry breadth. If the results in columns (1) through (6) were driven by idiosyncratic growth in treated cities, the coefficients should attenuate substantially once these matched cities serve as the counterfactual. Instead, the estimates remain stable with similar magnitude. The consistency suggests that the reallocation toward targeted industries reflects the stream launch itself rather than a coincident or pre-existing trend in cities selected for the program.

The identifying assumption underlying our DiD design is that treated and control industries within the same city would have followed parallel trends in entry absent the CDL stream launch. We assess this assumption using dynamic event-study specifications that trace the path of entry in treated relative to control industries in the years surrounding a launch. Figure 2 presents the results. The pre-period coefficients are small in magnitude and statistically indistinguishable from zero, providing no evidence of differential pre-trends between treated and control industries prior to the launch. Entry rises sharply in the post-period, consistent with the estimates reported in Table 3.

Figure 2: **Event-Study Estimates of Firm Entry Around New CDL Stream Launch.** This figure shows the dynamics of firm entry and the amount of financing attracted around a new stream launch by CDL, based on the following specification: θ_t of equation $Y_{i,c,t} = \sum_{t \neq -1} \theta_t Periods_t * Launch_{i,c,t} + FEs + \epsilon_{i,c,t}$, using Poisson regression. The specification includes city × industry FEs, city × year, and industry × year FEs. The year before the event is omitted as the base year. Standard errors are clustered by city × industry.



(a) Treated Cities



(b) Add Top Matched Never-Treated Cities

Appendix Figure A.5 presents additional robustness event-study graphs using alternative estimators, including the Callaway and Sant'Anna (2021) and Sun and Abraham (2021) estimators. To further examine the reallocation our main results suggest, Appendix Figure A.3 plots event-study estimates separately for targeted and untargeted industries, each benchmarked against the same industries' trend in never-treated cities. Following a CDL launch, entry in host cities rises in targeted industries and declines in untargeted industries relative to their respective control group. This divergence is consistent with launches

redirecting entrepreneurial activity toward targeted sectors.

3.2 Exposure to CDL launches and Non-Applicant Entry

The increase in new firm formation among non-applicants following a new launch raises the question of how these launches affect founders who neither seek nor receive direct program support. One possibility is that participants transmit information about a newly targeted industry through their personal networks, drawing other founders' attention to opportunities to apply their expertise. We examine this possibility by asking whether non-applicant entrants are disproportionately connected to cohorts in newly launched streams.

Table 4 compares entrants' schoolmate overlap with cohorts in newly launched streams and cohorts in already-operating streams. Schoolmate connections are defined by attending the same school in overlapping years. Cohorts in already-launched ongoing streams provide a benchmark for the connections that ordinarily exist between CDL participants and the surrounding entrepreneurial community.[3] The comparison therefore asks whether connections to the inaugural cohort are stronger than this baseline association with CDL. And if so, the evidence would be consistent with a new launch drawing non-applicant founders from participants' existing networks into the targeted industry who might otherwise not have considered entering it. All specifications include program, founding-year, and firm city $\times$ industry fixed effects.

Column (1) shows that non-applicant entrants exhibit greater schoolmate overlap with cohorts in newly launched streams. Columns (2) and (3) progressively control for similarity in academic credentials and professional experience to account for the possibility that founders and participants share training or industry exposure that makes the targeted industry attractive to both, regardless of any information transmitted through personal ties. The coefficient on schoolmate overlap remains positive and significant after these controls, suggesting that the pattern is not fully explained by observed similarities in the backgrounds of entrants and participants. This evidence is consistent with personal networks exposing founders to newly targeted industries and drawing attention to applications of their expertise. This evidence does not rule out the network itself facilitating access to resources. We therefore develop a model in which launches can affect both the consideration weight founders place on an industry and the expected payoff from entering it. Under explicit identifying assumptions, the model separates these channels.

[3]For example, founders may share educational backgrounds with participants because CDL draws from particular universities, because certain skills are relevant to its industries, or because local entrepreneurial networks are closely connected. These relationships can generate overlap even without a new launch.

Table 4: **Founding Team Overlap Between Non-Applicant Entrants and New CDL Cohorts.** This table examines the possible channel through which CDL launches can induce non-applicant entrants, who don't expect to receive the program's benefits ex-ante. Specifically, it asks whether these entrants are drawn disproportionately from the personal and professional network of the focal cohort, who are more likely to be exposed to these events. We interpret a positive premium as evidence of an information and attention channel: proximity to participants raises peers' attention to the opportunity. For each entrant fir-stream-year, we construct *founder-cohort overlap*, the share of the combined set of the founder's and the cohort's affiliations that the two share, bounded between zero and one (a Jaccard proximity index). For the purpose of display, the index is multiplied by 100 so that coefficients and the outcome mean are easier to read. Overlap is computed separately for each channel. The dependent variable is overlap through the direct-connection channel, former schoolmates, defined as attending at the same school in overlapping years. To isolate the spillover effect of a launch, treated observations measure a founder's overlap with a *newly launched* stream, while control observations measure overlap with *already-operating* streams. If an AI stream launched in 2015 and a Quantum stream launches in 2017, the 2017 treatment evaluates overlap with the new Quantum cohort and the control evaluates overlap with the ongoing AI cohort. Because the control group represents baseline CDL operations, the coefficient captures a "first-launch premium." Columns (2) and (3) progressively add *background overlap controls*: similarity between the entrant team and the cohort in academic credentials (degree level, field of study) and in professional experience (employer industry, kind of role). These absorb the tendency to enter simply because a founder holds the background the technology requires, which is the similarity a payoff-based explanation would predict. The premium that survives is therefore attributable to personal connection rather than to shared qualifications. All specifications include program, founding-year, and firm city × industry fixed effects. Standard errors are clustered by firm city × industry.

Dependent Variable: Founding-team overlap between non-applicant entrant and CDL cohort			
	Channel: former schoolmates		
	(1)	(2)	(3)
Treated × Post	0.114**	0.124**	0.095**
	(0.056)	(0.053)	(0.046)
Competing overlap channels			
Same degree level		0.013***	0.011***
		(0.002)	(0.002)
Same field of study		0.005**	0.003
		(0.002)	(0.002)
Same industry employer			0.004*
			(0.002)
Same kind of role			0.020***
			(0.004)
Program FE	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes
Firm City × Industry FE	Yes	Yes	Yes
Mean of DV	0.367	0.367	0.367
Observations	20,156	20,156	20,156
Adjusted R^2	0.029	0.067	0.077

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To further examine the entry induced by the payoff channel, we look at whether improved access to financing can explain the entry response. Appendix Tables A.2 and A.3 show that, although more young firms receive financing following a launch, there is no statistically significant increase in total capital raised in targeted industries across firm age groups, suggesting stronger competition for capital. Consistent with the aggregate evidence, post-launch entrants on average raise less capital.[4] These findings provide limited support that expansion in financing explains increased entry. Because financing is observed ex post, however, they do not rule out improved expectations at entry or other payoff spillovers, which we allow for in the model.

3.3 Accelerator Launches Redirect Science-Trained First-Timers

We exploit the staggered rollout of the CDL program across multiple cities and industries, and study the impact on the composition of entrepreneurial entry into different industries in those targeted cities. The unit of analysis in our main panel is the city-industry-year. In our main specification, we compare entry in treated city-industries with entry in not-yet-treated and never-treated industries in the same city. For the dependent variable, we count entry by founder type $g \in \{\text{science first-timers, non-science first-timers, serial founders}\}$ and estimate a Poisson pseudo-maximum-likelihood regression on the balanced city×industry×year×type panel:

$$\mathbb{E}\big[Y_{cstg}\big] = \exp\Big\{ \sum_{g} \big(\tau_{0g} + \tau_{1g}\, \tilde{A}_{cs}\big)\, D_{cst}\, \mathbf{1}\{g\} \;+\; \gamma_{csg} + \delta_{ctg} + \mu_{stg} \Big\}, \tag{2}$$

Where Y_{cstg} is the number of new firms of type g in city c, industry s, year t; D_{cst} is a post-launch dummy, and $\tilde{A}_{cs}$ is the local industry size before launch defined in Section 2. We also include city-industry, city-year, and industry-year fixed effects, all interacted with type of the founder. As all the independent variables and fixed effects are type-specific, the coefficients are similar to running separate regressions by type, while the joint covariance matrix allows for tests of cross-type differences of the estimated coefficients.

City-Industry fixed effects (γ_{csg}) controls for permanent local specialization. City-year fixed effects (δ_{ctg}) controls for everything happening to a target city in a year, so that τ measures within-city reallocation not the city-level shocks to all industries. Industry-year effects (μ_{stg}) control for the global industry waves, for example global AI boom. A city-industry-year effect is not included as it would absorb the treatment D_{cst}. Note that all fixed

[4] Please refer to Appendix Section A.2 for more detailed description.

effects are interacted with founder types, controlling for any of these confounders affecting each type differently.

We estimate (2) by PPML because entry counts are small and frequently zero in the emerging industry-cities that identify the maturity coefficient[5].

Treatment is at the city-industry-year level and shocks might be correlated within cities across industries and years, so we cluster by city. With ten cities, cluster-robust standard errors over-reject, we therefore report p-values from the Webb six-point wild cluster score bootstrap with 9,999 replications (Cameron et al., 2008; Webb, 2023). [6].

Table 5 reports estimates of equation (2). Each column is for each founder types. First row coefficients are the accelerator launch effect for a launch of average pre-existing local industry size(τ_{0g}), and the second row coefficients are the change in pre-launch local industry size (τ_{1g}). we also report the implied effect in an emerging cell, where the prior industry size is at $\tilde{A}$ = -2, the 10th percentile of launches (for example Toronto Quantum, Oxford Climate). We also report the statistical significance of cross-type differences in the estimated coefficients. We find that accelerator launches are associated with a large increase in entry into the targeted industry, and the effect declines with the size of the pre-existing local industry.

[5]Poisson quasi-likelihood is consistent for the conditional mean under weak assumptions (Santos Silva and Tenreyro, 2006)

[6]Standard errors in parentheses are cluster-robust; significance stars come from the bootstrap, so they might differ from the conventional t-statistic numbers. For the main industry composition result we additionally report design-based randomization inference, reassigning each launch to another host city that never launched that industry while holding the industry and year fixed, which fixes the wave structure to tests the null of no local launch effect

Table 5: **Entry into the Targeted Industry, by Founder Type.**

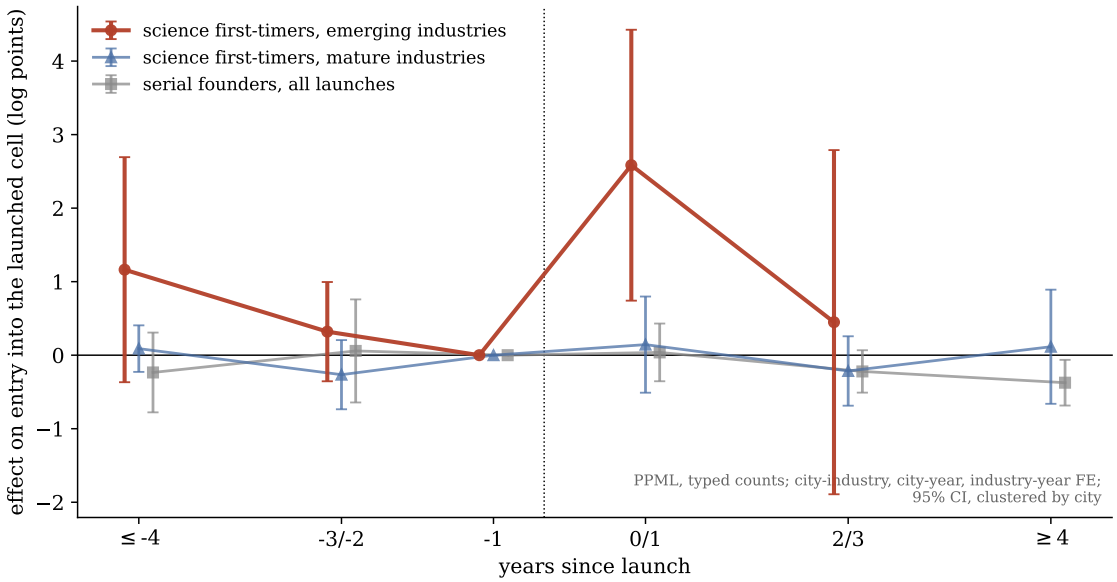
	Science first-timers	Non-science first-timers	Serial founders
Launch effect, mean launched cell	$+1.00^{**}$	-0.26	-0.33
	(0.37)	(0.33)	(0.42)
× pre-launch industry size	-0.45^{*}	$+0.23$	$+0.19$
	(0.24)	(0.17)	(0.21)
Implied effect, emerging cell	$+1.90^{**}$		
	(0.82)		
Cross-type contrasts (joint covariance): science − serial = $+1.33^{**}$ (0.52); science − non-science = $+1.26^{**}$ (0.55)			

Notes. Estimates of equation (2): a single stacked PPML regression on the balanced city×industry×year×type panel, fully interacted by founder type, with city×industry, city×year, and industry×year fixed effects by type. The sample is the 5,892 non-applicant founding teams for which a degree field is observed for at least one founder, so that non-science is a measured category rather than an absence of education data. The launch effect is evaluated at the mean launched cell, centering the log pre-launch local industry stock over the 23 launched city-industries; the maturity gradient is the interaction with that centered stock; the emerging-cell effect evaluates the two at $\tilde{A} = -2$, equivalent to about six existing local firms and within the observed launch support of two to 543. Cross-type contrasts are linear combinations of the stacked coefficients, computed from the joint covariance. Standard errors in parentheses are clustered by city (10 clusters), and significance is assessed against t_{G-1} critical values with $G = 10$: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

First column shows that for science-trained first-time founders, a launch raises entry into the targeted industry by 1.00 log points at the average launched cell, about 2.7-fold. This effect is decreasing in pre-launch industry size. Second and third columns show that non-science first-timers and also serial founders do not respond. The cross-type differences tests are large and significant: science-trained first-timers respond +1.33 log points more than serial founders and +1.26 more than non-science first-timers.

Figure 3 reports the event-study version of the same specification by splitting launches at the median pre-launch stock. The PPM regression aggregates at time level, this figure aggregates the effect at industry size level and shows the time trend. The figure shows that the response of science-trained first-timers in emerging industries is mostly in the first two years after the launch.

Figure 3: **Dynamics of the Typed Entry Response.** This figure reports the event-study version of equation (2), estimated by PPML on the balanced city×industry×year×type panel with city×industry, city×year, and industry×year fixed effects, all interacted with founder type. Event time is measured in two-year bins relative to the launch year, with the year before the launch omitted as the base period; launches are split at the median pre-launch local industry stock. Whiskers are 95 percent confidence intervals, clustered by city.



Taken together, our reduced-form analysis shows that entry by science-trained first-timers into launched industries rises relative to non-treated industries in those cities. We find no evidence of city-wide expansion at the aggregate level. These observations are therefore consistent with the inflow into launched industries coming from founders who would have started a business in other industries, suggesting that these accelerator launches have reallocated entrepreneurs across industries. These observations supports our assumptions in the structural model and our counterfactual analysis.

4 A Model of entrepreneurial Industry Choice

The reduced-form evidence shows that an accelerator launch redirects science-trained first-timers toward the targeted industry, while founders with prior founding experience do not respond. In this section, we develop a founder-level model of industry choice with limited consideration, that separates entrepreneurial entry incentives through two channels: *payoff channel*, where an accelerator launch raises the expected payoff to founding in the targeted industry, and *consideration channel* where the accelerator launch expands the set of indus-

tries potential founders consider. A payoff shift affects every payoff-constrained founder, but a consideration shift affects only uninformed founders.

4.1 Setup

Founding team i in city c and year t chooses an industry to start a business. Teams are described by by experience $e_i \in \{0, 1\}$, equal to one if any member in the team has founded before (serial entrepreneur on the team), and by a set of observable characteristics x_i including an indicator for a science degree on the team, the team's PhD share, its migrant founder share, and log team size.

We model industry choice in two conceptual layers. In the first layer, each founder assigns a positive consideration weight to each industry. This weight captures how prominently the industry enters the founder's consideration set. In the second layer, the founder chooses an industry given these consideration weights and the expected payoffs from the available alternatives. we let the attention weight a first-time founder places on industry s be:

$$\phi_{ist} \;=\; \begin{cases} 1, & e_i = 1 \quad \text{(serial founders)}, \\ \bar{\phi}_{cs} \, \exp\left\{\lambda(x_i, \tilde{A}_{cs}) \, D_{cst}\right\} & e_i = 0 \quad \text{(first-time founders)}, \end{cases} \tag{3}$$

Where $\bar{\phi}_{cs}$ is baseline awareness about industry s in city c that is common for everyone, D_{cst} is an indicator showing an active accelerator program , it becomes one in the year that CDL launches a program in city-industry cs and remains one for the following years. $\tilde{A}_{cs}$ is the industry's pre-launch local size relative to the average size of all launched cells (the log stock of accumulated existing local firms, centered at the average launched cell. we call $\tilde{A} < 0$ an emerging industry).

The function $\lambda(\cdot)$ is the *consideration technology* and captures the launch's proportional effect on the weight that a first-timer places on industry s. $\lambda(\cdot)$ captures the launch-induced change in the consideration weight among first-time founders. This effect depends on the pre-launch industry size as we expect the more salient industries in a city to already be on the menu of a potential founder. It also depends on founder characteristics, because a consideration boost to an industry is more informative for some founders than others, and permanent differences in which industries a given type of founder already considers are absorbed by the industry-by-characteristic controls in the second layer. We assume that launches do not change the consideration weights of serial founders, who serve as the informed benchmark. This assumption does not require serial founders to consider every

industry equally or to possess complete information before the launch.

In the second layer, founder chooses among industries, with consideration entering as a weight on each alternative.

Utility is $u_{ist} = v_{ist} + \varepsilon_{ist}$ with

$$v_{ist} = \delta_{cs}(x_i, e_i) + \tau(x_i, \tilde{A}_{cs})\, D_{cst},$$

ε i.i.d. Type-I extreme value, and choice probabilities

$$P_{ist} \;=\; \frac{\phi_{ist}\, e^{v_{ist}}}{\sum_k \phi_{ikt}\, e^{v_{ikt}}}.$$

[7]

$\delta_{cs}(x_i, e_i)$ captures many aspects of sorting: a team fixed effect, a city-industry fixed effect, and interactions of industry indicators with each of the elements of (x_i, e_i). These interactions of industries with characteristics control for the time-invariant taste of science-trained (or any other characteristic) teams to found in science-based (or any other specific) industries. The choice set could also include an outside option, not founding. Since we include a team fixed effect, the outside option is differenced out and the estimates do not depend on this assumption. Outside option matters only in the counterfactual, where removing an industry from someone's menu could in principle push her out of entrepreneurship. Since our reduced form analysis show that launches do not change city-wide founding , our counterfactual holds the population of founders fixed and allows for reallocation across industries.

$\tau(\cdot)$ is the launch's payoff effect that is common to all founders, conditional on their observable characteristics and conditional on pre-launch industry condition in that city. Note that we restrict our sample to teams that never apply to or join the program, so no founder receives direct program services. this alleviates the concerns about treatments inside the program and who benefits more from those services. We still include the payoff channel in the model to account for any residual externalities coming from actual improvement in the expected payoff of founding in the targeted industry.

for simplification, we consider both channels as linear functions:

$$\tau(x, \tilde{A}) = \pi_0 + \pi' x + \pi_A \tilde{A}, \qquad \lambda(x, \tilde{A}) = \lambda_0 + \lambda' x + \lambda_A \tilde{A}, \tag{4}$$

[7] This attention-weighted form is our current specification, with menus realized as independent draws, exact choice probabilities require summing over all 2^{J-1} possible menus (Goeree, 2008).

We center all the characteristics x in our sample, so the estimated coefficients are the responses of the average founding team, evaluated at the average pre-launch industry size cell.

The response of serial founders identifies the payoff component, while the excess response of first-time founders identifies the consideration component, under the assumption that launch-induced payoff effects do not vary with founding experience conditional on founder characteristics and pre-launch industry conditions.

4.2 Estimation

With consideration entering as a weight and extreme-value taste shocks, choice probabilities are $P_{ist} = \phi_{ist} e^{v_{ist}} \big/ \sum_k \phi_{ikt} e^{v_{ikt}}$. So, logarithm of choice probabilities are additive in $\ln \phi$ and v. The launch shifts i's log odds of choosing the targeted industry by

$$\tau(x_i, \tilde{A}_{cs}) \; + \; \mathbf{1}\{e_i = 0\}\, \lambda(x_i, \tilde{A}_{cs}), \tag{5}$$

this model is a fixed-effects conditional logit in which λ captures the launch-induced change in the log attention weight, $\lambda = \Delta \ln \phi$, and e^{λ} is the factor by which a launch multiplies the weight a first-timer places on the targeted industry. In the estimation, $\bar{\phi}_{cs}$ enters only for first-time founders and is absorbed by city-industry fixed effects interacted with founding experience, so only launch-induced changes are identified. The team fixed effect absorbs everything constant across a team's alternatives, including its city-year, because each choice is made within a city-year. Our estimation is by fixed-effects Poisson pseudo-maximum likelihood on the team-industry panel (one chosen industry per team), which is equivalent to the conditional-logit likelihood with high dimensional fixed effects. standard errors are clustered at the city level, and because there are ten cities, p-values are calculated using the Webb six-point wild cluster bootstrap with 9,999 draws (Cameron et al., 2008; Webb, 2023).

4.3 Identification

Two restrictions separate the channels.

Serial founders as attention-invariant benchmark who know their comparative advantage. We assume that the experience of having founded reveals the local industry space (a weaker assumption where serial founders know their best fit even if they don't consider

all options does not change our results). Their response identifies the payoff channel in the second layer, and the first-timer excess response identifies consideration channel. This logic of experienced agents as informed benchmark has been used to identify information frictions. For example expert consumers reveal full-information demand (Bronnenberg et al., 2015; Handel and Kolstad, 2015). Closest to our work, advertising that affects inexperienced consumers but not experienced ones separates the informativeness of the treatment from the persuasiveness (**?**Ackerberg, 2003). Also if launches expanded serial founders menus as well, their response would be largest where baseline familiarity is lowest. Estimated serial responses are near zero everywhere, including emerging cells.

The payoff channel does not depend on experience conditional on characteristics and industry condition, $\tau(x, e{=}0) = \tau(x, e{=}1)$. We assume that conditional on all the sorting controls, first timer and serial founders experience the same payoff from launch. This will be violated if for example payoff externalities conditional on all our sorting controls differ among first timers versus serial founders.

Because choice probabilities depend only on the product $\phi_{ist}e^{v_{ist}}$, consideration and preference are observationally equivalent in choice data which is common problem in the literature. Our decomposition does not depend on separating the level of ϕ_{ist} from the level of v_{ist}. What we identify is the launch-induced effect, split into a component common to all founders and a component affecting only first-time founders.

4.4 Results

The estimation sample is the 5,573 founding teams with complete founder characteristics, who never apply to or join the accelerator. We have seventeen industries in our sample. Table 6 reports the estimated parameters. Results are as follows:

Table 6: **Two-Channel Estimates.**

	Payoff $\tau(x)$	Consideration $\lambda(x)$
Level (mean founder)	$+0.033$ (0.395)	$+0.593^{*}$ (0.289)
Science degree	-0.655^{***} (0.133)	$+0.792^{***}$ (0.151)
PhD share	$+0.730^{**}$ (0.296)	-1.177^{***} (0.215)
Migrant share	$+0.771^{***}$ (0.175)	-0.615^{***} (0.179)
Log team size	$+0.045$ (0.191)	-0.235 (0.236)
$\times$ pre-launch industry size	$+0.090$ (0.143)	-0.241^{*} (0.111)
Founding teams	5,573	
Founder$\times$industry obs.	86,953	

Notes. Fixed-effects conditional logit (estimated as fixed-effects Poisson) on 5,573 non-applicant founding teams $\times$ 17 industry alternatives; team fixed effects, city$\times$industry fixed effects, and industry indicators interacted with each team characteristic and with experience (sorting controls). Characteristics are mean-centered and $\tilde{A}$ is centered at the mean launched cell, so the level row reports the response of a founding team with average characteristics at a typical launch. Standard errors in parentheses are clustered by city; stars are from the Webb wild cluster bootstrap (9,999 draws): * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$..

First, for an average founder team, the average payoff channel is not significantly different from zero. The estimated payoff level is $+0.03$ and insignificant. This is consistent with the fact that the sample is not directly connected to the program and this is the average payoff spillover which is reasonable to be small. The estimated consideration level is 0.59, which means a launch makes the average first-timer about 1.8 times as likely to consider the targeted industry relative to a scenario of no launch.

Second, the table shows that the two channels have opposite directions in most characteristics. For example, consideration channel rises with science training ($+0.79$) and falls with PhD share (-1.18) and migrant share (-0.61). This pattern suggests that PhD-heavy and migrant-heavy teams are more resource-constrained than uninformed, and that this is why most of their response comes through payoffs rather than information. This is intuitive: PhD-heavy teams are already highly specialized in a specific field and are probably well informed about their niche in the industry, so a shock can affect their entry only if it provides them with resources. Similarly, teams with many migrant founders are probably those with more specific specializations and educational backgrounds, but are likely less connected and more resource-constrained. On the other hand, science-trained founders with bachelor's or master's degrees in science or engineering are more likely to be uninformed about where their capabilities fit, and an information campaign can redirect them toward certain industries. For example, someone with a bachelor's degree in electrical engineering might not consider

founding a quantum startup, even though she is equally capable of founding either a quantum or an energy company. An electrical engineer with a PhD, by contrast, is more likely to be specialized in quantum or in energy, which gives her better information about where she fits best. Note that science training is a relevant margin in our work because the accelerator launch that we exploit focuses on science-based ventures and targets those industries.

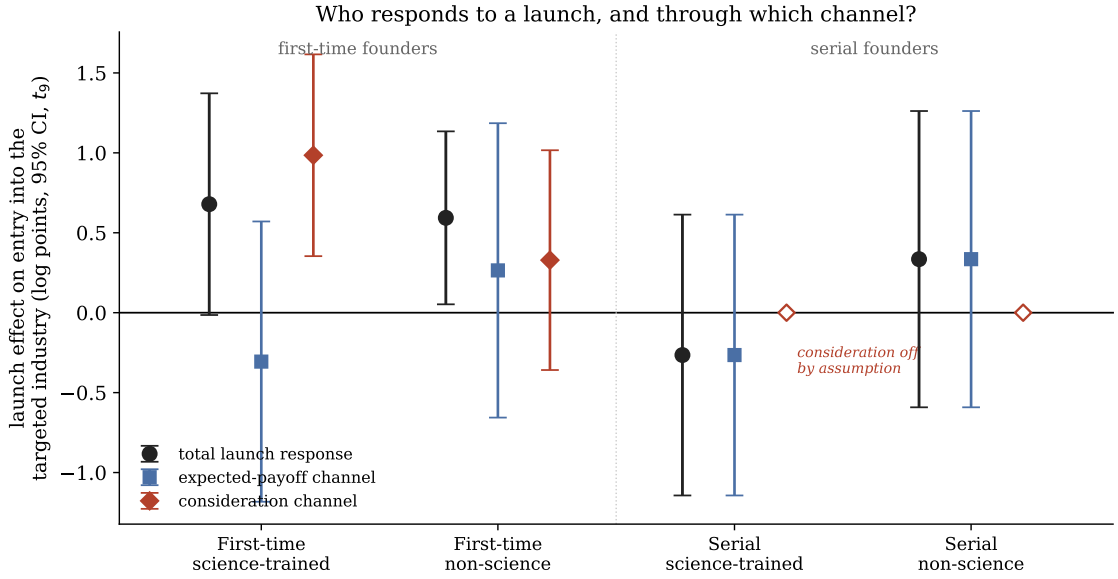
The payoff channel in our estimation is identified from the response of experienced founders, so, serial founders with PhD-heavy and migrant-heavy teams respond positively to the payoff channel, consistent with these groups being more resource constrained than an average serial founder. an interesting results is that the payoff response declines with science training that could suggest for an experienced founder in a science field, a launch also announces more competition in the future, whereas experienced non-science founders might see that as entry of more complement talents in the future. The mirror pattern is the table is consistent with an information friction with heterogeneous informedness. The two channels move different people: information is relatively more important for uninformed-relevant founders, while payoff matters relativeley more for the informed-but-resource-constrained ones.

Third, we find that consideration channel fades with the pre-launch industry size, however, the payoff channel does not depend on that. This is intuitive as information channel should be weaker for an industry that is already known in the local market. At the lower support of pre-launch industry size in our sample, the implied expansion for a science-trained first-timer is roughly 3.4×, decreasing to one (zero expansion effect) as the industry matures. Because a probability cannot exceed one, an expansion of 3.4× implies that at most about 30 percent of such founders considered that emerging industry before the launch, a.

The estimates are also robust to industry-level shocks.

Figure 4 summarizes the estimates at the group level. Science-trained first-timers respond mostly through consideration channel. By assumption,the consideration channel for serial founders is zero, the average payoff channel is not significantly different from zero for our main type. Looking at first-time non-science founders, the total effect for them is positive and significant, the point estimates show positive values for both channels, but we cannot separate the two channels for those founders as the confidence intervals are wide.

Figure 4: **Who Responds, and Through Which Channel.** This figure reports the total launch response and its decomposition into the payoff and consideration channels, by founder group, evaluated at group mean characteristics and at the mean launched cell. The consideration channel is zero for serial founders by construction. Whiskers are 95 percent confidence intervals, clustered by city.



The estimated saturation coefficient implies a window of emergence: in an established local industry, menus are already full and a launch is only as effective as its payoff channel. In a young industry the launch functions also as advertising influencing uninformed-relevant founders. In our counterfactual analysis, we use the estimated model, and remove the consideration channel, while payoffs, including the launch's payoff effects, are held at their estimated values.

The consideration channel is identified by the exclusion restrictions above, and a payoff channel specific to first-time- science trained founders only in young industries would violate our identification. The estimates limit how special such a mechanism would have to be. It could not be a general improvement in fundamentals that serial founders would also benefit, Because observationally similar founders with only founding experience do not respond, and aggregate financing in treated city-industries also does not change. It could not be a general first-timer advantage either, since non-science first-timers do not respond, for example financing , investor, mentoring or network effect that is similar to science-trained versus for example business background founders are not violating our identification. It would then be an indirect payoff spillover of the program, specific to science-trained first-time founders in emerging industries.

The consideration channel is identified by the exclusion restrictions above. Our model

allows the payoff channel to vary with founder characteristics, including science training. what would violate the identifying restriction is a payoff effect that varies with founding experience conditional on those characteristics, and science-trained first-time founders benefiting more from those payoff spillovers relative to science-trained serial founders in the same industry. The estimates limit how special such a mechanism would have to be. It could not be a general improvement in local fundamentals, since serial founders do not respond. It could not be a field-specific effect either, because any improvement in financing, investor attention, mentorship, or local networks that favors science-based over business-based ventures is absorbed by the payoff channel's science gradient. It would have to be an indirect spillover of the program that reaches science-trained founders without prior founding experience, in industries with little pre-existing local presence, at the moment of the launch. One would also expect that these payoffs to accumulate over time instead of peaking at the launch time especially for founders who are not directly connected to the program. Also note that we find no improvement in financing conditions for first time science trained entrants after a launch relative to before the launch. this is what the investors moved first and founders followed the money version of the payoff story would require, that we do not observe.

5 Counterfactual: The Role of Consideration Channel

We use the estimated model to quantify how the allocation of entrepreneurial entry would differ in the absence of the consideration channel, holding the payoff channel and the population of founders fixed. This counterfactual measures the contribution of consideration to entry into targeted industries and the associated reallocation of entrepreneurial activity across industries.

5.1 Entry Reallocation Through Consideration

We construct the counterfactual by setting $\lambda(\cdot) = 0$ for every founder while retaining all estimated payoff effects. We then recompute each founder's choice probabilities over industries. The set of potential entrants is held fixed, so the exercise measures the reallocation of a fixed population of founding teams.

Let $\mathcal{L}_{c(i)t(i)}$ be the set of accelerator-targeted industries active in founder i's city and

year. For founder group g, fitted entry into launched industries under $r \in \{F, NC\}$ is

$$\widehat{N}_g^r = \sum_{i \in g} \sum_{s \in \mathcal{L}_{c(i)t(i)}} \widehat{P}_{ist}^r,$$

where F denotes the factual environment and NC denotes the counterfactual without the consideration channel. We define entry attributable to consideration as

$$\widehat{\Delta N}_g = \widehat{N}_g^F - \widehat{N}_g^{NC},$$

and its share of factual entry as

$$\widehat{S}_g = \frac{\widehat{N}_g^F - \widehat{N}_g^{NC}}{\widehat{N}_g^F}.$$

A positive value indicates that the consideration channel redirects founders into launched industries; a negative value indicates that it redirects them away from those industries.

Table 7 reports the results. In the absence of the consideration channel, science-trained first-time entry into launched industries falls by 27.4 founders, or 30 percent of factual entry. Thus, the model attributes approximately one in four science-trained first-time entries into these industries to the consideration channel.

Table 7: **Entry Attributable to the Consideration Channel.**

	Entrants	95% interval	Share of base
Science-trained	+27.4	[+13.0, +37.5]	+22.8%
Non-science	−10.2	[−35.5, +9.3]	−8.2%
All founders	+17.1	[−16.3, +42.2]	+7.0%
Option value, log points × 100			
Science-trained first-timers	+7.38	[+3.18, +10.07]	
Non-science first-timers	−2.80	[−8.78, +1.34]	

Notes. Change in entry into launched city–industries when the consideration channel is removed, holding the payoff channel, the fixed effects, and the population of entrants fixed. The model is a discrete choice over industries, so it speaks to where founders build rather than how many; the absence of a total founding response to a launch is what makes reallocation the relevant margin. Intervals are 95 percent intervals from 2,000 parametric bootstrap draws of the channel parameters from a multivariate t_{G-1} distribution with the city-clustered covariance ($G = 10$ cities), with the counterfactual recomputed founder by founder inside each draw. Denominators are observed entrants and are held fixed across draws, so share intervals are count intervals rescaled. The option value is the change in the model's log inclusive value when the channel is switched off, in payoff-equivalent log points.

The effect for non-science first-time founders has the opposite sign. For this group, the estimated consideration channel redirects entry away from launched industries. This result

reflects their negative estimated consideration response and should not be interpreted as displacement caused by the inflow of science-trained founders. The positive effect for science-trained founders and the negative effect for non-science founders largely offset one another: the net effect on total entry into launched industries is 7.0 percent. The consideration channel therefore primarily changes the composition of entry: it redirects science-trained first-time founders into targeted industries and non-science first-time founders toward other industries, with little change in the total number of firms entering launched industries.

Figure 5 reports the change in entry related to the consideration channel by launched industry, as a percent of baseline fitted entry in that industry's launched city-years, split into science-trained and non-science first-time founders. Numerator and denominator are both restricted to launched cells, the only cells in which λ is non-zero and relevant; baseline entry counts appear in parentheses and industries with fewer than three baseline entrants are omitted. Diamonds are net effects with 95 percent intervals from 2,000 parametric bootstrap draws of the launch parameters from a multivariate t_{G-1} distribution with the city-clustered covariance ($G = 10$ cities), recomputed founder by founder within each draw. Negative bars reflect estimated negative consideration responses for the founders concerned, not displacement by the inflow of science-trained founders. All flows are first-time founders, and the channel does not move serial founders by construction. Percentages for Climate, Space and Commerce rest on baselines of four to five founding teams and should be read with their intervals.

Figure 6 reports the Science-trained first-time entry attributable to the Science-trained first-time entry attributable to the consideration channel, as a share of that industry's baseline entry in launched cells.

Figure 5: **Composition of the Reallocated Entry within Launched City-Industries.** This figure reports the change in entry attributable to the consideration channel by launched industry, as a percent of baseline fitted entry in that industry's launched city–years, split into science-trained and non-science first-time founders. Numerator and denominator are both restricted to launched cells, the only cells in which λ is non-zero; baseline entry counts appear in parentheses and industries with fewer than three baseline entrants are omitted. Diamonds are net effects with 95 percent intervals from 2,000 parametric bootstrap draws of the launch parameters from a multivariate t_{G-1} distribution with the city-clustered covariance ($G = 10$ cities), recomputed founder by founder within each draw. Negative bars reflect estimated negative consideration responses for the founders concerned, not displacement by the inflow of science-trained founders. All flows are first-time founders; the channel does not move serial founders by construction. Percentages for Climate, Space and Commerce rest on baselines of four to five founding teams and should be read with their intervals.

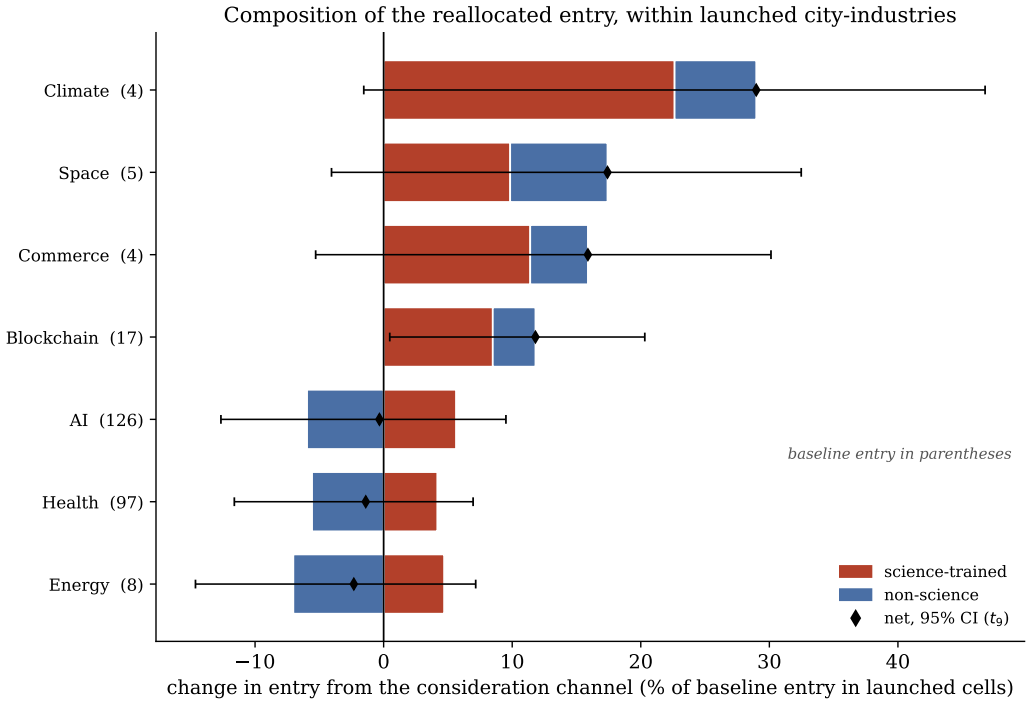
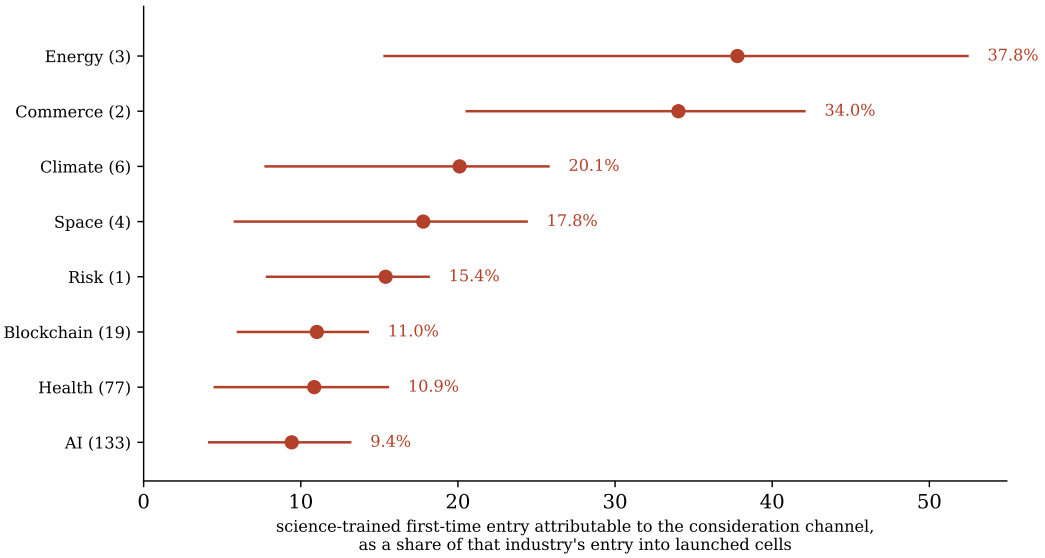


Figure 6: **Where the reallocation lands**

Notes. Each row reports the share of entry into a launched city–industry cell that is attributable to the consideration channel. The numerator is $\Delta_s = \sum_i \left(P_{ist} - P_{ist}^{\mathrm{CF}}\right)$, summed over science-trained first-time founding teams in the launched cells of industry s, where the counterfactual sets $\lambda(x_i, \tilde{A}_{cs}) = 0$ and leaves the payoff channel τ and all baseline terms unchanged; the denominator is the number of teams observed founding in those cells, given in parentheses beside each industry label. The pooled figure is 11.2% of the 245 entrants into launched cells, equivalently 30.1% of the 91 science-trained first-time entrants. Estimates come from a fixed-effects Poisson fit of the attention-weighted logit on 5,573 founding teams and 86,953 team × industry observations, with team and city × industry fixed effects and industry × characteristic sorting controls, across 141 city × industry cells of which 21 are launched. Horizontal bars are 95% intervals from a multivariate-$t(9)$ parametric bootstrap of the six consideration parameters (2,000 draws), holding the fitted choice probabilities fixed; standard errors are clustered by city. Because every interval is a transformation of the same parameter draws, they are strongly correlated across industries and should not be read as eight independent tests. Four industries have fewer than ten entrants, so their point estimates are sensitive to a single founding team.



5.2 Capital Associated with Consideration-Induced Entry Reallocation

We use the estimated choice model to measure the capital associated with the reallocation of entry through the consideration channel. This is an accounting exercise rather than a welfare calculation.

Attribution. For founder i observed entering launched city-industry s, let $\widehat{P}_{is}^{F}$ denote the fitted probability of choosing that industry under the estimated payoff and consideration

channels, and $\widehat{P}_{is}^{NC}$ the probability after setting $\lambda_i = 0$ while holding the payoff channel, founder characteristics, fixed effects, and the population of founders unchanged. We assign the entrant the attribution weight

$$\omega_i = \frac{\widehat{P}_{is}^{F} - \widehat{P}_{is}^{NC}}{\widehat{P}_{is}^{F}}, \tag{6}$$

the model-implied fraction of that entry attributable to the consideration channel. Two properties of (6) matter. It attributes the founder's industry *destination*, not the decision to become an entrepreneur, because the analysis conditions on founding. And it aggregates consistently with Section 5.1: the expectation of $\sum_i \omega_i$ over realized entrants is $\sum_i(\widehat{P}_{is}^{F} - \widehat{P}_{is}^{NC})$, so the two exercises must report the same total. They do—17 against 17.1 in Table 7.

The weight is negative for the 76 of 245 treated-cell entrants with $\widehat{\lambda}_i < 0$, whom the channel moved away from s and who entered anyway. We report gross ledgers, summing over $\omega_i > 0$, alongside net ledgers over all entrants, rather than truncating.

Gross capital. Multiplying each entrant's realized three-year early-stage capital K_i by ω_i and summing gives $K^{\text{gross}} = \sum_{i:\,\omega_i>0} \omega_i K_i$; dividing by the total raised by firms in launched city-industries gives the share of observed capital associated with attributed entries. This is not capital created by the launch. It is capital raised by firms whose industry destination the model attributes to the consideration channel.

Counterfactual destinations. Let $\widehat{m}_{ik}$ denote predicted log three-year capital for founder i in industry k. The exact change in expected predicted capital induced by the channel is

$$\Delta\widehat{m}_i = \sum\nolimits_k \big(\widehat{P}_{ik}^{F} - \widehat{P}_{ik}^{NC}\big)\widehat{m}_{ik}, \tag{7}$$

and we report the per-redirected-founder gap $\Delta\widehat{m}_i/(\widehat{P}_{is}^{F} - \widehat{P}_{is}^{NC})$, averaged across entrants with weights ω_i. When a launch changes the consideration weight of the targeted industry alone, (7) is equivalent to comparing $\widehat{m}_{is}$ with $\sum_{k\neq s} q_{ik}\widehat{m}_{ik}$, using diversion weights $q_{ik} = (\widehat{P}_{ik}^{NC} - \widehat{P}_{ik}^{F})/(\widehat{P}_{is}^{F} - \widehat{P}_{is}^{NC})$ that sum to one. We work with (7) directly because that equivalence is not the general case in our data: 148 of the 245 treated-cell entrants face more than one simultaneous launch in their city–year, and for 68 attributed entrants at least one q_{ik} is negative, because the other launched industries also lose consideration in the counterfactual. Equation (7) requires no monotonicity and no reading of q as a distribution.

Since $\widehat{m}$ is additively separable and its founder-level terms do not vary across indus-

tries, those terms cancel in (7), and the gap splits exactly into an industry-composition component—differences in average capital intensity across industries—and a match component—the science×science-based interaction, the only term in which the particular founder–industry pairing appears.

Table 8: Capital associated with consideration-induced reallocation

	Estimate	95% interval
Entries attributed to the consideration channel		
Attributed entrants, gross	36.8	[21.2, 56.7]
Attributed entrants, net	16.7	[−18.8, 41.0]
Three-year early-stage capital raised by those entrants		
Gross ($M)	34.8	[19.7, 57.6]
as % of treated-cell capital	8.4%	[4.8, 13.9]
Net ($M)	10.0	[−30.8, 42.3]
as % of treated-cell capital	2.4%	[−7.4, 10.2]
Predicted capital versus counterfactual destinations (log points × 100)		
Composition (industry capital intensity)	+13.5	[6.3, 17.3]
Match (science × science-based)	+9.6	[1.3, 17.6]
Total	+23.0	[12.3, 32.0]

Notes. Three-year early-stage capital (Crunchbase, winsorized at the 99th percentile); 245 treated-cell entrants raising $414M in total. Attribution weights from (6) and destination comparisons from (7), both computed founder by founder and then aggregated. Gross ledgers sum over entrants with $\omega_i > 0$; net ledgers sum over all treated-cell entrants, including those the channel moved away from the launched industry. Capital model: $\log(1 + K)$ on a science-degree indicator, first-time-founder status and the science×science-based interaction, with industry, city and cohort-year fixed effects, clustered by city; the estimated match premium is +0.069 (0.026). Intervals: 2,000 parametric bootstrap draws, jointly over the consideration parameters and the capital-model match premium, from multivariate t_{G-1} distributions with the city-clustered covariance ($G = 10$ cities), with every quantity recomputed founder by founder inside each draw. Given no aggregate financing response, the ledger measures reallocation of capital, not creation.

Results. Table 8 reports the ledger. Entries attributed to the consideration channel account for 8.4 percent of the three-year capital raised in launched city-industries, and 2.4 percent net of the entries the channel moved the other way. Among attributed entries, predicted log capital in the observed industry is 23.0 log points higher than at the model-implied counterfactual destinations, of which 9.6 points are associated with the founder–industry match rather than with differences in average capital intensity across industries. The match component is the part that cannot be produced by sorting founders into richer industries, and it is separated from zero at conventional levels.

Figure 7 reports heterogeneity across industries. In emerging launched industries, a

comparatively large share of observed capital is associated with entries attributed to the consideration channel. Because aggregate local financing does not increase detectably following launches, we interpret these results as a reallocation of founders, and of the capital that follows them, across industries—not as evidence that the accelerator creates an equivalent amount of new aggregate capital.

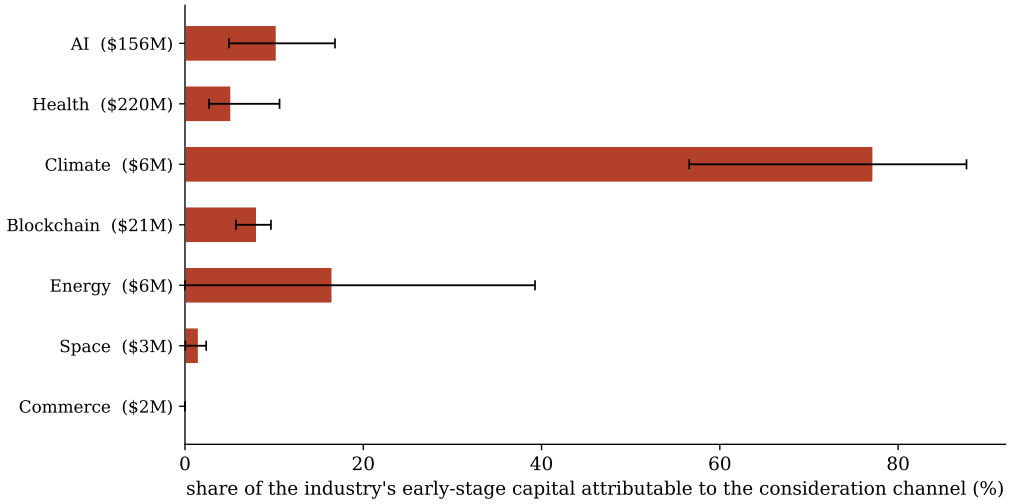


Figure 7: Share of each industry's early-stage capital associated with consideration-induced entry

Notes. Gross attributable capital, $\sum_{i:\,\omega_i>0} \omega_i K_i$ within each launched industry, as a share of that industry's total three-year early-stage capital raised in launched city–industries. Industry capital totals in parentheses. Whiskers are 95 percent intervals from 2,000 parametric bootstrap draws of the consideration parameters from a multivariate t_{G-1} distribution with the city-clustered covariance ($G = 10$ cities). Shares for Climate, Energy, Space and Commerce rest on industry totals of $6M or less and should be read with their intervals; AI and Health account for 91 percent of capital raised in launched city–industries.

Two assumptions behind $\widehat{m}$. First, $\widehat{m}_{ik}$ for $k \neq s$ is an out-of-sample prediction: it assumes founders do not select into industries on unobservables that also predict capital. Second, the destination comparison is predicted against predicted, so realized outcomes enter only through the estimated coefficients and no individual firm's fundraising luck is attributed to the channel.

6 Conclusion

This paper provides empirical evidence on how targeted place-based innovation policy shapes local entrepreneurial ecosystems. By exploiting the staggered rollout of industry-specific

streams within the Creative Destruction Lab, and by studying only founders who never apply to or join the program, we separate the payoff channel such policies, where they raise the expected payoff off of entry intro targeted industries, from consideration channel, where the policy redirects the attention of potential founders towards targeted industries.

Our reduced-form analysis show that accelerator launches redirect entrepreneurial entry towards targeted industries, and this reallocation is stronger in younger emerging industries in each region. Entry by science-trained first-time founders into the targeted industry rises,while experienced founders do not respond, and city-wide founding does not change. This increase in entry of not-connected firms is stronger among founders connected to program participants, and is not accompanied by an aggregate financing increase. These facts motivate a model separating payoff changes from changes in consideration. We build and estimate a model of entrepreneurial industry choice that separates the payoff and consideration channels of accelerator launches. We find that first-time science-trained entrepreneurs are reallocated towards targeted industries mostly through consideration channel, and this effect declines if the targeted industry is already salient in the local market.

Quantitatively, the consideration channel accounts for 30 percent of science-trained first-time entry into launched city-industries. Entry by founders without science training moves in the opposite direction, and the net change in total entry into these industries is indistinguishable from zero. Firms whose industry choice the model attributes to the attention channel account for 8.4 percent of the early-stage capital raised in launched city-industries, and their expected capital is 23.0 log points higher than at the destinations they would otherwise have chosen, with the improved match between founder and industry contributing 9.6 of those points.

These results have implications for the design of place-based policies. Debates over such policies focus on whether programs work and where they should locate; our results show that the impact of such policies on entrepreneurial activity goes beyond the fundamentals they provide, and can potentially relocate entrepreneurial attention toward viable opportunities.

While we document the short-run response of entry, financing, and matching, the long-run welfare implications remain an open question. Whether the reallocation of talent toward targeted sectors leads to more productive regional specialization or to fragile clusters depends on the long-term performance, and our early evidence is that the redirected founders land in matches investors fund more heavily. A second open question concerns the margin we cannot measure. Our model conditions on founding, because the founding margin does not respond to launches in our data; whether information of this kind moves the decision to found at all, rather than the decision of what to found, requires data on potential founders who never

enter. Both questions seem to us worth pursuing, and both are natural extensions of the framework developed here.

References

Ackerberg, D. A. (2003). Advertising, learning, and consumer choice in experience good markets: An empirical examination. *International Economic Review*, 44(3):1007–1040.

Acs, Z., Åstebro, T., Audretsch, D., and Robinson, D. T. (2016). Public policy to promote entrepreneurship: A call to arms. *Small Business Economics*, 47(1):35–51.

Adelino, M., Ma, S., and Robinson, D. T. (2017). Firm age, investment opportunities, and job creation. *The Journal of Finance*, 72(3):999–1038.

Babina, T. (2020). Destructive creation at work: How financial distress spurs entrepreneurship. *The Review of Financial Studies*, 33(9):4061–4101.

Balsmeier, B., Fleming, L., Marx, M., and Shin, S. R. (2025). Startups, unicorns, and the local influx of inventors. *Review of Economics and Statistics*, pages 1–44. Published online.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., and Van Reenen, J. (2019). Who becomes an inventor in america? the importance of exposure to innovation. *Quarterly Journal of Economics*, 134(2):647–713.

Bloom, N., Van Reenen, J., and Williams, H. (2019). A toolkit of policies to promote innovation. *Journal of Economic Perspectives*, 33(3):163–184.

Brady, R. L. and Rehbeck, J. (2016). Menu-dependent stochastic feasibility. *Econometrica*, 84(3):1203–1223.

Bronnenberg, B. J., Dubé, J.-P., Gentzkow, M., and Shapiro, J. M. (2015). Do pharmacists buy bayer? informed shoppers and the brand premium. *Quarterly Journal of Economics*, 130(4):1669–1726.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427.

Chatterji, A., Glaeser, E., and Kerr, W. (2014). Clusters of entrepreneurship and innovation. *Innovation Policy and the Economy*, 14(1):129–166.

Chatterji, A. K. (2009). Spawned with a silver spoon? entrepreneurial performance and innovation in the medical device industry. *Strategic Management Journal*, 30(2):185–206.

Chen, J. and Roth, J. (2024). Logs with zeros? Some problems and solutions. *The Quarterly Journal of Economics*, 139(2):891–936.

Decker, R., Haltiwanger, J., Jarmin, R., and Miranda, J. (2014). The role of entrepreneurship in us job creation and economic dynamism. *Journal of Economic Perspectives*, 28(3):3–24.

Ellison, G. and Glaeser, E. L. (1997). Geographic concentration in U.S. manufacturing industries: A dartboard approach. *Journal of Political Economy*, 105(5):889–927.

Ellison, G., Glaeser, E. L., and Kerr, W. R. (2010). What causes industry agglomeration? evidence from coagglomeration patterns. *American Economic Review*, 100(3):1195–1213.

Ewens, M., Nanda, R., and Rhodes-Kropf, M. (2018). Cost of experimentation and the evolution of venture capital. *Journal of Financial Economics*, 128(3):422–442.

Fehder, D. C. and Hochberg, Y. V. (2014). Accelerators and the regional supply of venture

capital investment. *Working Paper*.

Foster, L., Haltiwanger, J., and Krizan, C. J. (2001). Aggregate productivity growth: Lessons from microeconomic evidence. In *New developments in productivity analysis*, pages 303–372. University of Chicago Press.

Glaeser, E. L., Kerr, S. P., and Kerr, W. R. (2015). Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics and Statistics*, 97(2):498–520.

Goeree, M. S. (2008). Limited information and advertising in the u.s. personal computer industry. *Econometrica*, 76(5):1017–1074.

Gompers, P., Lerner, J., and Scharfstein, D. (2005). Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999. *The Journal of Finance*, 60(2):577–614.

Gonzalez-Uribe, J. and Leatherbee, M. (2018). The effects of business accelerators on venture performance: Evidence from start-up chile. *Review of Financial Studies*, 31(4):1566–1603.

Hallen, B. L., Cohen, S. L., and Bingham, C. B. (2020). Do accelerators work? if so, how? *Organization Science*, 31(2):378–414.

Handel, B. R. and Kolstad, J. T. (2015). Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105(8):2449–2500.

Heydari-Nejad, M. (2024). A structural model of mentorship in startup accelerators: Matching, learning, and value creation. Technical report, Univerity of Toronto Working Paper.

Honka, E., Hortaçsu, A., and Vitorino, M. A. (2017). Advertising, consumer awareness, and choice: Evidence from the u.s. banking industry. *RAND Journal of Economics*, 48(3):611–646.

Khanna, G. and Morales, N. (2025). The it boom and other unintended consequences of chasing the american dream.

Kline, P. and Moretti, E. (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *Quarterly Journal of Economics*, 129(1):275–331.

Lee, S. R. and Glennon, B. (2023). The effect of immigration policy on founding location choice: Evidence from canada's start-up visa program. Technical report, National Bureau of Economic Research.

Lerner, J. (2009). *Boulevard of Broken Dreams: Why Public Efforts to Boost Entrepreneurship and Venture Capital Have Failed—and What to Do About It*. Princeton University Press, Princeton, NJ.

Lyons, E. and Zhang, L. (2018). Who does (not) benefit from entrepreneurship programs? *Strategic Management Journal*, 39(1):85–112.

Manzini, P. and Mariotti, M. (2014). Stochastic choice and consideration sets. *Econometrica*, 82(3):1153–1176.

Muendler, M.-A., Rauch, J. E., and Tocoian, O. (2012). Employee spinoffs and other entrants: Stylized facts from brazil. *International Journal of Industrial Organization*, 30(5):447–458.

Murphy, K. M., Shleifer, A., and Vishny, R. W. (1991). The allocation of talent: Implications for growth. *Quarterly Journal of Economics*, 106(2):503–530.

Nanda, R. and Rhodes-Kropf, M. (2013). Investment cycles and startup innovation. *Journal of Financial Economics*, 110(2):403–418.

Santos Silva, J. M. C. and Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics*, 88(4):641–658.

Sariri, A. (2025). The economics of advice: Evidence from start-up mentoring. *Management science*, 71(12):10022–10046.

Sariri, A., Gatov, E., Neal, G., Robinson, K., Sennik, S., Tham, W. Y., Vertolli, M., and Goldfarb, A. (2025). Database, methodological tools, and research opportunities: Creative destruction lab and early-stage technology ventures. Technical report, National Bureau of Economic Research.

Savage, J. and Ziedonis, A. A. (2024). Testing the waters: founding team composition and search heuristics in academic entrepreneurial ventures. *Industrial and Corporate Change*, 33(1):152–171.

Saxenian, A. (1994). *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge, MA.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Webb, M. D. (2023). Reworking wild bootstrap-based inference for clustered errors. *Canadian Journal of Economics*, 56(3):839–858.

Appendix

A.1 Additional Empirical Results

This appendix collects the robustness exercises, all of which are produced by the replication package accompanying the paper.

Leave-one-city-out. Re-estimating equation (2) while dropping each of the ten cities in turn, the launch effect for science-trained first-time founders ranges from +0.73 to +1.25 log points. Excluding Toronto, the program's home city, which accounts for the largest number of launches, strengthens the estimate.

Randomization inference. We reassign each launch to a randomly chosen host city that never launched that industry, holding the industry and the launch year fixed. This preserves the timing of launches, the number of treated cells per city, and the industry-wave structure exactly, while breaking the link between the launch and its actual location. Across 499 placebo assignments, the distribution of the composition effect is centered at zero and the estimate from the true assignment lies in its upper tail (one-sided $p = 0.02$ in emerging cells).

Anticipation and pre-decision placebos. Adding an announcement-year indicator ($t_0 - 1$) and a pre-decision window ($t_0 - 3$ to $t_0 - 2$) to equation (2), the response for science-trained first-timers appears at launch and not before: the announcement-year coefficient is small and insignificant, and the pre-decision coefficient is significantly negative. Because CDL commits to a stream roughly a year before the first cohort, this pattern is inconsistent with placement in anticipation of local surges in science-founder activity.

The announcement is local. Restricting attention to city-industry cells that never received a launch, we estimate the response of typed entry to the first launch of that industry *anywhere*. Science-trained first-time founders in these cells do not respond, while serial founders do. Informed founders track the industry's global moment wherever they are located; uninformed founders respond only where the announcement is made.

Composition at the firm level. Estimated as a linear probability model on entrants rather than as counts, the probability that an entrant in a launched cell is a science-trained

first-time founder rises by about fifteen percentage points at the average launched cell, against a base of seventeen percent, with the same negative gradient in pre-launch local industry size. The share of entrants who are science-trained *serial* founders is unchanged, which is the placebo implied by the information interpretation.

Industry×year fixed effects in the structural model. Re-estimating the model of Section 4 with industry×year fixed effects added to the team and city×industry fixed effects, main results hold.

A.2 CDL launches and access to financing

The most direct payoff channel that supports the increase in entry into targeted industries is that founders anticipate improved access to financing following a CDL launch. Even founders who do not participate in the program could benefit if a launch attracts investors or increases their willingness to fund local ventures in the targeted industry. We examine this possibility using financing outcomes following specification (1).

Table A.2 reports the results. Panel A reports the number of firms receiving disclosed financing. Although the estimated change in the total number of financed firms is not statistically significant, the age-specific estimates indicate a shift toward younger firms. The number of financed firms aged zero to three years increases by approximately 29%, while the number aged over seven years declines by approximately 39%. These patterns remain similar when accepted CDL participants are excluded. Panel B examines whether this change is accompanied by an increase in the amount of capital raised. We find no statistically significant increase in total financing in targeted industries, either in the full sample or after excluding accepted participants. In particular, the increase in the number of young firms receiving financing is not accompanied by a detectable increase in the capital raised by that age group.

Table A.3 complements these aggregate results by examining subsequent outcomes among startups founded before and after a launch in the corresponding city-industry. The sample covers firms founded between 2010 and 2019 and excludes accepted CDL participants. The specifications control for founding-team characteristics and include city × industry and founding-year fixed effects. Post-launch entrants raise less capital, conditional on observed financing information, while the estimated difference in the total number of financing rounds is not statistically significant. They also complete fewer VC-led rounds. These patterns

provide little evidence that post-launch entrants enjoy systematically better financing outcomes.

The financing results offer little support for an explanation in which entry rises because a CDL launch brings a broad expansion in available capital. Instead, these results are consistent with greater competition for financing as entry expands. However, these results are realized outcomes, thus cannot rule out ex-ante optimistic financing expectations or improvements in other expected payoffs, such as access to customers, advice, or collaborators. We therefore use these results as supporting evidence against a broad financing-expansion explanation, while allowing for indirect payoff spillovers in the model.

A.3 Program Placement and Academic Networks

CDL's expansion required a local business school willing to host the program, a process that typically unfolds over months or years and is not tightly linked to short-run entrepreneurial conditions in the host city. Using OpenAlex data on publications coauthored by University of Toronto faculty and researchers at institutions in potential host cities, restricted to business and management journals, we document that academic connectivity to Toronto predicts both whether a city hosts a CDL site and when it does so. We report this relationship as evidence on *how* placement occurred rather than as an instrument. What academic connectivity captures is a slow-moving feature of a city, the presence of a research-intensive business school, and such features are absorbed by the city and city×industry fixed effects in our design. Used as an instrument for aggregate city-level entry, the same variation is weak and fragile: with ten cities, wild-bootstrap p-values for the aggregate entry response are far from conventional thresholds, weak-instrument-robust confidence sets are wide, and the estimates are driven by the home city. We therefore rely on the design described in Section 3.3 rather than on instrumental variables, and we report the city-level analysis as a null on the founding margin rather than as a positive result.

Table A.1: **Academic Connectivity and CDL Placement.** This table reports cross-sectional regressions of an indicator for whether a city hosts a CDL site on measures of academic connectivity in business-related fields between institutions in the focal city and the University of Toronto. *Has_Academic_Connection* is an indicator for any prior collaboration; the remaining measures capture the strength of the connection. Columns (1)-(3) report OLS with robust standard errors; columns (4)-(5) add country fixed effects with standard errors clustered by country. The table documents how placement occurred; it is not used as an instrument in the paper's estimates.

y = Whether a City Has CDL Site						
	(1)	(2)	(3)	(4)	(5)	(6)
Has_Academic_Connection	0.027***			0.027***		
	(0.008)			(0.001)		
Connection_Strength		0.001**			0.001***	
		(0.000)			(0.000)	
Academic_Connectivity_Instrument			0.020***			0.024***
			(0.008)			(0.006)
Country FE	No	No	No	Yes	Yes	Yes
Observations	79052	377	377	79021	357	357

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.4 Supplementary Figures and Tables

Figure A.1: **CDL Launching Dynamics.** This figure shows the launch timing of each CDL city-stream pair.

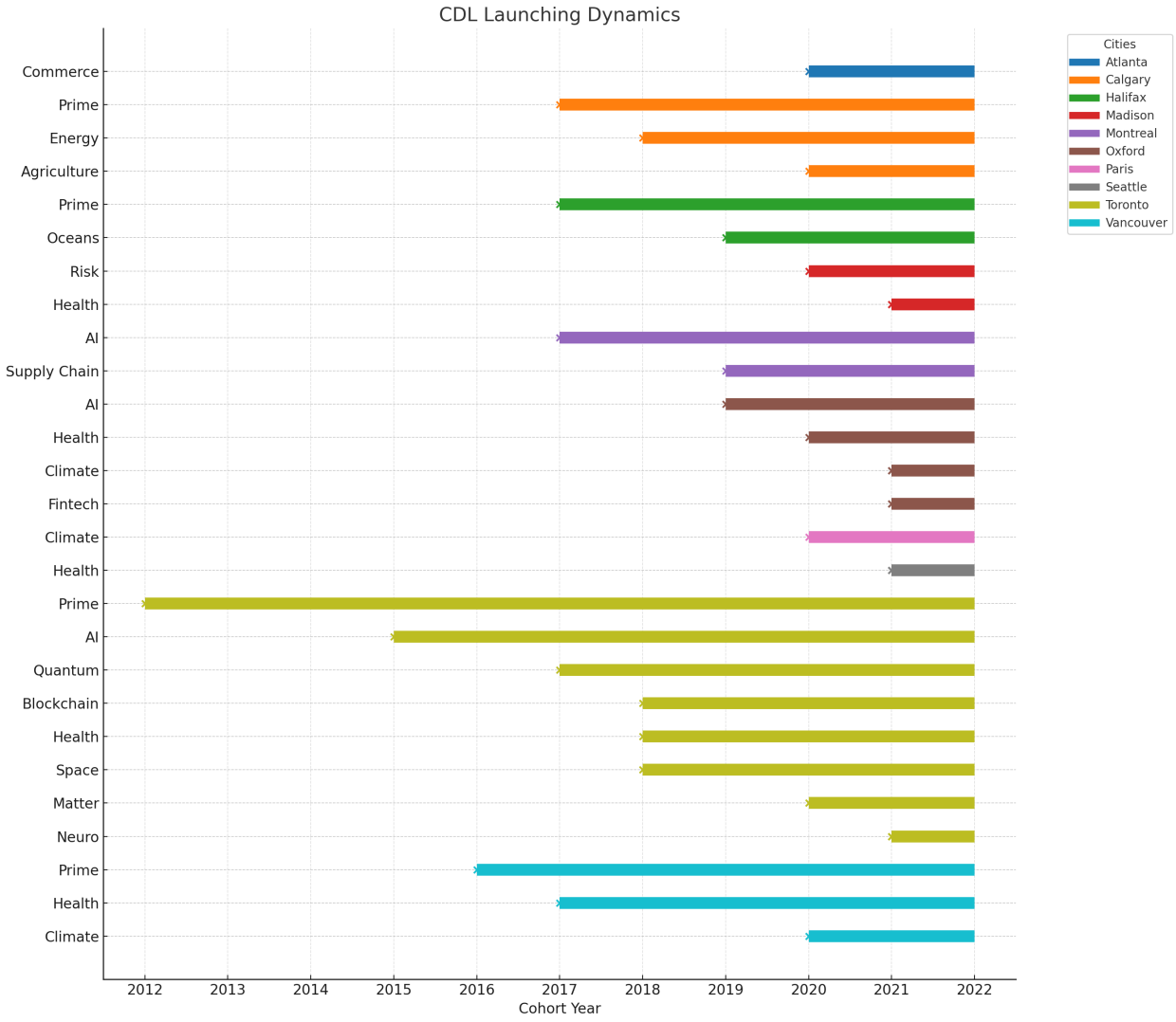


Figure A.2: **Size of the CDL Program in Focal Cities.** This figure shows the average number of firms accepted into CDL across focal cities, from the first cohort year (2012) to the most recent cohort (2021) in our data.

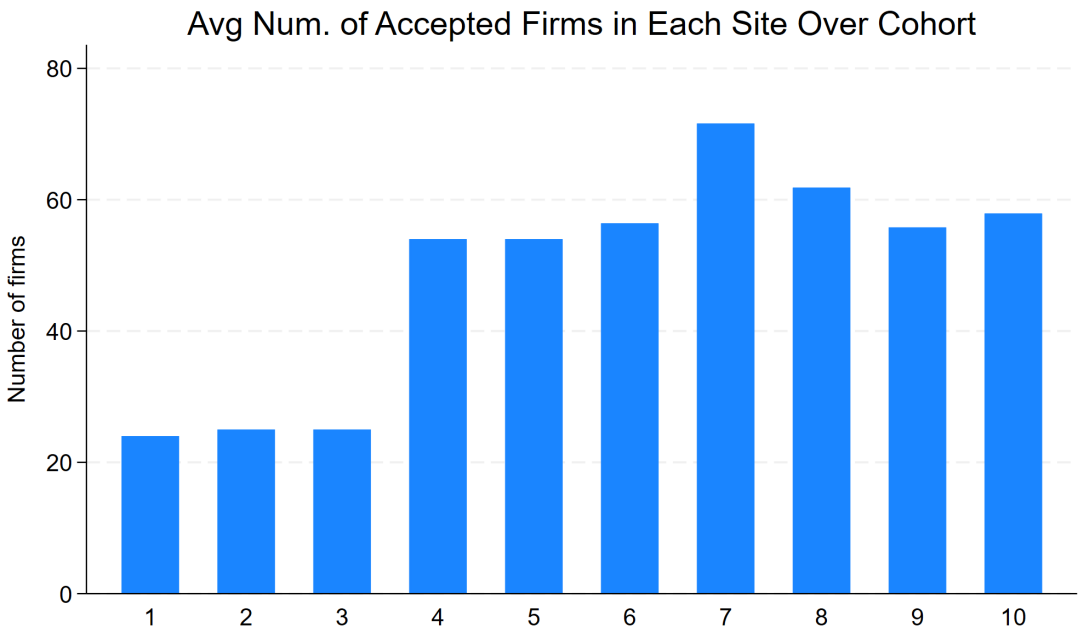


Figure A.3: **Reallocation Effect of Firm Entry around New CDL Stream Launch: Decomposing the effect between treated and untreated industries.** The figure shows reallocation patterns across industries around the CDL Stream Launch in Canada, by plotting the Callaway-Sant'Anna estimator. This figure compares firm entry into an industry within the focal city to entry into the same industry in never-treated cities, assuming that trends in never-treated cities provide a valid counterfactual for that industry's average attractiveness to entrepreneurs. Any difference reflects a change in an industry's attractiveness around the shock. We present results separately for (1) treated industries and (2) never-treated industries. Standard errors are clustered by city × industry.

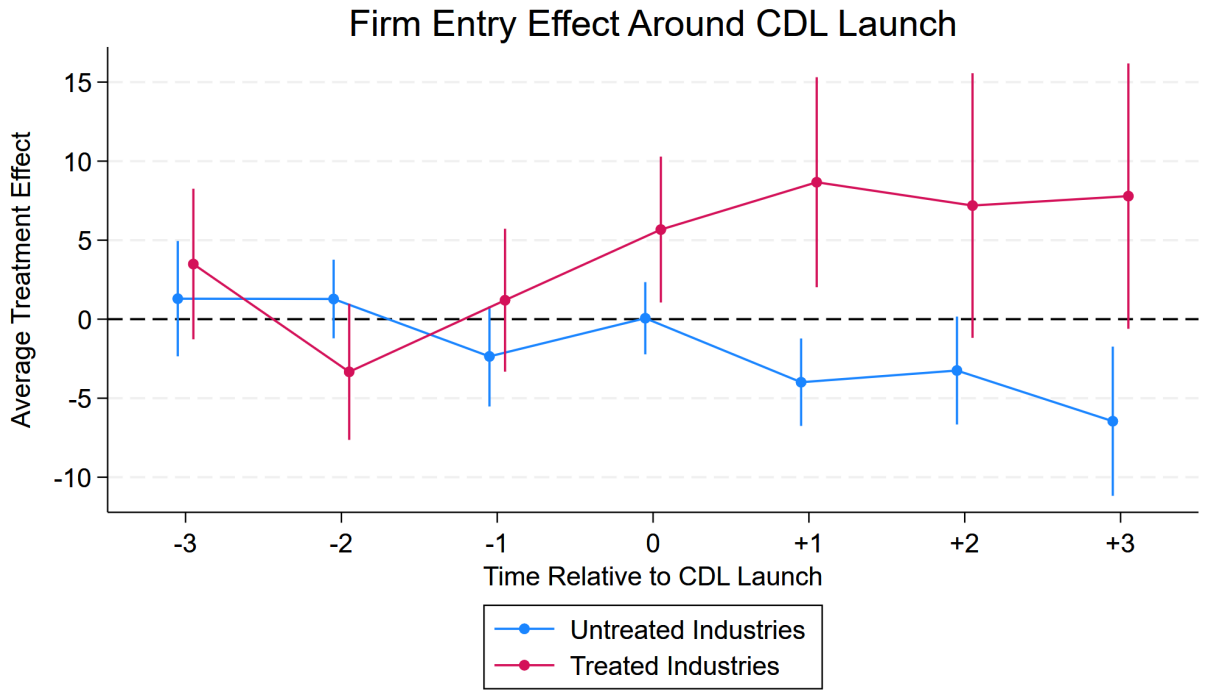
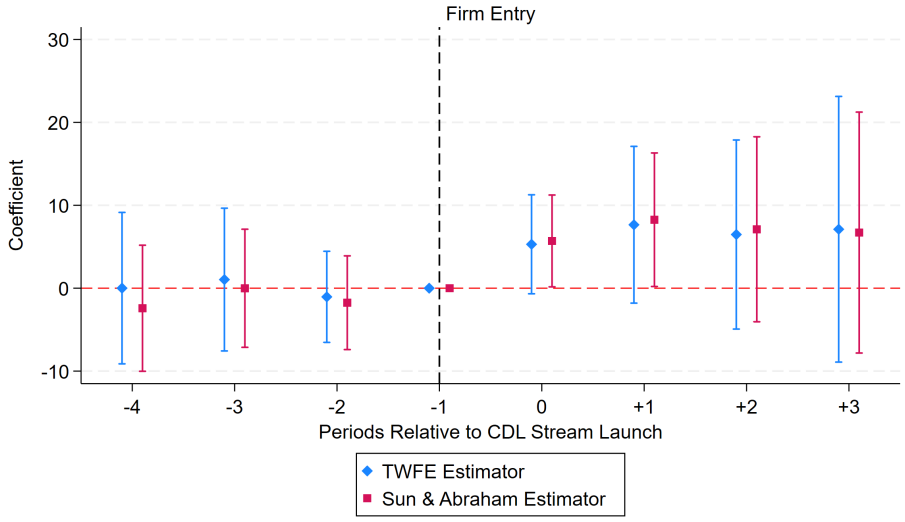
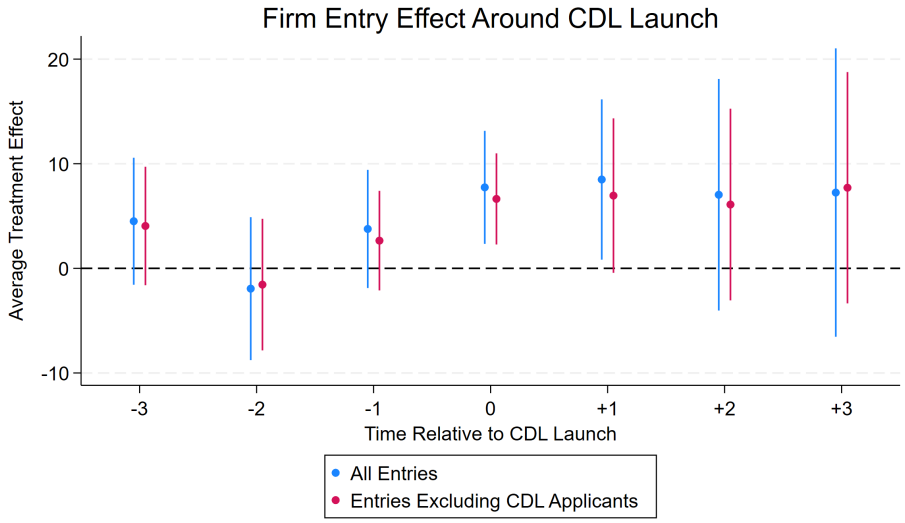


Figure A.4: Firms Reallocate Across Industries

Figure A.5: **Reallocation Effect of Firm Entry around New CDL Stream Launch: Robustness Event Graphs.** This figure shows the dynamics of firm entry and the amount of financing attracted around a new stream launch by CDL. We estimate a DID with high-level dimension fixed effects to control for potential confounding factors by following θ_t of equation $Y_{i,c,t} = \sum_{t \neq -1} \theta_t Periods_t * Launch_{i,c,t} + \alpha_{i,c} + \alpha_{i,t} + \alpha_{c,t} + \epsilon_{i,c,t}$. The estimates compare differences in firm entry and financing between the focal stream and other streams in the focal city with those in other cities around the stream launch year. To allow for an adequate post-period, we restrict our analysis to launches before 2020. Panel (a) shows TWFE and Sun & Abraham estimators; Panel (b) shows the Callaway and Sant'Anna estimators. The specification includes city × industry FEs, city × year FEs, and industry × year FEs. The year before the event is omitted as the base year. Standard errors are clustered by city × industry.



(a) Alternative Esitmators



(b) Callaway and Sant'Anna Estimators

Table A.2: **Capital Reallocation Around CDL Stream Launch in Treated Cities.** This table reports the estimated effects of CDL stream launches on the allocation of startup financing within treated cities. Panel A examines the number of firms receiving financing by firm age group. Panel B examines the amount of financing raised by firm age group. Columns compare the full sample with the sample excluding accepted CDL participants. These estimates test whether CDL launches are associated with a reallocation of financing toward younger versus older local firms. All specifications include city × industry fixed effects, city × year fixed effects, and industry × year fixed effects. Standard errors are clustered by city × industry.

Panel A: Number of Firms Received Financing

	Number of Firms Received Disclosed Financing (Poisson)							
	(1) Total	(2) Age 0-3	(3) Age 3-7	(4) Age >7	(5) Total	(6) Age 0-3	(7) Age 3-7	(8) Age >7
Treated × Post	0.098	0.254***	0.057	-0.497**	0.077	0.246***	0.070	-0.490**
	(0.063)	(0.093)	(0.110)	(0.196)	(0.063)	(0.090)	(0.138)	(0.194)
Implied % effect	10.3	28.9	5.8	-39.2	8.0	27.9	7.2	-38.7
Sample	Full	Full	Full	Full	Excl. Acc.	Excl. Acc.	Excl. Acc.	Excl. Acc.
City × Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1762	1712	1490	1355	1763	1705	1487	1352
Pseudo R-squared	0.761	0.701	0.562	0.474	0.801	0.701	0.562	0.472

Panel B: Total Amount Financed

	Total Log Amount Raised by Firms (OLS)							
	(1) Total	(2) Age 0-3	(3) Age 3-7	(4) Age >7	(5) Total	(6) Age 0-3	(7) Age 3-7	(8) Age >7
Treated × Post	-0.393	-0.281	-0.459	0.224	-0.431	-0.236	-0.484	0.286
	(0.395)	(0.560)	(0.613)	(0.555)	(0.411)	(0.523)	(0.614)	(0.583)
Implied % effect	-32.5	-24.5	-36.8	25.1	-35.0	-21.1	-38.4	33.1
Sample	Full	Full	Full	Full	Excl. Acc.	Excl. Acc.	Excl. Acc.	Excl. Acc.
City × Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1612	1015	748	1225	1594	995	737	1209
Adjusted R-squared	0.655	0.492	0.383	0.516	0.654	0.483	0.380	0.515

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: **CDL-induced Entrants and Subsequent Financing Activities.** This table compares startups founded after a targeted CDL stream launch with startups founded before the launch within the same city-industry cells. The sample is restricted to firms founded between 2010 and 2019 and excluding firms that were accepted to the program in order to capture the outcome of the remaining induced entrants. The treatment indicator equals one for startups founded after the CDL launch in the corresponding city-industry cell. The outcomes cover three dimensions of new firm performance: venture failure, financing activity, and extreme success. Failure outcomes include closure and website inactivity. Financing outcomes include whether the firm receives financing, the number of financing rounds, VC-led rounds, top-investor participation, financing within three years, and acquisition. Amount and valuation outcomes are conditional on observed financing or exit information. The table reports cross-sectional estimates with founder-team controls, including team size, serial entrepreneurship, gender, education, and demographic characteristics. All specifications include city × industry fixed effects and founding-year fixed effects. Standard errors are clustered by city × industry.

	Closed in CB	**No Financing Since 2022**	**Website Inactive**	**Amt Raised**	**N. Rounds**	**N. Rounds Led by VC**	**N. Top Investors**	**Exit Dummy**	**Disclose Exit Value**	**Log Exit Value**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated × Post	0.028*	-0.041	0.021	-0.249***	-0.016	-0.019	0.006	0.041*	0.168	2.412
	(0.016)	(0.032)	(0.042)	(0.084)	(0.055)	(0.023)	(0.010)	(0.024)	(0.129)	(1.741)
Founding Team Feature Controls										
N. Founder	0.055	0.015	0.020	0.223*	0.041	0.040	0.009	0.030	-0.148*	0.680
	(0.037)	(0.062)	(0.046)	(0.129)	(0.092)	(0.040)	(0.033)	(0.034)	(0.083)	(1.076)
N. Serial Entre	0.016**	-0.013	0.029**	0.059*	-0.003	-0.009	0.006	0.001	-0.032	1.824***
	(0.007)	(0.014)	(0.012)	(0.035)	(0.027)	(0.013)	(0.009)	(0.008)	(0.029)	(0.573)
N. Female	-0.012	0.019	0.007	-0.184***	-0.064**	-0.028***	-0.012***	-0.042***	-0.031	0.684
	(0.008)	(0.014)	(0.011)	(0.039)	(0.026)	(0.008)	(0.004)	(0.009)	(0.042)	(0.800)
N. PhD	0.019	0.056*	-0.006	0.150	0.083	0.028	0.022	0.030	0.073	-0.617
	(0.014)	(0.029)	(0.020)	(0.093)	(0.051)	(0.019)	(0.018)	(0.021)	(0.071)	(0.816)
N. Master	0.002	0.041***	0.001	-0.018	0.062*	0.016	0.009	0.022*	-0.044	-1.290*
	(0.009)	(0.012)	(0.010)	(0.038)	(0.037)	(0.011)	(0.011)	(0.011)	(0.027)	(0.640)
N. Bachelor	-0.000	0.041***	-0.011	0.000	0.057**	0.011	-0.001	0.020	-0.031	-0.493
	(0.008)	(0.014)	(0.011)	(0.035)	(0.025)	(0.014)	(0.007)	(0.013)	(0.030)	(0.660)
N. Asian	-0.056	0.018	-0.032	-0.200*	0.087	-0.032	-0.004	-0.003	0.177**	-0.180
	(0.038)	(0.060)	(0.042)	(0.118)	(0.099)	(0.046)	(0.033)	(0.034)	(0.083)	(0.825)
N. Black	-0.062	0.028	-0.075*	-0.191	0.126	-0.009	-0.004	0.010	0.142**	0.188
	(0.040)	(0.061)	(0.045)	(0.136)	(0.094)	(0.043)	(0.030)	(0.031)	(0.071)	(0.382)
N. Hispanic	-0.051	-0.055	-0.078	-0.289**	-0.054	-0.065	0.002	-0.020	0.001	2.037**
	(0.045)	(0.073)	(0.050)	(0.144)	(0.107)	(0.057)	(0.035)	(0.034)	(0.100)	(0.911)
N. White	-0.068*	0.010	-0.072	-0.117	0.126	-0.003	0.010	0.019	0.145**	0.000
	(0.040)	(0.056)	(0.043)	(0.122)	(0.093)	(0.041)	(0.032)	(0.033)	(0.069)	(.)
N. Bus_Degree	0.015*	-0.003	0.025**	-0.012	0.015	-0.002	0.004	-0.016**	-0.010	-0.308
	(0.008)	(0.015)	(0.011)	(0.029)	(0.029)	(0.013)	(0.005)	(0.008)	(0.035)	(0.599)
N. Sci_Degree	0.013**	0.000	0.014	-0.042	0.029	0.013	0.013	-0.004	0.004	0.366
	(0.006)	(0.014)	(0.011)	(0.028)	(0.024)	(0.009)	(0.009)	(0.010)	(0.032)	(0.941)
Number of Founder Missing Education Info	0.035***	0.061***	0.021	0.062	0.141***	0.064***	0.017	0.003	0.015	-0.785
	(0.011)	(0.023)	(0.014)	(0.044)	(0.039)	(0.017)	(0.016)	(0.014)	(0.046)	(0.930)
City × Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Founding Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5114	5114	5085	5114	5114	5114	5114	5114	531	77

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.5 Firm-Level Variable Construction

This appendix details the construction of the primary variables used in the analysis, derived from the raw data files (`organizations.csv`, `people.csv`, `degrees.csv`, and `jobs.csv`).

Sample Filtering Filtered to include only firms founded between 2000 and 2023. Founder and job samples are restricted to individuals associated with the filtered firms.

Founding Team and Executive Identification Founders and executives are identified by applying regular expression (regex) patterns to job titles. Then we classify an individual as a founder if the cleaned title contains the pattern `founder`, `cofounder`, or `founding` (regex-based matching). Next, we construct five executive flags from the cleaned title (e.g., contains `chief`+`officer`, `executive`+`officer`, any C.O title such as CEO/CFO/CTO, or phrases such as `executive director`). We also flag non-executive board-type titles (e.g., `non-executive`, `executive committee`) and subtract these from the executive count. Let `ExecSum` denote the resulting net executive score; then

$$\texttt{Executive} = \mathbb{1}\{\texttt{ExecSum} > 0\}.$$

For job records, we also set `Executive` to one when `job_type` is `executive`.

Demographic Proxy Construction Because gender and ethnicity are often missing in raw data, they are proxied using name-based algorithms:

- *Gender*: First names are processed through the `gender_guesser` package. Results are mapped to "male" or "female," while "mostly_male" and "mostly_female" are collapsed into their respective categories.
- *Ethnicity*: Full names are processed using the `ethnicolr2` package to predict ethnic groups based on the highest probability assigned to the name parts.

Education Level and Specialization Degrees are mapped to a five-point ordinal scale: `PhD`, `Masters`, `Bachelor`, `College`, and `Other`. The degree `subject` field is classified via keyword matching:

Science (STEM)

'biolog', 'chemi', 'physi', 'math', 'statistics', 'auto', 'robotics', 'tronics', 'electric', 'quant', 'neuro', 'biochemistry', 'neuro', 'cognitive', 'zoology', 'genetics', 'earth', 'environment', 'ecology', 'dna', 'photonics', 'medic', 'geology', 'astronomy', 'biomedical', 'biophysics', 'engineering', 'computer', 'mechanic', 'civil', 'software', 'robotics', 'systems', 'technology', 'informatics', 'network', 'ai', 'data', 'computer', 'aero', 'defense', 'geomatics', 'system', 'cyber', 'telecom', 'iot', 'machine learning', 'artificial intellilgence', 'medicine', 'psychology', 'clinical', 'biomedical', 'pharmacy', 'nutrition', 'denti', 'olog', 'disease', 'anatomy', 'health', 'pathology', 'neurobiology', 'physi', 'natural science', 'nano', 'traffic', 'ee', 'ece', 'comput', 'cogsci', 'medical'.

Business

'business', 'commerce', 'marketing', 'management', 'finance', 'accounting', 'strategy', 'administration', 'entrepreneur', 'supply chain', 'leadership', 'analytics', 'economics', 'mba', 'corporate', 'econ', 'executive', 'entrepreneurship', 'vc', 'venture capital', 'pe', 'm&a', 'hr'.

Degree information is recorded at the founder–firm level and may vary over time. For each founder-firm pair, we retain the highest degree attained as of the firm's founding date. In addition, we construct indicators for whether the founder has ever obtained a business degree or a science degree, regardless of whether it is associated with the highest degree.

Experience and Career History Prior experience is calculated by aggregating job history relative to the firm's founding date. `Has Experience Before` is a binary indicator for any job record prior to the current founding role. `Entrepreneurial Experience` is a binary indicator for whether the individual has previously served as a founder and the cumulative count of previous founding roles.

Geographic Proximity to CDL Sites Distances are calculated using the `geopy` library. Latitude and longitude for cities are retrieved via the `Nominatim` geocoder. Geodesic distance (in kilometers) is calculated between the firm location and nine CDL benchmark sites: Toronto, Vancouver, Montreal, Calgary, Halifax, Oxford, Paris, Atlanta, and Wisconsin. `Closest_CDL_Site` identifies the nearest CDL location, and `Closest_CDL_Distance` records that minimum distance.

Website Validation To determine firm activity, homepage URLs are probed asynchronously using the `requests` library. A site is marked as active if a HEAD or GET request returns a status code < 400. The `Last-Modified` header is extracted where available to proxy the date of the firm's last digital update.

Stream Mapping Keywords An industry tag is assigned to a CDL stream if the stream name or description matches that of a stream in the program list. If no direct match is found, the following semantic keyword map is applied. Each stream is associated with the corresponding set of keywords:

- *AI:* artificial intelligence; machine learning; computer vision; robotics; automation; ai
- *Health:* health; biotechnology; pharmaceutical; medical; life science; digital health
- *Blockchain:* blockchain; cryptocurrency; ethereum; bitcoin; distributed ledger; web3
- *Energy:* energy; battery; renewable energy; oil and gas; electric vehicle; energy storage
- *Space:* space; aerospace; satellite; satellite communication
- *Quantum:* quantum; quantum computing
- *Supply Chain:* supply chain; logistics; warehousing; shipping
- *Oceans:* marine; oceans; aquaculture
- *Climate:* cleantech; green; sustainability; environmental engineering; climate
- *Agriculture:* agriculture; farming; agtech; horticulture
- *Fintech:* fintech; payments; insurance; banking; financial services; wealth management; insurtech
- *Neuro:* neuro; neuroscience; brain; mental health

An industry tag is classified as `tech` if it satisfies either of the following conditions: (i) it contains any of the long-form technology phrases listed below (substring matching), or (ii) it contains any of the listed acronyms as a standalone word (exact word-boundary matching).

Technology phrases: "technology", "cloud computing", "cloud infrastructure", "cloud security", "saas", "paas", "machine learning", "artificial intelligence", "blockchain", "cryptocurrency", "web3", "quantum computing", "autonomous vehicles", "electric vehicle", "augmented reality", "virtual reality", "robotics", "drones", "wearables", "satellite communication", "big data", "analytics", "predictive analytics", "cyber security", "biotechnology",

"digital health", "health diagnostics", "telemedicine", "renewable energy", "energy storage", "smart home", "smart cities", "cleantech", "fintech", "insurtech", "digital banking", "mobile apps", "developer apis", "gaming", "social media marketing", "online games", "digital media", "content delivery network", "e-commerce platforms", "digital entertainment", "automation", "automative"

Acronyms (whole-word matching): "ai", "ar", "vr", "ev", "iot", "5g"

An industry tag is classified as "other" if it matches none of the above keywords and doesn't belong to any of the steam.

IPOs and Acquisitions While IPOs and acquisitions represent canonical "successful" exits, many transactions do not disclose sales price, effectively low-value exits. Consistent with prior work, we treat no-price deals as low-price sales by imputing a conservative $1\times$ Multiple on Invested Capital (MOIC), which implies zero net gain to investors. We merge exit records from Crunchbase and Refinitiv SDC Platinum M&A, assign $1\times$ MOIC where prices are missing, and then apply MOIC thresholds, from $1\times$ to $4\times$, to classify exits as successful (MOIC $\geq$ threshold) or unsuccessful (MOIC $<$ threshold).

Interim Success For firms not yet exiting, we adopt early-stage success proxies. These include total funding raised, number of funding rounds, number of funding rounds led by VC investors, backing by a top-tier investor within three years of founding, as well as securing a follow-up round within two years.

Failure We use Crunchbase's operating status-specifically, firms listed as inactive or closed-as our primary failure indicator. We then supplement this with web-traffic data: any company whose website remains persistently inactive or exhibits chronically low traffic is likewise classified as a failure.

A.6 Firm-Level Variable Construction

This appendix details the construction of the primary variables used in the analysis.

Outcome variables.

- *Firm entry*: the number of new firms founded in each city-industry-year cell, by founder type where indicated.
- *Entrant composition*: the share of entrants in a cell that are of a given founder type.
- *Aggregate financing*: total financing raised by startups in a city-industry-year or city-year cell, in USD.
- *Number of firms receiving financing*: the count of startups receiving at least one disclosed financing round in a cell.
- *Early-stage capital*: total capital raised by a firm within three years of founding, used in the capital-value calculation of Section 5.2.

Treatment variables.

- D_{cst} *(launch)*: an indicator equal to one for all periods after the launch of a given industry stream in a focal city, and zero for city-industry pairs that are not yet treated or never treated. Each city-industry can be launched only once, and treatment is absorbing.
- $\tilde{A}_{cs}$ *(pre-launch local industry size)*: the log stock of existing local firms in the city-industry at the time of the launch, centered at the mean launched cell, so that negative values denote industries that were emerging locally and positive values denote established ones. This variable is the empirical counterpart of "how visible was this industry already?" and is the key moderator throughout the paper.
- *Treatment intensity*: the cumulative number of CDL launches in a city, used in the city-level analyses of the founding margin.

Sample Filtering Filtered to include only firms founded between 2000 and 2023. Founder and job samples are restricted to individuals associated with the filtered firms.

Founding Team and Executive Identification Founders and executives are identified by applying regular expression (regex) patterns to job titles. Then we classify an individual as a founder if the cleaned title contains the pattern `founder`, `cofounder`, or `founding` (regex-based matching). Next, we construct five executive flags from the cleaned title (e.g., contains `chief`+`officer`, `executive`+`officer`, any C.O title such as CEO/CFO/CTO, or phrases such as `executive director`). We also flag non-executive board-type titles (e.g., `non-executive`, `executive committee`) and subtract these from the executive count. Let `ExecSum` denote the resulting net executive score; then

$$\texttt{Executive} = \mathbf{1}\{\texttt{ExecSum} > 0\}.$$

For job records, we also set `Executive` to one when `job_type` is `executive`.

Demographic Proxy Construction Because gender and ethnicity are often missing in raw data, they are proxied using name-based algorithms:

- *Gender*: First names are processed through the `gender_guesser` package. Results are mapped to "male" or "female," while "mostly_male" and "mostly_female" are collapsed into their respective categories.
- *Ethnicity*: Full names are processed using the `ethnicolr2` package to predict ethnic groups based on the highest probability assigned to the name parts.

Education Level and Specialization Degrees are mapped to a five-point ordinal scale: `PhD`, `Masters`, `Bachelor`, `College`, and `Other`. The degree `subject` field is classified via keyword matching:

Science (STEM)

```
'biolog', 'chemi', 'physi', 'math', 'statistics', 'auto', 'robotics',
'tronics', 'electric', 'quant', 'neuro', 'biochemistry', 'neuro', 'cognitive',
'zoology', 'genetics', 'earth', 'environment', 'ecology', 'dna', 'photonics',
'medic', 'geology', 'astronomy', 'biomedical', 'biophysics', 'engineering',
'computer', 'mechanic', 'civil', 'software', 'robotics', 'systems',
'technology', 'informatics', 'network', 'ai', 'data', 'computer', 'aero',
'defense', 'geomatics', 'system', 'cyber', 'telecom', 'iot', 'machine
learning', 'artificial intellilgence', 'medicine', 'psychology', 'clinical',
'biomedical', 'pharmacy', 'nutrition', 'denti', 'olog', 'disease', 'anatomy',
```

'health', 'pathology', 'neurobiology', 'physi', 'natural science', 'nano', 'traffic', 'ee', 'ece', 'comput', 'cogsci', 'medical'.

Business

'business', 'commerce', 'marketing', 'management', 'finance', 'accounting', 'strategy', 'administration', 'entrepreneur', 'supply chain', 'leadership', 'analytics', 'economics', 'mba', 'corporate', 'econ', 'executive', 'entrepreneurship', 'vc', 'venture capital', 'pe', 'm&a', 'hr'.

Degree information is recorded at the founder–firm level and may vary over time. For each founder-firm pair, we retain the highest degree attained as of the firm's founding date. In addition, we construct indicators for whether the founder has ever obtained a business degree or a science degree, regardless of whether it is associated with the highest degree.

Experience and Career History Prior experience is calculated by aggregating job history relative to the firm's founding date. `Has Experience Before` is a binary indicator for any job record prior to the current founding role. `Entrepreneurial Experience` is a binary indicator for whether the individual has previously served as a founder and the cumulative count of previous founding roles.

Geographic Proximity to CDL Sites Distances are calculated using the `geopy` library. Latitude and longitude for cities are retrieved via the `Nominatim` geocoder. Geodesic distance (in kilometers) is calculated between the firm location and nine CDL benchmark sites: Toronto, Vancouver, Montreal, Calgary, Halifax, Oxford, Paris, Atlanta, and Wisconsin. `Closest_CDL_Site` identifies the nearest CDL location, and `Closest_CDL_Distance` records that minimum distance.

Website Validation To determine firm activity, homepage URLs are probed asynchronously using the `requests` library. A site is marked as active if a HEAD or GET request returns a status code < 400. The `Last-Modified` header is extracted where available to proxy the date of the firm's last digital update.

Stream Mapping Keywords An industry tag is assigned to a CDL stream if the stream name or description matches that of a stream in the program list. If no direct match is

found, the following semantic keyword map is applied. Each stream is associated with the corresponding set of keywords:

- *AI:* artificial intelligence; machine learning; computer vision; robotics; automation; ai
- *Health:* health; biotechnology; pharmaceutical; medical; life science; digital health
- *Blockchain:* blockchain; cryptocurrency; ethereum; bitcoin; distributed ledger; web3
- *Energy:* energy; battery; renewable energy; oil and gas; electric vehicle; energy storage
- *Space:* space; aerospace; satellite; satellite communication
- *Quantum:* quantum; quantum computing
- *Supply Chain:* supply chain; logistics; warehousing; shipping
- *Oceans:* marine; oceans; aquaculture
- *Climate:* cleantech; green; sustainability; environmental engineering; climate
- *Agriculture:* agriculture; farming; agtech; horticulture
- *Fintech:* fintech; payments; insurance; banking; financial services; wealth management; insurtech
- *Neuro:* neuro; neuroscience; brain; mental health

An industry tag is classified as `tech` if it satisfies either of the following conditions: (i) it contains any of the long-form technology phrases listed below (substring matching), or (ii) it contains any of the listed acronyms as a standalone word (exact word-boundary matching).

Technology phrases: "technology", "cloud computing", "cloud infrastructure", "cloud security", "saas", "paas", "machine learning", "artificial intelligence", "blockchain", "cryptocurrency", "web3", "quantum computing", "autonomous vehicles", "electric vehicle", "augmented reality", "virtual reality", "robotics", "drones", "wearables", "satellite communication", "big data", "analytics", "predictive analytics", "cyber security", "biotechnology", "digital health", "health diagnostics", "telemedicine", "renewable energy", "energy storage", "smart home", "smart cities", "cleantech", "fintech", "insurtech", "digital banking", "mobile apps", "developer apis", "gaming", "social media marketing", "online games", "digital media", "content delivery network", "e-commerce platforms", "digital entertainment", "automation", "automative"

Acronyms (whole-word matching): "ai", "ar", "vr", "ev", "iot", "5g"

An industry tag is classified as "other" if it matches none of the above keywords and doesn't belong to any of the steam.

IPOs and Acquisitions While IPOs and acquisitions represent canonical "successful" exits, many transactions do not disclose sales price, effectively low-value exits. Consistent with prior work, we treat no-price deals as low-price sales by imputing a conservative 1× Multiple on Invested Capital (MOIC), which implies zero net gain to investors. We merge exit records from Crunchbase and Refinitiv SDC Platinum M&A, assign 1× MOIC where prices are missing, and then apply MOIC thresholds, from 1× to 4×, to classify exits as successful (MOIC $\geq$ threshold) or unsuccessful (MOIC $<$ threshold).

Interim Success For firms not yet exiting, we adopt early-stage success proxies. These include total funding raised, number of funding rounds, number of funding rounds led by VC investors, backing by a top-tier investor within three years of founding, as well as securing a follow-up round within two years.

Failure We use Crunchbase's operating status-specifically, firms listed as inactive or closed-as our primary failure indicator. We then supplement this with web-traffic data: any company whose website remains persistently inactive or exhibits chronically low traffic is likewise classified as a failure.