The Effect of Disagreement on Learning: Evidence from Business Mentoring

Luca Gius
Bocconi University

Amir Sariri
Purdue University

November 9, 2026

Click here for the latest version

Abstract

What are the nature and consequences of expert disagreement in mentoring startups? We study this question at Creative Destruction Lab, an accelerator where exited entrepreneurs and investors advise science-based startups on which business objectives to prioritize, then debate them to reach consensus. In records of 13,148 objectives from 4,508 mentoring sessions of 1,762 startups, mentors disagree over at least one objective in more than 70% of sessions. Using quasi-random assignment of mentors who differ in how often they disagree, we find that disagreement increases the likelihood that founders pursue a different kind of activity than planned by more than 40%. About two-thirds of this shift is a step back from costlier, more committing activities, such as launching a product, to low-cost analysis such as market research. Founders complete objectives finalized after disagreement at least as often as the others. Disagreement is associated with less mentor support and investment interest, especially for later-stage ventures, but causal estimates do not confirm these associations. In mentors' public debates, the most intense disagreements are followed by dropping the venture from the program more often than milder ones, without a corresponding increase in reorientation. The results suggest that expert disagreement acts as an anti-commitment device, prompting ventures to reconsider commitments and return to analysis.

We thank Andy Wu, Ashish Arora, Gary King, Hong Luo, Jacqueline Lane, Joshua Gans, Joshua Krieger, Katherine Collins, Kevin Bryan, Mu-Jeung Yang, Ramana Nanda, Rebecca Karp, and Scott Stern for helpful comments, suggestions, and discussions. Thanks to participants at the 2026 Bayesian Entrepreneurship Conference and the SDA Bocconi DBA Research Evening, and to seminar participants at Harvard Business School. We acknowledge support from the Institute for Humane Studies and the Strategic Innovation Fund of the Federal Government of Canada. We thank the staff of Creative Destruction Lab for their generous support of this project. Melissa Naranjo provided excellent research assistance. We have used AI tools in this research.

1 Introduction

Advice is a relatively low-cost way for founders to learn about their opportunity before committing scarce resources to any one path (Sariri, 2025; Miller *et al.*, 2024). Relative to other forms of learning, however, advice is inherently subjective because advisors draw on their own experiences and mental models of what makes ventures succeed (Cohen & Koning, 2026). It is therefore natural for different advisors to hold genuinely different beliefs about whether and how a startup business can succeed (Felin & Zenger, 2017; Agrawal *et al.*, 2026). Such disagreement creates a trade-off. On one hand, disagreement can be valuable, because it may signal a promising opportunity that is not obvious to everyone (Gius, 2025) and that multiple viable paths to commercialization are plausible (Gans *et al.*, 2019), and considering competing alternatives can sharpen the assessment of their relative strengths and weaknesses (Eisenhardt *et al.*, 1997). On the other hand, founders must eventually concentrate attention, people, and resources around a chosen path (Gans *et al.*, 2019). Disagreement can create confusion and make advice harder to act on. Founders may struggle to reconcile conflicting recommendations (Cohen *et al.*, 2019), while mentors may be less willing to commit their own time and capital when there is no clear direction to which they would apply their support.

This paper aims to shed light on the nature and consequences of disagreement between experts who are the systematic suppliers of advice and capital to startups, namely, exited entrepreneurs, angel investors, and venture capitalists. Specifically, we ask three questions. How often do they disagree, and what do they disagree about? When they disagree, do founders change their planned objectives, and toward what kind of activity? And do mentors keep supporting the venture?

Answering these questions presents two empirical challenges. First, strategic disagreement among experts is often selectively observed, if at all. Organizations need not preserve the process through which a strategic recommendation is reached, especially when it is of a sensitive nature. And even when the process is preserved, some experts might prefer not to participate in the discussion or might otherwise withdraw support rather than express dissent. Second, disagreement is not randomly distributed across ventures. More ambiguous or novel ventures may be more likely to elicit disagreement, while the same characteristics may also shape subsequent decisions.

We address these questions at Creative Destruction Lab (CDL), a selective global accelerator in which mentors advise science-based startups through repeated working sessions. The program focuses on seed-stage startups, where strategic ambiguity and flexibility are both unusually high. These ventures face several plausible paths to commercialization but vary in the extent to which they have made costly commitments that prevent them from changing course. Ahead of each session, founders propose three objectives for the next eight-week cycle. On the morning of the session, a Lead mentor proposes revisions, and an assigned Critique mentor records a response to each. These assigned roles require two experts to take positions on the same objectives, with their recommendations recorded separately before the rest of the group meets. The group then finalizes those objectives, and mentors decide whether to commit their time to supporting the venture. Ventures receiving no support are usually cut from the program. This sequence allows us to observe disagreement before it is reconciled, as well as its consequences—namely,

whether it redirects founders' planned objectives and whether mentors continue to support the venture. For identification, we exploit quasi-random assignment of Critique mentors driven by logistical constraints that curtail the organization's control in matching ventures to mentors. Because mentors differ in their stable propensities to disagree, this assignment generates plausibly exogenous variation in ventures' exposure to mentors who are more likely to express their disagreement.

Looking behind the ultimately agreed-upon objectives, we find that disagreement between advising experts is common and substantive. Across 13,148 objectives from 4,508 mentoring sessions in four cohorts, the two assigned mentors disagree over 41% of objectives, and more than 70% of mentoring sessions contain disagreement over at least one objective. A disagreement means the Critique either materially changes the scope of the Lead's objective or replaces it with a different one.

Founders respond to contrasting advice by revising their plans—the finalized objectives they actually pursue. Following Sariri (2025), we categorize objectives into four conceptual categories: analysis, experimentation, implementation, and resource acquisition. When the assigned mentors disagree with each other, finalized objectives are more likely to redirect the venture toward a different type of activity than the one initially proposed by the founders. Using quasi-random exposure to disagreement-prone Critique mentors, we estimate that disagreement raises the probability of such reorientation by 19 percentage points, more than 40% of the average reorientation rate. About two-thirds of the estimated increase in reorientation is toward analysis, meaning low-cost, low-commitment learning activities such as developing a go-to-market strategy or conducting market research. Descriptively, much of this movement comes from objectives initially aimed at implementation and resource acquisition.

Disagreement is not resolved mechanically, for example by always replacing the Lead's recommendation with the Critique's. When the mentors disagree, the finalized objective follows the Critique's recommendation more often than the Lead's, and in more than two of five cases it matches neither mentor's activity. Nor is it immediately resolved in the morning session, when it first surfaces: morning disagreement is disproportionately associated with an afternoon public debate between the same two assigned mentors. Because mentors decide whether to keep supporting the venture, founders may accept revisions that they do not intend to pursue. However, we find that founders complete objectives finalized after a disagreement at least as often as those finalized after agreement. In other words, revisions do not appear to be cheap talk.

At the same time, disagreement is associated with less mentor support for some ventures. Ventures that elicit disagreement are about 15% more likely to be cut, receive 6% fewer mentoring hours, and attract 32% fewer expressions of investment interest. These associations are concentrated among ventures at or beyond the validation stage. A potential explanation is that later-stage ventures have existing commitments that make reorientation less appealing even as disagreement weakens conviction in the existing path. Causal estimates, however, do not confirm these associations. Estimates of effects on later financing and survival are likewise inconclusive on average.

To understand the nature of disagreement, we also study the staff notes and transcripts of the public debates between mentors. We label each disagreement by whether it is about value (the overall viability of the venture) or strategy. We find that sessions are four times as likely to contain

a debate about strategy as one about value. To proxy for the distance between opposing views, we also label each disagreement by intensity. The most intense disagreements are followed by cut more often than milder ones, without a corresponding increase in reorientation.

Together, these results suggest that disagreement among informed experts can act as an anti-commitment device, reopening proposed commitments and redirecting ventures toward learning. Weaker mentor support accompanies disagreement mainly among later-stage ventures and after the most intense public debates.

The paper makes three contributions. First, the paper contributes to research on how entrepreneurs make strategic choices under uncertainty (Ghemawat & Levinthal, 2008; Gans *et al.*, 2019; Novelli & Spina, 2024). Past work has shown that advice offers a relatively inexpensive and effective way to reduce uncertainty, as founders can learn from others' experiences and perspectives without first committing resources to pursuing each alternative (Sariri, 2025; Cohen & Koning, 2026). Yet the same diversity of perspectives often generates conflicting recommendations, raising questions about how founders work through them and how they shape strategic choices (Cohen *et al.*, 2019). We provide causal field evidence that disagreement redirects ventures' prioritized tasks toward low-cost, low-commitment learning. Descriptively, much of this additional movement comes from objectives originally aimed at implementation (e.g., launching a product) and resource acquisition (e.g., fundraising, hiring). In this sense, disagreement can act as an anti-commitment device.

Second, the paper contributes to research on how institutions evaluate and develop polarizing ideas (Liu & Kovács, forthcoming; Åstebro *et al.*, 2026; Lane *et al.*, working paper; Fehder & Murray, 2018). One stream of research has focused on the notion that disagreement among evaluators about an opportunity's value can be a marker of future potential (Gius, 2025), and how institutions can harness this signal through champion-based rules that favor selecting polarizing opportunities (Malenko *et al.*, 2024). But institutions—including accelerators and venture capital funds—often care not only about selecting valuable ventures, but also about helping them develop and scale. Our results show that (i) much disagreement concerns which direction a venture should take, rather than only whether it is valuable; (ii) disagreement has consequential effects on the strategic path of ventures, rather than solely on whether ventures are selected in or out of the program; (iii) while a second opinion may surface considerations overlooked in the initial review—as suggested by the greater adoption of the Critique's recommendations relative to the Lead's—working through disagreement often involves distilling competing recommendations rather than simply selecting one advisor's opinion (see also Miller *et al.*, 2024). Harnessing disagreement therefore involves more than choosing how to aggregate evaluations to select ventures. Even under a champion-based model, institutions must also consider how competing recommendations are reconciled and how they shape ventures' subsequent development.

Third, the paper contributes a direct, quantitative measure of strategic disagreement. Prior research has examined disagreement qualitatively (Eisenhardt *et al.*, 1997) or inferred it from subsequent actions (Klepper & Thompson, 2010; Wu, 2016). Part of the challenge is that judgments about an opportunity's value or whether to pursue it are recorded in evaluators' scores and expressions of investment interest (Kerr *et al.*, 2014; Howell, 2021; Gius, 2025), while competing views

about how to develop it are much less readily observed. We compare two advisors' recommendations about the same concrete, time-bound, measurable objectives before collective deliberation, directly observing differences in recommendations about how the venture should prioritize its scarce resources. We also systematically measure open debates among mentors, a setting close to the partners' meetings at which most venture capital firms make investment decisions by vote or consensus (Gompers *et al.*, 2020). These measures help bring theories grounded in heterogeneous beliefs about opportunities and how to create value to the data (Felin & Zenger, 2017; Gambardella *et al.*, 2026).

2 Conceptual Framework

In this section, we discuss three ways in which disagreement between expert mentors can influence early firm development, in terms of (i) whether founders change their planned objectives, (ii) which kind of activity they move toward, and (iii) whether mentors keep supporting the venture. On each, we highlight how the existing work points in opposite directions.

2.1 Reorientation

Expert advice is an important input into entrepreneurial learning (Sariri, 2025). A key reason is that it allows founders to draw on knowledge and experience beyond their own. But different advisors bring different perspectives (Cohen & Koning, 2026) and may therefore give conflicting advice (Cohen *et al.*, 2019). What happens when those perspectives conflict?

Each advisor's perspective may contain valuable information about the opportunity alongside idiosyncratic judgments. One possibility is that entrepreneurs interpret contrasting advice as "noisy" and less reliable. As such, they would rely more on their existing beliefs and be more likely to retain their proposed course.

Prior theoretical work nevertheless shows that conflicting judgments can be more—not less—informative than consensus, either because the combined evidence pins down the truth more precisely (Welch, 1939; Denrell, 2025) or because one evaluator recognizes an exceptional strength that others do not (Malenko *et al.*, 2024). Empirical evidence corroborates the notion that disagreement can carry useful information about an opportunity (Gius, 2025). Founders may therefore give conflicting advice greater weight and revise their existing beliefs more substantially. When the advice points away from the planned course of action, stronger updating can make reorientation to a different set of priorities more likely. However, reorientation may also be merely nominal. Founders who depend on the mentors' support may accept revisions they do not intend to pursue. If so, they would complete these objectives less often.

2.2 Reorientation to Analysis

Following Sariri (2025), we sort the activities a venture can prioritize into four conceptual categories, which Appendix Table A2 summarizes. Analysis and experimentation are activities

intended for generating information. Analysis refers to low-cost, low-commitment learning activities, consisting of search and planning, such as conducting market research or developing a product roadmap. Experimentation consists of higher-fidelity tests that are not commitment-free, such as validating a technology or product-market fit by running a pilot with a potential customer. Implementation and resource acquisition are not meant for learning. Implementation executes ideas, such as launching a product or getting new customers, and resource acquisition secures financial, human, and intellectual capital, such as raising money or hiring.

If disagreement prompts reorientation, it is not obvious what form it will take. One possibility is that working through competing recommendations helps the venture identify a particularly promising direction that can be immediately implemented. An alternative possibility is that the debate surfaces alternative directions without resolving which to pursue, prompting the venture to step back and re-evaluate its possibilities.

Consider a venture in our sample that proposed to finish developing two more sensors for its first product. One mentor agreed, but a second mentor proposed replacing the objective with research into whether consumers would pay the product's intended price. In the discussion that followed, the first mentor conceded having taken the market for granted and agreed that more validation was a good idea. The founders' finalized objective was to establish clear evidence of product demand, which CDL recorded as completed at the next session.

This example offers two lessons. First, a different perspective can help expand the set of alternatives considered, rather than only help evaluate a known set (Felin & Zenger, 2009; Miller *et al.*, 2024). This is useful because the set of strategic possibilities is large (Knudsen & Levinthal, 2007), founders often consider an overly limited subset of it (Gruber *et al.*, 2008), and advice can therefore help avoid premature satisficing (Cohen *et al.*, 2019).

Second, disagreement need not generate a clearly superior alternative that can be immediately adopted. Competing theories can expose assumptions that need examination (Ehrig & Schmidt, 2022), and uncertainty among plausible theories can make experimentation worthwhile before commitment (Camuffo *et al.*, 2024). As such, mentor disagreement might encourage ventures to step back into evaluating their options rather than outright committing to one of the surfaced paths.

2.3 Mentor Support

Advisors provide time, connections, and other resources alongside advice (Sariri, 2025; Hallen *et al.*, 2020). Compared with uniformly lukewarm assessments, polarized views are more likely to include at least one mentor willing to champion the venture despite others' skepticism. When one champion is sufficient for the group to support the venture, polarizing ventures should therefore be more likely to secure support (Malenko *et al.*, 2024).

Alternatively, the fact that others disagree can exacerbate even an optimistic mentor's doubts about the value of the company or the right path to pursue. For example, disagreement may signal downstream difficulty in obtaining resources such as financing from the market. As such, disagreement can weaken mentors' willingness to commit time, connections, and resources to moving forward.

To conclude, research on entrepreneurial strategy distinguishes analysis and choice as sequential and complementary tasks: analysis evaluates alternatives, while choice entails committing to a path (Gans *et al.*, 2019). The arguments above suggest two ways mentor disagreement may enter this process. First, disagreement may substitute for further analysis by helping identify a particularly promising direction, allowing entrepreneurs and mentors to commit to a path more quickly and confidently. Alternatively, disagreement may inject doubt, encouraging ventures to step back into analysis; in this case, disagreement acts as an anti-commitment device. The same doubt may also weaken the mentors' willingness to provide support. We turn to the empirical analysis to distinguish between these two views.

3 Empirical Setting: Creative Destruction Lab

The Creative Destruction Lab (CDL) is a selective, eight-month accelerator for science-based startups that operates inside leading business schools at seventeen sites across ten countries and admits ventures into specialized technology streams such as artificial intelligence, space, quantum, and climate. CDL admits ventures through a competitive process that includes an application, business assessments, and technical interviews, and assigns successful applicants to a stream-site pair we refer to as a track. Each track has its own pool of mentors, drawn from angel investors, venture capitalists, and exited entrepreneurs with sector-specific experience. Sariri (2025) describes the institution and its admission process in detail.

The pedagogical core of CDL is a sequence of four working sessions held roughly every eight weeks at which the founders of each venture and the mentors of their track meet to design three measurable business objectives to prioritize for the next eight-week cycle. A fifth and final graduation meeting brings all sites together in Toronto for venture demos and concludes the program year. In the week before each session, staff circulate an updated dossier that reports progress on the previously finalized objectives and founders' proposed objectives for the next eight weeks.

A session has three parts. In the morning, founders meet privately with small groups of their track's mentors in what CDL calls small group meetings (SGMs). CDL assigns two mentors to formal roles called "Lead" and "Critique", who meet the venture in separate small group meetings. The Lead proposes a first round of revisions to the objectives the founders proposed, and the Critique responds, indicating whether they agree with the Lead's suggested changes. Both mentors record their positions on the venture's objectives in the venture's Slack channel before the afternoon session.

In the afternoon, all mentors and founders in a track convene for the large room discussion, which we call the plenary, and proceed venture by venture. For each venture, the Lead introduces the venture and briefly summarizes the proposed revisions, the Critique responds with their position, and the discussion opens to the rest of the mentors in the track. The plenary is moderated by a business-school professor and concludes with finalizing the three objectives that founders agree to prioritize, in CDL parlance, "at the expense of everything else." CDL keeps staff notes and transcripts of each venture's discussion.

In the evening, the moderator asks all founders to leave the room, then asks all mentors, one venture at a time, whether they will commit four hours of their personal time over the upcoming cycle to help that venture achieve its finalized objectives. Ventures that receive no formal commitment from mentors are usually dropped from subsequent sessions. Ventures that survive the four working sessions graduate from CDL at the fifth meeting.

3.1 Sample Construction

We assemble our data from four CDL cohorts between 2018/19 and 2022/23 using two main sources. We exclude the 2019/20 cohort disrupted by the COVID-19 pandemic.

CDL maintains a large relational database that contains detailed records of every venture's participation, from before joining the program to tracking alumni ventures' market outcomes after it. Sariri *et al.* (2025) describe the database and the data collection pipeline. From these records we observe each venture's proposed and finalized objectives, the identities of all mentors, including Leads and Critiques, and the staff notes and transcripts of each plenary. We also observe which mentors make formal time commitments to each venture, survival outcomes at the end of each session and graduation, and the completion status of each finalized objective. We also obtain detailed funding data from CDL's internal tracking system. CDL has two reasons to track funding closely. First, ventures keep CDL apprised of funding rounds to obtain introductions to the organization's network of investors. Second, as a non-profit, CDL's top-line performance metric reported to its board is equity value created, which is tied to funding events. Thus, designated staff at each site are responsible for routinely following up with alumni and mentors for any financing updates. To further ensure completeness, we also scan and merge financing data from PitchBook and company websites.

Our first measure of disagreement comes from differences between the Lead and Critique views on which objectives to pursue before the following session. We extract these data from the export of all ventures' Slack channels. We then bridge venture identifiers from the relational database to Slack channel IDs using a hand-validated mapping (Appendix A.1). In total, our estimation sample contains 13,148 Lead–Critique pairs across 4,508 venture-sessions and 1,762 ventures.

We separately measure disagreement as captured in the debates that surface in the public plenary, which we extract from the staff notes and transcripts described above. Within the estimation sample, 3,844 venture-sessions across 1,609 ventures have a plenary record with at least one comment about the venture. We call this the plenary sample. See Appendix A.1 for a detailed account of sample construction and attrition.

3.2 Startups & Objectives

Table 1 describes the 1,762 startups in the estimation sample, the mentoring support they receive during the program, and the finalized objectives that we observe. Panel A shows that the startups in the sample are predominantly high-technology ventures with small founding teams. The median venture has two founders, 56% have at least one PhD founder, 92% have a prototype, and 57% use

Table 1: Summary Statistics of Startups and Finalized Objectives

	Mean	Median	SD	Min	Max
Panel A: Venture characteristics and outcomes (N = 1,762)					
Team Size	2.54	2.00	1.09	1.00	8.00
Any PhD Founder	0.56	1.00	0.50	0.00	1.00
Has Prototype	0.92	1.00	0.27	0.00	1.00
Patent IP	0.57	1.00	0.49	0.00	1.00
Validation Stage or Later	0.59	1.00	0.49	0.00	1.00
Has Revenue	0.45	0.00	0.50	0.00	1.00
Pre-program Capital (CAD M)	1.00	0.21	2.32	0.00	55.33
Graduated	0.39	0.00	0.49	0.00	1.00
Total Raised (CAD M)	3.61	0.00	18.12	0.00	431.66
Panel B: Venture-session outcomes (N = 4,508)					
Mentoring Hours	6.31	8.00	4.78	0.00	28.00
Investment Interest	0.35	0.00	1.28	0.00	18.00
Cut	0.21	0.00	0.40	0.00	1.00
Panel C: Finalized-objective characteristics and outcomes (N = 8,283)					
Broad Reorientation	0.45	0.00	0.50	0.00	1.00
Objective Completed	0.62	1.00	0.49	0.00	1.00
Founder-Control Score	0.67	0.87	0.37	0.00	1.00
Subjectivity Score	2.46	2.40	0.97	1.00	5.00

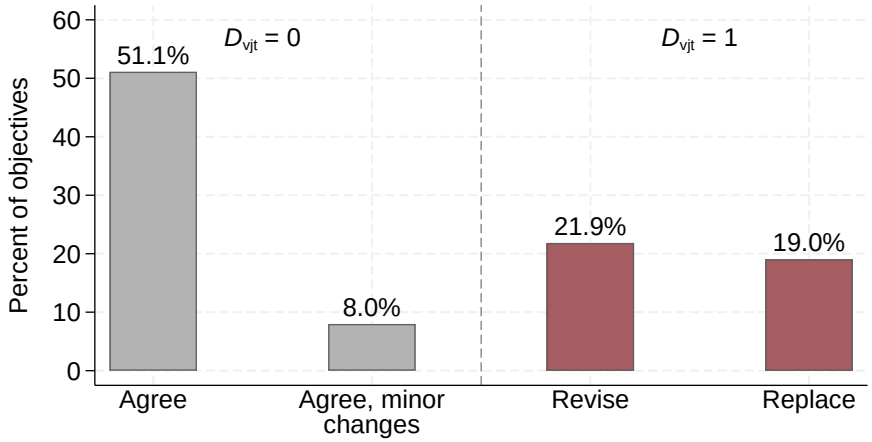
Notes: This table shows the characteristics of the estimation sample. Team size, any PhD founder, prototype, patent, stage, revenue, and pre-program capital come from the venture's application to CDL. Validation Stage or Later is 1 where the venture's development stage was validation, early revenue or profitable, and 0 where it was concept, technology or prototype. Pre-program capital and total raised are in millions of Canadian dollars.

or plan to use patents as an intellectual-property strategy. The ventures vary in their commercial development at entry, with 59% at the validation stage or later, and 45% reporting revenue. The financing measures are strongly right-skewed. Average pre-program capital is CAD $1.00 million, compared with a median of CAD $0.21 million. By the end of the observation period, total funding recorded averages CAD $3.61 million, while the median is zero and the maximum is CAD $431.66 million.

Panel C of Table 1 describes the 8,283 finalized objectives with a classified Lead–Critique pair. An LLM assigns each objective one to three of the 34 granular activities described in the classification system of early-stage startup activities (Sariri, 2025), of which we use the first, each mapped to one of the four conceptual categories: analysis, experimentation, implementation, and resource acquisition. We also measure how likely each objective is to be under the founders' control and how subjective the objective is, using a supervised machine learning method (Appendix A.3).[1] The average founder-control score is 0.67, and the average subjectivity score is 2.46 on the one-to-five scale.

[1] We thank Hong Luo and Laura Katsnelson for their contributions to developing the original version of these scores.

Figure 1: Distribution of Lead–Critique Classifications



Notes: This figure shows the distribution of the Critique's responses to the Lead's objectives. The unit of analysis is a classified Lead–Critique pair of the estimation sample. The dashed line separates the two agreement classifications from the two disagreement classifications.

4 Measuring Mentor Disagreement

Our first measure of disagreement is an indicator that, for each objective of each venture, equals 1 when the Critique revises or replaces the Lead's objective. We call this assigned-pair disagreement.

Our second measure captures debates in the afternoon plenary. Unlike assigned-pair disagreement, public disagreement surfaces organically, since mentors are not required to debate the same topic or objective, or to disagree at all. Appendix Figure C1 provides an anonymized example of a Lead–Critique disagreement carried into the plenary.

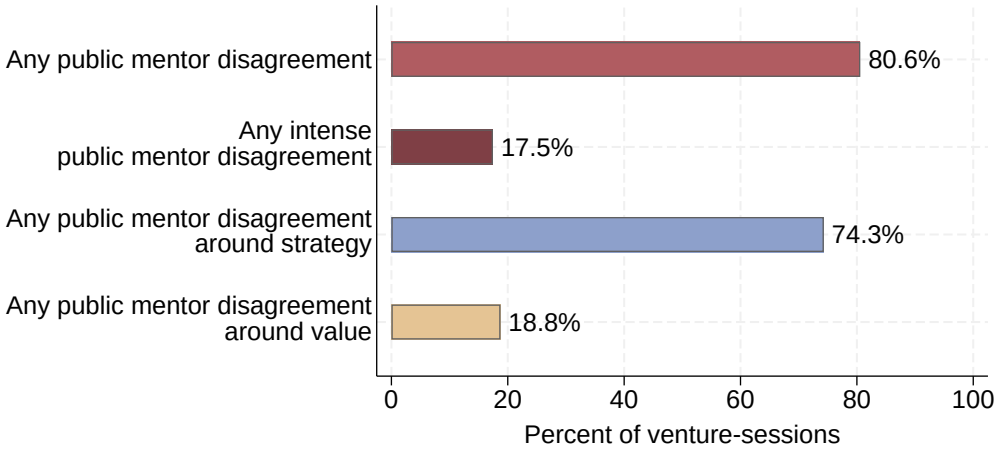
4.1 Assigned-Pair Disagreement

For each venture v in session t, we observe the Lead's proposed change to each objective j and the Critique's response to it. We use a large language model (LLM) to first extract the Lead and Critique views from the JSON exports of company Slack channels, then use another LLM to classify the relationship between the two positions as Agree, Agree with minor changes, Revise, or Replace. The mentors agree when the response is Agree or Agree with minor changes, and they disagree when it is Revise or Replace. We denote disagreement over objective j by D_{vjt}. For session-level analyses, D_{vt} equals one when the Lead and Critique disagree over at least one objective.

Appendix Table C1 gives examples from the data and Appendix Figures A1 and A2 show the exact prompts used. Appendix A.2 reports a hand audit of both steps.

We find that assigned-pair disagreement is common. The Lead and Critique disagree over 40.9% of the objectives. Aggregated to the venture-session level (D_{vt}), the Lead and Critique disagree over at least one objective in 70.7% of venture-sessions. Figure 1 shows that substantive disagreement is divided almost evenly between revision and replacement, which account for 21.9 and 19.0% of objectives, respectively.

Figure 2: Prevalence, Intensity, and Topic of Public Mentor Disagreement



Notes: This figure shows the share of venture-sessions with at least one debate of each kind between mentors. Intense debates have intensity four or five out of five. A venture-session can contain both a strategy and a value debate.

Appendix Figure C2 shows that the Lead and Critique disagree at similar rates whatever the category of the founders' proposed objective, and somewhat more often for analysis. The aggregate prevalence is therefore not driven by one category of venture activity.

4.2 Public Disagreement During the Plenary

For each venture-session in the plenary sample, we observe the staff notes and transcripts of the plenary discussion. We use an LLM to first extract each debate from this record, identifying the issue under dispute, the two opposing positions, and the speakers on each side with their roles. The LLM also rates each debate's intensity on a five-point scale, as the share of the founding team's working time over the next two months that would be spent differently under the two sides' courses of action. We then use the same LLM to classify each debate by whether the two sides disagree on the venture's overall viability (value debates) or the path it should take (strategy debates).

Appendix Table C2 gives an example of each type from the data and Appendix Figures A3 and A4 show the exact prompts used.

The LLM extracts 6,576 debates, and we count the 5,196 with at least one mentor on each side. Of these 5,196 debates, 84.1% are about strategy, 15.2% are about value, and the remaining 0.8% are about neither. Figure 2 shows that 80.6% of venture-sessions contain at least one debate, 74.3% a strategy debate, and 18.8% a value debate. At least one intense debate, with intensity four or five, occurs in 17.5% of venture-sessions.

Appendix Figure C3 shows how intense the debates are and how often a founder takes a side in those with a mentor on each side.

5 Empirical Strategy

We estimate linear regressions relating Lead–Critique disagreement to two sets of outcomes: the support ventures receive and their continuation in the program, and the objectives they prioritize. Table 1 reports their means.

The first set of outcomes is observed for every venture-session. A venture is cut when CDL records it as dropped from the program after session t. Mentoring hours are the hours that business mentors commit to assist the venture before the next session. Investment interest is the number of mentors who indicate interest in taking a financing meeting with the venture. Graduation indicates whether CDL classifies the venture as a graduate. For these outcomes, we estimate

$$Y_{vt} = \beta D_{vt} + \alpha_{c(v,t)} + \varepsilon_{vt},$$

where $\alpha_{c(v,t)}$ are track-session fixed effects, which compare ventures of the same cohort, site, technology stream, and session, and so absorb differences in the available mentor pool and the mix of admitted ventures. Graduation is a venture-level outcome, so we weight each venture-session by one over the venture's number of sessions in the sample.

The second set of outcomes is observed for each finalized objective. Reorientation equals one when a finalized objective belongs to a different conceptual category from the objective the founders proposed. Crossing a category boundary captures a substantive change in the kind of work the venture will undertake. Completion equals one when the finalized objective is completed before the next session. For these outcomes, we estimate

$$Y_{vjt} = \beta D_{vjt} + \alpha_{c(v,t)} + \lambda_{a(vjt)} + \varepsilon_{vjt},$$

where $\lambda_{a(vjt)}$ denotes fixed effects for the conceptual category of the founders' proposed objective. The proposed-category fixed effects account for differences across categories in both the frequency of disagreement and the probability that an objective is reoriented. For completion, every specification additionally includes the proposed objective's founder-control and subjectivity scores, to control for differences across objectives in how hard the task is to perform. Standard errors are clustered by venture.[2]

5.1 Causal Inference

Some ideas are inherently more polarizing than others, for example because they are more novel or uncertain (Felin & Zenger, 2009; Gius, 2025). Yet the same characteristics that make these ideas polarizing also make it difficult to isolate the consequences of disagreement. Greater uncertainty

[2] Standard errors reflect variation induced by both the venture-session-level assignment of Critique mentors and the venture-level sampling of organizations admitted to CDL. Because our conclusions are intended to apply to comparable ventures beyond those observed in the four sample cohorts, and the observed ventures constitute a small fraction of that broader population, conventional venture-clustered standard errors are asymptotically correct under this sampling interpretation (Abadie *et al.*, 2023).

about the best path forward may increase entrepreneurs' propensity to experiment and revise their strategy (Murray & Tripsas, 2004), while novel ideas may be more likely to be discounted by experts (Zuckerman, 1999; Boudreau *et al.*, 2016; Criscuolo *et al.*, 2017), independently of the disagreement they generate. The association we measure between disagreement and outcomes therefore need not identify a causal effect of disagreement.

To address this issue, we use an instrumental variable (IV) strategy that isolates variation in disagreement arising from the mentor CDL assigns as Critique. Mentors differ in their propensity to disagree, and we can estimate these differences from their repeated assignments as Critiques to other ventures. These differences provide a useful source of variation if Critique assignment is unrelated to unobserved venture characteristics. While CDL seeks to match each venture with mentors it deems effective, strict logistical constraints such as the limited number of mentors and the roles each must fill force compromises. CDL staff prioritize Lead assignments, then draw the Critique from the mentors who remain available. To gauge how much room this process leaves, we simulate it for a planner who knows how well every mentor fits every venture. Appendix Table B1 shows that the planner usually gives a venture its best-fit Lead but rarely its best-fit Critique. This process motivates treating Critique assignment as quasi-random within a track-session.

Assigned Critiques IV: We operationalize this idea using a cluster jackknife instrumental variable approach (CJIVE) (Chyn *et al.*, 2025; Frandsen *et al.*, 2025). Our preferred Assigned Critiques specification uses the assigned Critique's estimated propensity to disagree as an instrument for Lead–Critique disagreement. For each venture v, we estimate these propensities by regressing disagreement D_{us} on indicators for the assigned Critique W_{us}, using sessions s of other ventures $u \neq v$,

$$\widetilde{D}_{us} = \widetilde{W}'_{us}\pi + \nu_{us}, \qquad u \neq v,$$

where tildes denote variables net of track-session fixed effects.

Operationally, we first partial out these fixed effects from disagreement and the mentor-assignment indicators using the full assignment sample. Then, we estimate the regressions leaving out all observations belonging to the focal venture. For objective-level outcomes, we repeat this procedure using objective-level disagreement, also partialling out proposed-category fixed effects and leaving out all objectives belonging to the focal venture. The resulting coefficients, $\widehat{\pi}_{-v}$, capture differences in mentors' propensities to disagree, estimated from their behavior as Critiques when assigned to other ventures. The predicted disagreement for venture v in session t is

$$Z_{vt} = \widetilde{W}'_{vt}\widehat{\pi}_{-v},$$

where Z_{vt} is the instrument for disagreement we use in two-stage least squares. This specification compares the propensity to disagree of the focal venture's Critique with those of the other Critiques assigned in the same track-session.[3]

[3] The assigned Critique must have at least four classified venture-sessions involving other ventures, counting service as either Lead or Critique across the estimation sample. Both specifications also require at least two eligible mentors, including the assigned Critique, who meet the same history requirement.

Eligible Mentors IV: Since a venture's Lead cannot also be its Critique, the mentors who could critique it depend on whom CDL chose as its Lead. If CDL tends to give certain ventures Leads who often disagree, those ventures are left with Critiques less prone to disagree. The Assigned Critiques instrument would then partly reflect CDL's choice of Lead rather than the draw of the Critique. As a robustness check, the Eligible Mentors instrument compares the assigned Critique with the mentors who could have been drawn for that venture, namely those who attended a small-group meeting in the same track-session that morning. For each venture-session, we first identify this pool of M_{vt} eligible mentors, which includes the assigned Critique and excludes the venture's Lead and mentors without sufficient histories. We then replace the Critique indicators W_{vt} with centered indicators: the assigned Critique's entry is $1 - 1/M_{vt}$, each other eligible mentor's entry is $-1/M_{vt}$, and mentors outside the pool get zero. In a pool of four mentors, for example, the assigned Critique's entry is $3/4$ and each other mentor's entry is $-1/4$. We estimate mentors' propensities and the instrument exactly as before, but use centered indicators in place of W in both equations. The instrument then equals the assigned Critique's estimated propensity to disagree minus the average estimated propensity of the M_{vt} mentors in the pool.

Both specifications use the same assignment sample of 3,573 venture-sessions for venture-session outcomes and 6,526 finalized objectives in 2,254 venture-sessions for objective-level outcomes. We report corresponding OLS estimates on that sample.

5.2 Instrument Validity

Our IV estimates can be interpreted as treatment effects under four conditions (Chyn *et al.*, 2025): (i) a strong first stage, (ii) Critique assignment within a track-session is quasi-random, (iii) the exclusion restriction, meaning that Critique assignment affects outcomes only through its effect on Lead–Critique disagreement, and (iv) average monotonicity, meaning that, for any given venture or objective, the average disagreement propensity of Critiques who would disagree is at least as high as that of Critiques who would agree.

Appendix B examines each condition, and we report its key results here. First, the first stage is strong, with Kleibergen–Paap first-stage F-statistics of 126.3 for cut and 129.9 for reorientation outcomes. Second, the assigned Critique's disagreement propensity is unrelated to venture characteristics, individually and jointly, with the exception of a small negative correlation with venture stage. Third, we detect no relationship between the instrument and whether the assigned Critique commits time to the venture they are assigned to. Fourth, the first stage is positive across cohorts, sessions, sites, and proposed categories, consistent with average monotonicity. Lastly, the results are qualitatively similar under other minimum mentor histories, with two-sample IV, and with instruments that leave out the focal venture-session rather than the entire venture.

Table 2: Broad Reorientation and Completion Following Lead–Critique Disagreement

	OLS	First Stage	Second Stage	OLS	Second Stage
DV=	2-1 Broad Reorientation	2-2 Disagreement	2-3 Broad Reorientation	2-4 Objective Completed	2-5 Objective Completed
Disagreement	0.146***		0.191**	0.026**	0.137*
	(0.013)		(0.078)	(0.012)	(0.079)
Critique's disagreement propensity		0.387***			
		(0.034)			
Comparison set	—	Assigned Critiques	Assigned Critiques	—	Assigned Critiques
N	6,526	6,526	6,526	6,526	6,526
Mean of DV	0.452	0.402	0.452	0.619	0.619
Kleibergen–Paap rk Wald *F*		129.9	129.9		131.9
Anderson–Rubin *p*			0.017		0.080
Track-Session FE	X	X	X	X	X
Proposed Category FE	X	X	X	X	X
Proposed Text's Scores				X	X

Notes: This table shows the relationship between disagreement, and broad reorientation and completion of the finalized objective. The unit of analysis is a finalized objective of a venture-session in the assignment sample. The proposed text's two scores are the subjectivity and founder-control scores of the proposed objective's text, measured before the session. Column 5's first stage is nearly identical to column 2's. Standard errors clustered by venture are in parentheses. Statistical significance is *(10%), **(5%), or ***(1%).

6 Results

6.1 Reorientation and Completion

We first examine whether disagreement prompts ventures to reconsider their proposed course of action. Columns 1 and 3 of Table 2 show that disagreement is associated with a 15 percentage point (pp) higher probability of reorientation in OLS and raises it by 19 pp in IV estimates. Panel A of Appendix Table C3 shows that the findings are robust to adding venture fixed effects–comparing more and less polarizing objectives within the same venture–and to using the Eligible Mentors IV. Appendix Table C4 shows that disagreement is also associated with less near-verbatim adoption of the founders' proposed objectives.

We further examine completion to infer the level of effort and resources founders commit to different objectives. Columns 4 and 5 of Table 2 show that disagreement is associated with a 2.6 pp increase in completion, while the IV estimate implies an increase of 13.7 pp. The IV estimate is imprecise, however, and its 95% confidence interval includes the OLS estimate.[4] Panel B of Appendix Table C3 shows that the findings are robust to using the Eligible Mentors instrument, adding venture fixed effects, and controlling for finalized objective characteristics

[4] There are two possible explanations for the larger IV results. First, OLS estimates may be downward biased because more ambitious proposals may be both more polarizing and, whether kept or revised, harder to complete. Second, the larger IV estimate could reflect treatment-effect heterogeneity. IV identifies the effect for objectives that spark disagreement for some Critiques but not others. An additional perspective may be particularly useful for these objectives, whereas objectives that divide almost every mentor may remain difficult to pursue.

and category. Appendix Table C5 shows that disagreement is also associated with the finalized objective's characteristics, so these controls may themselves respond to disagreement. Since the IV estimate is similar within the finalized category, the higher completion is not explained by disagreement moving objectives into categories that are easier to complete. These results suggest that reconsidering plans in response to disagreement need not hinder ventures' progress toward their objectives.

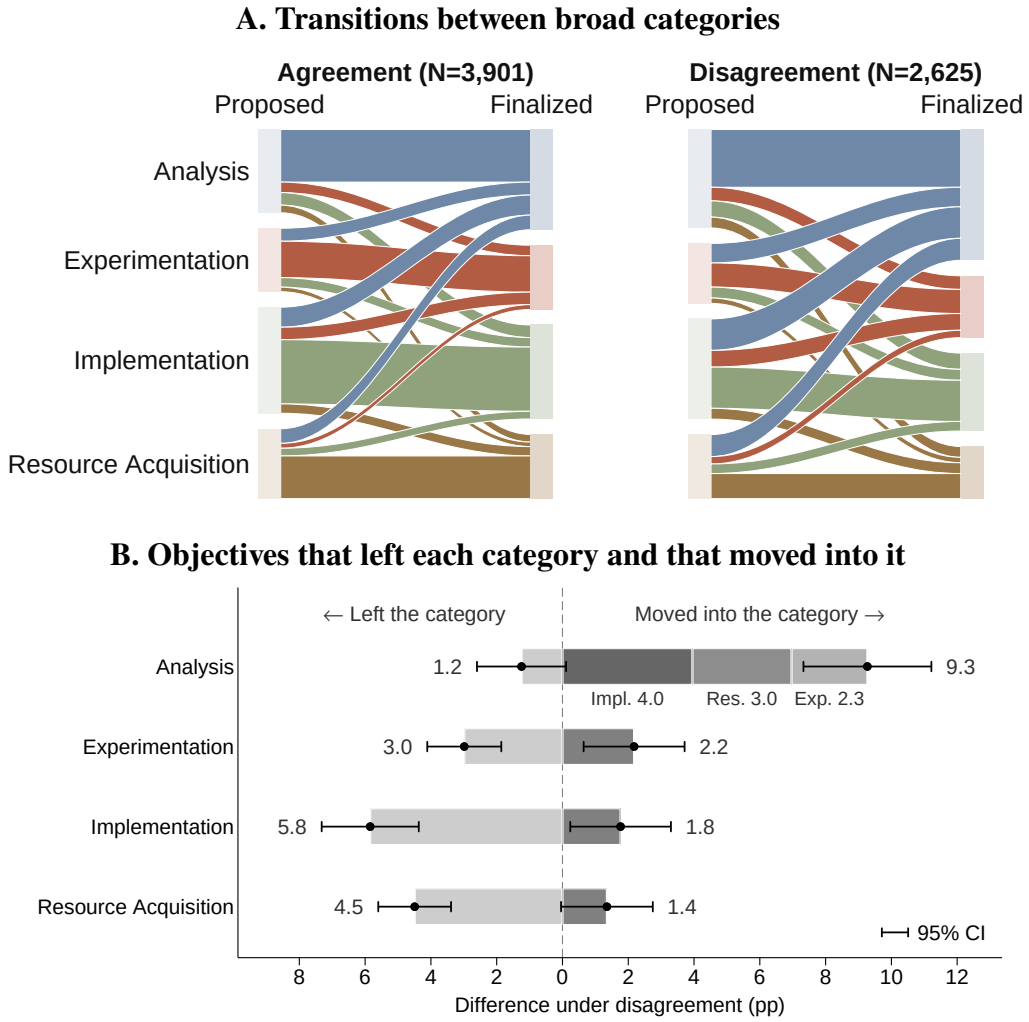
Figure 3 shows the origins and destinations of the reorientation that follows disagreement. Panel A traces objectives from their proposed to their finalized categories, separately under agreement and disagreement. Panel B shows that, under disagreement, objectives move mostly out of implementation (5.8 pp) and resource acquisition (4.5 pp), which together account for about 70% of the 15 pp increase. Most of the increase goes into analysis, 9.3 pp or 64%. Implementation and resource acquisition supply 7.0 of the 9.3 pp that move into analysis (75%). Descriptively, this pattern suggests that disagreement nudges ventures to re-examine their assumptions, more than bringing to light an alternative they can implement immediately. Overall, these patterns are consistent with disagreement acting as an anti-commitment device, bringing objectives back into analysis when founders had proposed implementation or resource acquisition. Appendix Figure C4 shows that the move into analysis, mostly from implementation and resource acquisition, is significant with either instrument and holds with venture fixed effects. For illustration, Appendix Figure C5 shows that when founders' proposed objective to raise money sparks disagreement, the finalized objective is less likely to keep raising money and more likely to turn to the analytical task of preparing for financing (e.g., preparing financial or pitch documents) or to another activity.

Panel A of Figure 4 shows that among objectives on which the Lead and Critique disagree, finalized objectives more often match the Critique's recommended category than the Lead's. This may reflect an advantage of being second, potentially identifying information overlooked in the initial review. However, finalized objectives match neither mentor's category in 23.9% of cases–a share that rises to 41.9% when distinguishing objectives at the finer 34 activity level. Finalized objectives thus often differ from both recommendations, a pattern consistent with mentors working together to distill coherent objectives from contrasting advice, potentially through the afternoon's public discussion. Panel B of Figure 4 shows that objectives matching the Critique's recommendation have an estimated completion rate at least as high as those matching the Lead's.

6.2 Mentor Support

We next examine whether disagreement affects mentors' willingness to commit time and resources to a venture. Here, disagreement is measured at the venture-session level and indicates disagreement over at least one proposed objective. We find that disagreement is associated with weaker support. The OLS estimates in Panel A of Table 3 show that ventures eliciting disagreement are 3.1 pp more likely to be cut, receive 0.36 fewer mentoring hours per session and 0.12 fewer expressions of investment interest–reductions of 6% and 32% relative to their respective sample means–and are 3.2 pp less likely to graduate. The association between disagreement and weaker support is concentrated among later-stage ventures. Figure 5 shows that, for these ventures, disagreement

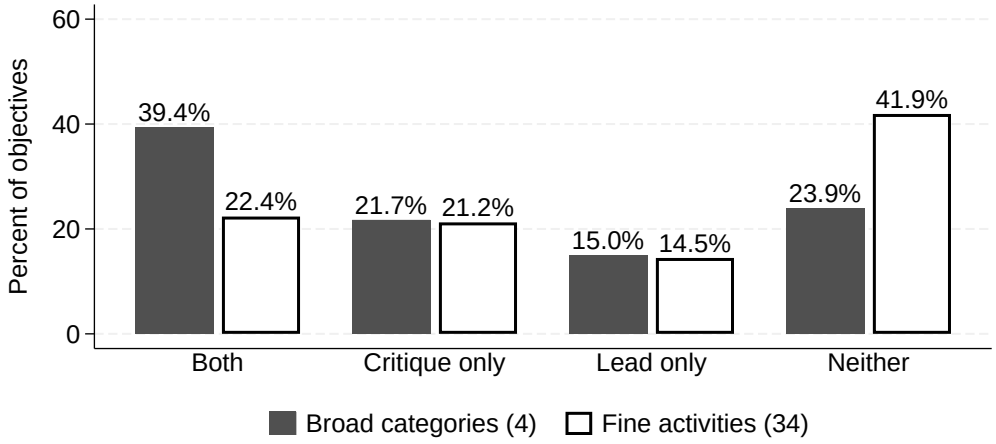
Figure 3: The Origins and Destinations of Reorientation



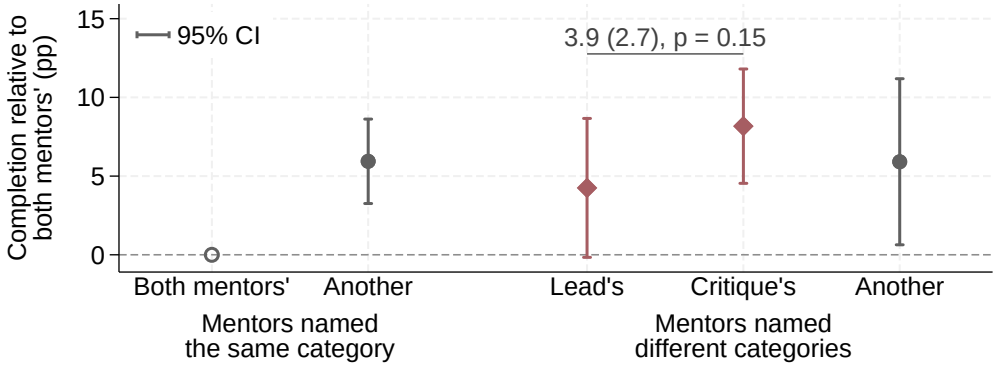
Notes: This figure shows the origins and destinations of reoriented objectives. The unit of analysis and the sample are those of Table 2. Panel A reports the raw transitions. Each estimate in panel B comes from an OLS regression that stacks four rows per objective, one per category, of the outcome on Disagreement, with the coefficient and the track-session and proposed-category fixed effects specific to each category. Estimates are in percentage points of all objectives. For the bars to the left, a row's outcome is 1 where the proposed objective is in that category and the finalized one is in another. For the bars to the right, it is 1 where the finalized objective is in that category and the proposed one is in another. For analysis's segments, it is 1 where the proposed objective is in the segment's category and the finalized one is in analysis. Standard errors are clustered by venture.

Figure 4: Finalized Objective Category by Lead and Critique Category & Completion

A. Whether the finalized objective matches the Lead's or the Critique's category



B. Completion of finalized objectives matching each mentor's category



Notes: This figure shows which mentor's category the finalized objective matches, and completion by match. Its unit of analysis is a finalized objective of the classified sample. Among objectives with disagreement, Panel A shows the fraction of finalized objectives' categories matching the objectives suggested by the Lead and by the Critique, by whether objectives are categorized into broad conceptual or fine categories. Panel B shows point estimates in pp from one regression of completion on indicators for whether the broad category of the finalized objective matches the Lead's, the Critique's, neither's or, when both mentors chose one category, another category, relative to the category both mentors chose. The regression includes track-session and proposed-category fixed effects and the proposed text's two scores. Standard errors are clustered by venture.

Table 3: Mentor Support Following Lead–Critique Disagreement

DV=	3-1 Cut	3-2 Mentoring Hours	3-3 Investment Interest	3-4 Graduated
Panel A: OLS				
Disagreement	0.031*	−0.361**	−0.122**	−0.032*
	(0.016)	(0.176)	(0.053)	(0.018)
Panel B: IV, Assigned Critiques				
Disagreement	−0.005	0.784	0.313	0.045
	(0.082)	(0.936)	(0.200)	(0.092)
First stage: Critique's	0.370***	0.370***	0.370***	0.378***
disagreement propensity	(0.033)	(0.033)	(0.033)	(0.037)
Kleibergen–Paap rk Wald F	126.3	126.3	126.3	104.9
Anderson–Rubin p	0.951	0.401	0.114	0.626
N	3,573	3,573	3,573	3,573
Mean of DV	0.208	6.400	0.376	0.414
Track-Session FE	X	X	X	X

Notes: This table shows the relationship between disagreement and mentor support. The unit of analysis is a venture-session in the assignment sample. Graduated weights each venture-session by the inverse of the venture's number of sessions. Standard errors clustered by venture are in parentheses. Statistical significance is *(10%), **(5%), or ***(1%).

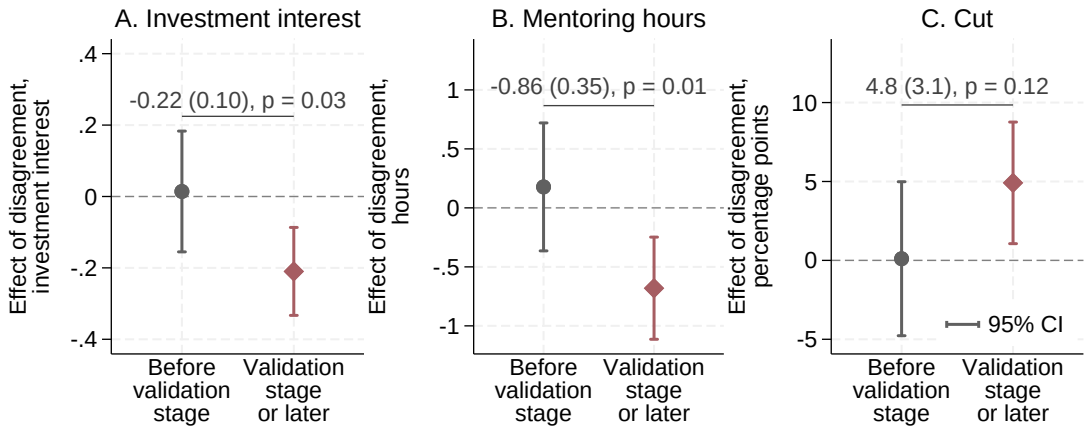
is associated with more cuts, fewer mentoring hours and less investment interest, while among ventures entering before validation the associations are close to zero.

Panel B of Table 3 shows that the IV estimates are too imprecise to confirm the associations with cut, mentoring hours and graduation, and rule out a reduction in investment interest as large as the association. If anything, the IV point estimates suggest that disagreement might *increase* rather than decrease support. Appendix Table C6 shows similar, less precise IV estimates with the Eligible Mentors instrument and with venture fixed effects. Appendix Tables B2 and B3 show the same without the Critique's own mentoring hours and, for cut, with a two-sample IV. The IV estimates in Panel D of Appendix Table C7 also suggest that disagreement has less favorable effects on support for later-stage ventures. The differences between earlier- and later-stage ventures are significant only for being cut from the program, at the 10% level, and for eventually graduating, at the 5% level. Before the validation stage, the estimated effects are favorable. Appendix Table C7 shows that disagreement is also associated with a lower chance of raising a financing round within 24 months, which the IV estimates do not confirm. CDL's record of later rounds is more complete for ventures that stay in the program. As with support, the IV estimate is more favorable before the validation stage. Survival to June 2026 shows no relationship with disagreement.

6.3 How Disagreement Is Resolved in the Plenary

The morning's disagreement often does not stay private. Figure 6 shows that a Lead and Critique who disagreed over objectives in the morning oppose each other in the afternoon plenary more

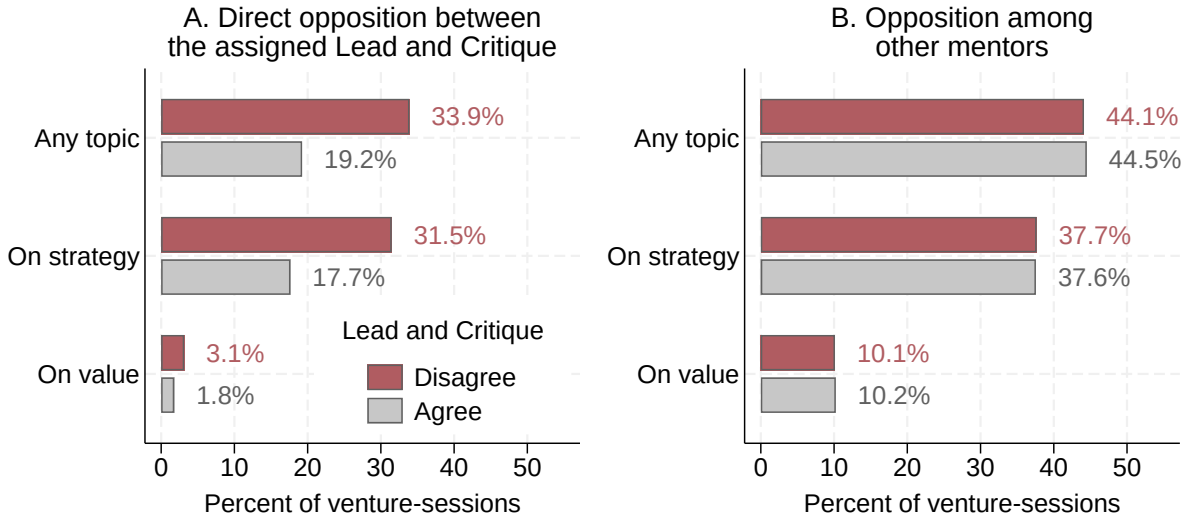
Figure 5: Mentor Support Following Disagreement, by Venture Stage



Notes: This figure shows the relationship between disagreement and mentor support, by venture stage. The unit of analysis is a venture-session in the assignment sample. Each panel shows OLS estimates from one regression of the outcome on Disagreement, Validation Stage or Later and their interaction, with track-session fixed effects (Appendix Table C7). The bracket reports the difference between the two stage groups. Standard errors are clustered by venture.

often, mostly in debates about the venture's strategy, while opposition among the other mentors does not differ. The plenary is where this disagreement becomes public. It lets us observe whether public disagreement concerns the venture's value or the direction it should take, and how it relates to the decisions to keep the venture in the program and which objectives to finalize.

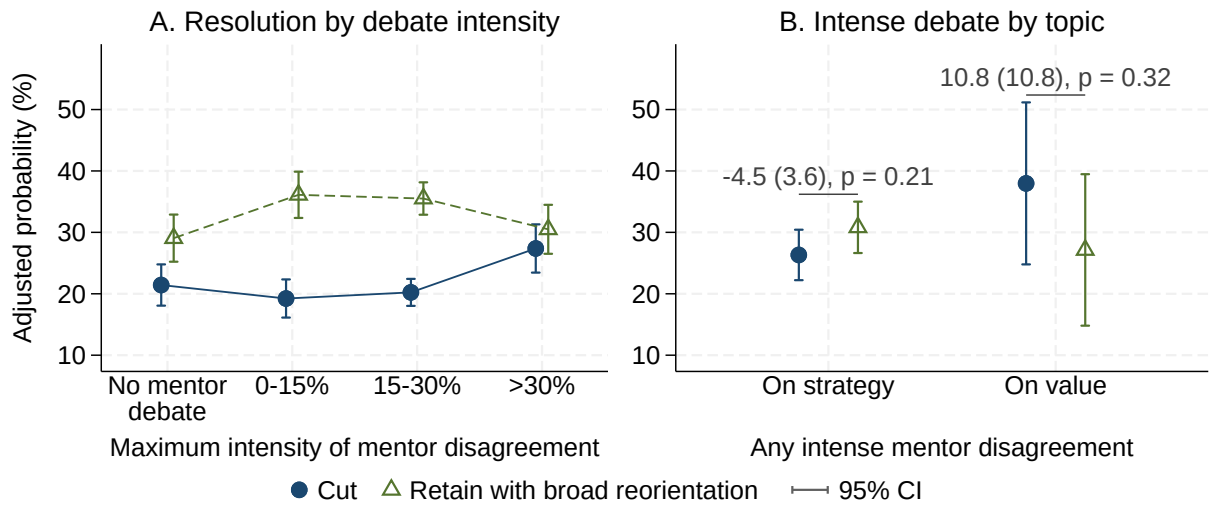
Figure 6: Assigned-Pair Disagreement and Public Opposition



Notes: This figure shows how often mentors take opposing sides in the plenary, by whether the morning's Lead and Critique disagreed. The unit of analysis is a venture-session of the plenary sample whose assigned Lead and Critique are on record and distinct. Panel A requires the assigned Lead and Critique to stand on opposing sides of the same public debate. Panel B first removes both assigned mentors from each debate, then requires a mentor to remain on each opposing side. Topic rows indicate at least one such debate on strategy or on value, and are nonexclusive.

First, we find that the probability of the venture being cut rises when plenary debates are most

Figure 7: Public Escalation and Venture Resolution



Notes: This figure shows the relationship between public disagreement and the venture's cut or retention with broad reorientation. The unit of analysis is a venture-session of the plenary sample at sessions one to three with the outcome on record. Each panel shows adjusted probabilities from linear probability models of each outcome on the plenary indicators and Lead–Critique disagreement, with track-session fixed effects. Retention with broad reorientation means the venture is retained and more than half of its finalized objectives changed conceptual category. Panel A groups sessions by the maximum intensity of debates with a mentor on each side. Panel B shows sessions with an intense debate, at level four or five with a mentor on each side, on strategy or on value, each estimated separately. The bracket reports cut minus reorientation. Standard errors are clustered by venture.

intense, but not at lower intensities. By contrast, we find no corresponding increase in reorientation. Panel A of Figure 7 shows that, among sessions with mentor debate at low or intermediate intensity (under 15% and 15–30%, respectively), reorientation is more likely than cut by approximately 15–17 pp, whereas at high intensity the gap narrows to about 3 pp and cut is about as likely as retention with broad reorientation.

Second, Panel B of Figure 7 shows that cut follows intense debates about the venture's value more often than reorientation does. Reorientation follows intense debates about strategy more often than cut does. In both cases, however, differences are not statistically significant.

One read of these results is that disagreement mostly redirects ventures rather than ending them, because most of it concerns strategy. Most public disagreement, including the part carried over from the morning, is about strategy, and intense strategy debates are followed by reorientation at least as often as by cut. Intense debates about the venture's value lean toward cut. This read matches our earlier results indicating disagreement causes reorientation and has no detectable effect on cut. The topic of debate does not explain why cut rises at the highest intensity, however. The rise in cut at the highest intensity remains when the intense value debates are set aside. One possibility is that under high intensity of disagreement, reconciling the two sides into one direction becomes harder, and dropping the venture becomes about as likely as redirecting it.

7 Discussion

Strategic disagreement is natural around entrepreneurial opportunities. Reasonable people may disagree about the value and the best way to commercialize new ideas (Felin & Zenger, 2017; Agrawal *et al.*, 2026; Gambardella *et al.*, 2026). This disagreement matters, especially in the early stages of a firm's development, when strategic choices are consequential (Agrawal *et al.*, 2021) and persistent (Kaplan *et al.*, 2009). It is then surprising how little we know about the nature and consequences of disagreement among experts who are looked upon by novice founders for guidance on how to build their companies. Part of the challenge is the difficulty of observing competing judgments among real experts. The key contribution of this paper is to measure strategic disagreement directly, at scale. We document, to our knowledge for the first time, key systematic patterns in when experts disagree and how it affects entrepreneurial decision-making.

First, we find that disagreement is common, both in private and in public, and centers primarily around the development choices that firms make. Second, we find that disagreement between mentors drives a large and significant change in how founders prioritize the use of their scarce time and resources. The largest share of this change is away from costly, high-commitment activities such as product development and hiring, to the least costly, least committing learning activities such as market research and hiring planning. Third, these changes do not stall founders' progress, as measured by success in achieving prioritized objectives. Fourth, disagreement is associated with weaker mentoring support, especially among later-stage ventures, although our causal estimates do not confirm this. Fifth, in mentors' public debates, the most intense disagreements are followed by cut more often than milder ones, without a corresponding increase in reorientation.

7.1 Implications

Entrepreneurs are constrained by the resources they can apply to reducing uncertainty, as well as by the partial commitments that arise from learning itself. Strategic disagreement can act as an anti-commitment device, shaping where ventures direct their efforts within the entrepreneurial choice process (Gans *et al.*, 2019). Founders seeking external advice should therefore consider whether their resources and existing commitments can accommodate taking a step back. Responding to conflicting recommendations also often involves distilling them into a coherent direction, rather than simply choosing which one to follow (Miller *et al.*, 2024). Although our evidence comes from an entrepreneurial setting, related lessons may also apply to managers seeking external advice on contentious issues (Van den Steen, 2018).

Our results also speak to institutions that evaluate and help develop novel ideas, including accelerators and venture capital funds. Advisors' tendency to disagree can shape ventures' development and is worth weighing in advisor selection alongside traits such as expertise. Organizational choices that favor selecting polarizing ventures may also tilt portfolios toward ventures that need more assistance with early-stage planning. Champion-based rules, commonly used by early-stage venture capital firms, empower optimistic evaluators to greenlight polarizing ideas despite others' skepticism (Malenko *et al.*, 2024). Finalized objectives often differ from both mentors' proposals,

which suggests that CDL's champion-based model—where one mentor can secure a venture's continued participation—does not eliminate the need to work through competing recommendations about how the venture should develop.

7.2 Limits and Research Ahead

A first limit of our study is that it leaves considerable uncertainty about the causal effect of disagreement on ventures' longer-term financial performance. Part of the reason comes from constraints associated with our causal design, which limit our power to detect downstream effects that are further removed from the session's immediate consequences. In part, this is because the two margins we focus on—whether ventures remain in the program and whether they reorient—may be intertwined in the long run. Losing continued support may constrain a venture's strategic options, while strategic reconsideration may change its resource needs and ability to attract backing. Disentangling these two margins and tracing out their effect on a venture's future success remains an important direction for future research.

A second limit of our study concerns the nature of disagreement that we focus on. Our causal estimates capture a weighted average of effects among ventures whose exposure to disagreement depends on which Critique they are assigned. But much of the interest in polarizing opportunities concerns disagreement rooted in the novelty or ambiguity of the idea itself (Felin & Zenger, 2017; Gambardella *et al.*, 2026)—ideas that are "inherently" contested and contrarian. While we expect our findings to be broadly informative about how disagreement affects the strategic development of these ideas as well, extending our estimates to them requires assumptions about how treatment effects vary across ventures.

Finally, our findings come from a particular institutional setting. At CDL, ventures are encouraged to follow strategic advice, and the resources they receive depend, in part, on doing so. Elsewhere, entrepreneurs may shop around for advice they are free to disregard. CDL also brings multiple mentors together to discuss ventures' plans and offer support, even though one champion can secure continued participation. In other settings, advisors may never interact, and a single backer may provide sufficient resources for the venture to proceed without cooperation from others. These institutional differences may shape how our findings generalize to other settings. Exactly which features of CDL—whether the fact that Lead and Critique must record positions on the same concrete objectives, or the afternoon plenary—help disagreements surface and become part of collective deliberation remains an open question.

References

Abadie, Alberto, Athey, Susan, Imbens, Guido W, & Wooldridge, Jeffrey M. 2023. When Should You Adjust Standard Errors for Clustering? *The Quarterly journal of economics*, **138**(1), 1–35.

Agrawal, Ajay, Gans, Joshua S., & Stern, Scott. 2021. Enabling Entrepreneurial Choice. *Management Science*, **67**(9), 5510–5524.

Agrawal, Ajay, Camuffo, Arnaldo, Gambardella, Alfonso, Scott, Erin L., Stern, Scott, & Gans, Joshua (eds). 2026. *Bayesian Entrepreneurship*. Cambridge, MA, USA: MIT Press.

Angrist, Joshua D., & Fernández-Val, Iván. 2013. ExtrapoLATE-ing: External Validity and Overidentification in the LATE Framework. *Pages 401–434 of:* Acemoglu, Daron, Arellano, Manuel, & Dekel, Eddie (eds), *Advances in Economics and Econometrics: Tenth World Congress*. Cambridge University Press.

Boudreau, Kevin J., Guinan, Eva C., Lakhani, Karim R., & Riedl, Christoph. 2016. Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science. *Management Science*, **62**(10), 2765–2783.

Camuffo, Arnaldo, Gambardella, Alfonso, & Pignataro, Andrea. 2024. Theory-Driven Strategic Management Decisions. *Strategy Science*, **9**(4), 382–396.

Chyn, Eric, Frandsen, Brigham, & Leslie, Emily. 2025. Examiner and Judge Designs in Economics: A Practitioner's Guide. *Journal of Economic Literature*, **63**(2), 401–439.

Cohen, Susan, & Koning, Rembrand. 2026. Advice and the Bayesian Entrepreneur. *Pages 231–246 of:* Agrawal, Ajay, Camuffo, Arnaldo, Gambardella, Alfonso, Gans, Joshua S., Scott, Erin L., & Stern, Scott (eds), *Bayesian Entrepreneurship*. Cambridge, MA: MIT Press.

Cohen, Susan L., Bingham, Christopher B., & Hallen, Benjamin L. 2019. The Role of Accelerator Designs in Mitigating Bounded Rationality in New Ventures. *Administrative science quarterly*, **64**(4), 810–854.

Criscuolo, Paola, Dahlander, Linus, Grohsjean, Thorsten, & Salter, Ammon. 2017. Evaluating Novelty: The Role of Panels in the Selection of R&D Projects. *Academy of Management Journal*, **60**(2), 433–460.

Denrell, Jerker. 2025. *When Does Variability Strengthen Bayesian Updating?* Working paper, University of Warwick.

Ehrig, Timo, & Schmidt, Jens. 2022. Theory-Based Learning and Experimentation: How Strategists Can Systematically Generate Knowledge at the Edge Between the Known and the Unknown. *Strategic Management Journal*, **43**(7), 1287–1318.

Eisenhardt, Kathleen M., Kahwajy, Jean L., & Bourgeois, L. J. 1997. Conflict and Strategic Choice: How Top Management Teams Disagree. *California Management Review*, **39**(2), 42–62.

Fehder, Daniel, & Murray, Fiona. 2018. Evaluation of Early-Stage Ventures: Bias across Different Evaluation Regimes. *Academy of Management Proceedings*, **2018**(1), 17519. Conference abstract.

Felin, Teppo, & Zenger, Todd R. 2009. Entrepreneurs as Theorists: On the Origins of Collective Beliefs and Novel Strategies. *Strategic Entrepreneurship Journal*, **3**(2), 127–146.

Felin, Teppo, & Zenger, Todd R. 2017. The Theory-Based View: Economic Actors as Theorists. *Strategy Science*, **2**(4), 258–271.

Frandsen, Brigham, Leslie, Emily, & McIntyre, Samuel. 2025. Cluster Jackknife Instrumental Variables Estimation. *The Review of Economics and Statistics*.

Gambardella, Alfonso, Gius, Luca, & Stern, Scott. 2026. Homo Entrepreneuricus. *Pages 63–86 of:* Agrawal, Ajay, Camuffo, Arnaldo, Gambardella, Alfonso, Gans, Joshua S., Scott, Erin L., & Stern, Scott (eds), *Bayesian Entrepreneurship*. Cambridge, MA: MIT Press.

Gans, Joshua S., Stern, Scott, & Wu, Jane. 2019. Foundations of entrepreneurial strategy. *Strategic management journal*, **40**(5), 736–756.

Ghemawat, Pankaj, & Levinthal, Daniel. 2008. Choice Interactions and Business Strategy. *Management Science*, **54**(9), 1638–1651.

Gius, Luca. 2025. Disagreement Predicts Startup Success: Evidence from Venture Competitions. *Strategy science*, **10**(2), 93–108.

Gompers, Paul A., Gornall, Will, Kaplan, Steven N., & Strebulaev, Ilya A. 2020. How Do Venture Capitalists Make Decisions? *Journal of Financial Economics*, **135**(1), 169–190.

Gruber, Marc, MacMillan, Ian C, & Thompson, James D. 2008. Look Before You Leap: Market Opportunity Identification in Emerging Technology Firms. *Management Science*, **54**(9), 1652–1665.

Hallen, Benjamin L, Cohen, Susan L, & Bingham, Christopher B. 2020. Do Accelerators Work? If So, How? *Organization science (Providence, R.I.)*, **31**(2), 378–414.

Howell, Sabrina T. 2021. Learning from Feedback: Evidence from New Ventures*. *Review of Finance*, **25**(3), 595–627.

Kaplan, Steven N., Sensoy, Berk A., & Strömberg, Per. 2009. Should Investors Bet on the Jockey or the Horse? Evidence from the Evolution of Firms from Early Business Plans to Public Companies. *The Journal of Finance*, **64**(1), 75–115.

Kerr, William R., Lerner, Josh, & Schoar, Antoinette. 2014. The Consequences of Entrepreneurial Finance: Evidence from Angel Financings. *The Review of financial studies*, **27**(1), 20–55.

Klepper, Steven, & Thompson, Peter. 2010. Disagreements and Intra-industry Spinoffs. *International Journal of Industrial Organization*, **28**(5), 526–538.

Knudsen, Thorbjørn, & Levinthal, Daniel A. 2007. Two Faces of Search: Alternative Generation and Alternative Evaluation. *Organization Science*, **18**(1), 39–54.

Lane, Jacqueline N., Rietzler, Nathan, & Zhang, Yuqi. working paper. *Noise or Signal? Professional Backgrounds and Prediction of Startup Success*. Working paper; undated author webpage.

Liu, Chengwei, & Kovács, Balázs. forthcoming. *Smart Contrarians: Thinking Differently about Thinking Differently*. Harvard Business Review Press.

Malenko, Andrey, Nanda, Ramana, Rhodes-Kropf, Matthew, & Sundaresan, Savitar. 2024. Catching Outliers: Committee Voting and the Limits of Consensus When Financing Innovation. *Journal of Finance*. Forthcoming.

Miller, Amisha, O'Mahony, Siobhan, & Cohen, Susan L. 2024. Opening the Aperture:

Explaining the Complementary Roles of Advice and Testing When Forming Entrepreneurial Strategy. *Organization Science*, **35**(1), 1–26.

Murray, Fiona, & Tripsas, Mary. 2004. The Exploratory Processes of Entrepreneurial Firms: The Role of Purposeful Experimentation. *Pages 45–75 of:* Baum, Joel A. C., & McGahan, Anita M. (eds), *Business Strategy over the Industry Life Cycle*. Advances in Strategic Management, vol. 21. Elsevier JAI.

Novelli, Elena, & Spina, Chiara. 2024. Making business model decisions like scientists: Strategic commitment, uncertainty, and economic performance. *Strategic Management Journal*.

Sariri, Amir. 2025. The Economics of Advice: Evidence from Start-up Mentoring. *Management science*, **71**(12), 10022–10046.

Sariri, Amir, Gatov, Evgenia, Neal, Geneva, Robinson, Kyle, Sennik, Sonia, Tham, Wei Yang, Vertolli, Michael, & Goldfarb, Avi. 2025. Database, Methodological Tools, and Research Opportunities: Creative Destruction Lab and Early-Stage Technology Ventures. *NBER Working Paper Series*.

Van den Steen, Eric. 2018. Strategy and the Strategist: How It Matters Who Develops the Strategy. *Management Science*, **64**(10), 4533–4551.

Welch, B. L. 1939. On Confidence Limits and Sufficiency, with Particular Reference to Parameters of Location. *The Annals of Mathematical Statistics*, **10**(1), 58–69.

Wu, Andy. 2016. Organizational Decision-Making and Information: Angel Investments by Venture Capital Partners. *Academy of Management Proceedings*, **2016**(1), 11043.

Zuckerman, Ezra W. 1999. The Categorical Imperative: Securities Analysts and the Illegitimacy Discount. *American Journal of Sociology*, **104**(5), 1398–1438.

Åstebro, Thomas, Funck, Andrew, & Giordano, Francesco. 2026. *Accuracy and Uncertainty in Venture Selection: The Roles of Experience and Complexity*. Working paper.

Appendix A Data Appendix

A.1 Sample Construction and Attrition

In this section we describe the sample construction and attrition from the raw data to the analysis tables and figures. Table A1 provides a high-level summary.

We use four cohorts of the CDL program: 2018/19, 2020/21, 2021/22, and 2022/23. We exclude the 2019/20 cohort, which the COVID-19 pandemic disrupted.

The four cohorts contain 2,177 admitted ventures. We exclude 415 of them for six reasons, applied in order so that each venture is excluded once: a session marked as paused (10); mentor commitments recorded on a different track than the one the venture was admitted to (31); no session attended (50); an incomplete objective record, meaning a session without a complete set of three proposed objectives or with an incomplete set of finalized ones (160); no attended session in which the Lead–Critique exchange could be classified (141); and a second admission of a company admitted in two cohorts (23). For the 27 companies admitted twice, we keep only the first admission. Of the 27 second admissions, 23 are dropped under this rule and the other four under earlier rules. We also drop 11 venture-sessions, at sessions one to three of ventures not cut, in which a proposed or finalized objective cannot be assigned an activity category, even by hand. In this case, the venture's other sessions are kept. The remaining 1,762 ventures attended 5,045 venture-sessions.

For 2020/21 to 2022/23, we extract the Lead–Critique exchange from each venture's Slack channel. We match each channel to a venture by company name within its site, cohort, and track. A research assistant resolved the ambiguous matches by hand and confirmed the 40 least certain automatic ones. Of the 1,772 ventures admitted in these cohorts, 1,663 (94%) have a matched channel. The other 109 drop out under the exclusions above, 53 of them because they have no classified Lead–Critique exchange. For 2018/19, research assistants coded the same exchange by hand from CDL's structured session records. Each exchange yields Lead–Critique pairs, meaning an objective the Lead posted and the Critique's response to it. Of the 5,045 venture-sessions, 4,508 have at least one classified pair. These 4,508 venture-sessions form the estimation sample. Of the other 537, 487 have a plenary record but no classified pair and 50 have neither. Those whose assigned Critique meets the assignment sample's history requirement enter only the reduced form of the two-sample IV estimates. From the estimation sample's 13,689 pairs, we drop 533 that are coded Unclear and 8 that could not be parsed, leaving us with 13,148 pairs.

Finalized objectives come from the following session's dossier. They are missing when a venture is cut, after the fourth session, and for 70 venture-sessions at sessions one to three of ventures that were not cut. The estimation sample has 8,616 finalized objectives. We drop the 333 without a classified Lead–Critique pair, leaving 8,283 in 2,872 venture-sessions. We call these the classified sample.

For the main results, OLS and IV alike, we use the 3,573 venture-sessions of 1,596 ventures whose Critique served as Lead or Critique in at least four venture-sessions of other ventures in the estimation sample, and for which CDL's attendance records list at least one mentor other than the Critique and the Lead with that history at the session's morning meetings. We call this the

assignment sample. The reorientation and completion results use its 6,526 finalized objectives, in 2,254 venture-sessions of 1,199 ventures.

For the plenary results, we use the 3,844 venture-sessions of 1,609 ventures whose plenary record contains at least one comment about the venture. Figure 7 uses the 3,144 of these, from 1,573 ventures, with the session's outcome on record, at sessions one to three.

Table A1: Sample Attrition in Stages

	Ventures	Venture-sessions	Objectives
Panel A: Exclusions			
Admitted ventures, four cohorts	2,177		
Session marked as paused	−10		
Commitments recorded on another track	−31		
No session attended	−50		
Incomplete objective record	−160		
No classifiable Lead–Critique exchange	−141		
Second admission of the same company	−23		
Objective without an activity category		−11	
Retained	1,762	5,045	
Panel B: Samples			
Estimation sample (Lead–Critique pairs)	1,762	4,508	13,148
Finalized objectives with a classified pair	1,353	2,872	8,283
Assignment sample	1,596	3,573	
Finalized objectives	1,199	2,254	6,526
Plenary sample	1,609	3,844	
Session outcome on record	1,573	3,144	

Notes: This table shows the construction of the analysis samples. Panel A's exclusions are applied in the order listed. The objective count of the estimation sample is its classified Lead–Critique pairs.

A.2 Hand-Coded Validations of LLM Extraction

A research assistant manually validated the extraction and classification of the Lead–Critique exchange against the original session records. The extraction audit covers 153 objective-level entries from the 2020/21 cohort, each pairing a Lead-proposed objective with the Critique's response. For each entry the audit checked the extracted Lead objective, the extracted Critique response, and whether each mentor was matched to the correct role. Table A3 shows that 96.7% of Lead objectives and 90.8% of Critique responses were extracted correctly. The observed failures reflect messages that depart from the standard three-objective format, such as a mentor posting a single consolidated comment. In 9 of 153 cases the Critique commented before the Lead posted, so the recorded response addresses the founders' original objectives rather than the Lead's revisions.

The same audit validated the five-category agreement labels: the four classifications of Section 4.1 and *Unclear*. Table A3 also shows that, among adjudicated pairs, the five-category label matched the auditor's in 86.4% of cases and the binary disagreement indicator in 91.2%.

Table A2: Four Conceptual Categories of Start-up Activities

Category	Features	Examples
Analysis	Low-commitment learning	Examine size of the market
	Standard templates	Develop product road map
	Noisier than experimentation	
Experimentation	High-commitment learning	Validate technology
	No standard template	Validate product-market fit
	Less noisy than analysis	
Implementation	Involves selecting ideas	Launch product
	Intent is not learning	Get new customers
Resource acquisition	Financial capital	Raise capital
	Human capital	Hire CEO
	Intellectual capital	Submit patent application

Notes: This table shows each conceptual category's key features and stylized examples (Sariri, 2025, p. 8).

Table A3: Human Validation of Lead–Critique Extraction and Classification

Check	Correct/Total	Percent
Lead extraction	148/153	96.7
Critique extraction	139/153	90.8
Five-category classification	127/147	86.4
Binary disagreement classification	125/137	91.2

Notes: This table shows how often a human auditor agrees with our extraction and classification of Lead–Critique pairs. The unit of analysis is a Lead–Critique objective pair, for 153 pairs in 61 venture-sessions of the 2020/21 cohort, 43 drawn at random and 18 challenging cases. The five categories are Agree, Agree with minor changes, Revise, Replace, and *Unclear*. The two classification rows use the 147 audited pairs that the auditor rated and that match exactly one pair in our data. The five-category row compares the paper's label with the auditor's on all 147. The binary row compares only whether each label means agreement or disagreement, so it drops the 10 pairs that the paper or the auditor labels *Unclear*.

A.3 Founder-Control and Subjectivity Scores

We score all proposed and finalized objectives for founder control and subjectivity with text models trained on human ratings. The founder-control score is the predicted probability that completing the objective is under founders' control. The subjectivity score is the predicted rating on a 1-5 scale, with higher values meaning a more subjective objective. For training, human raters scored 4,737 objective texts. Then, we fit two regressions on the words, word pairs, and short strings of three to five characters in each objective's text. A logistic regression predicts the founder-control flag. A linear regression predicts the subjectivity rating. To assess prediction quality, we test machine learning predictions on a held-out fifth of the rated ventures. The founder-control model agrees with the raters 81% of the time, and the subjectivity prediction is off by 0.7 points on average.

A.4 LLM Extraction and Classification Prompts

Figure A1: Objective Extraction Prompt

```
q1 = QuestionFreeText(
    question_text="""
    Below is a Slack transcript between CDL mentors, founders, and staff. The conversation should contain exactly one set of three objectives for each mentor role:

    - 1st set: Objective Setter, also known as Lead / Proposed New Objectives - 1st mentor's proposed revisions to the founders' proposed objectives
    - 2nd set: Critique, also known as Comment on Proposed New Objectives - second mentor's critique of the first mentor's suggested revisions.

    Task:

    - Find 1st (Objective Setter) and 2nd (Critique) objective sets, posted by mentors in that order.
    - Copy each objective verbatim (strip control characters like \n only). Do not invent content.
    - Do Not extract objectives proposed by the founder(s)/venture or objectives completed in the previous period.
    -- Mentors' display name often include ".Associate" or ".Fellow" and sometimes ".Guest".

    Formatting:

    - Output must be raw JSON WITHOUT the code fence ```json ```.
    - No additional notes or commentary before or after the json block.
    - Return a json array with two objects. Each object must have:
    -- "role": "Objective Setter" or "Critique".
    -- "user_id": Slack user ID responsible for that objective set.
    --- If the objectives appear in a consolidated message (e.g., CDL staff summarizing with <@U123> tags), use the tagged user ID.
    --- If the objectives come directly from the mentor's own post, use that poster's user ID (provided in the transcript header).
    -- "objectives": array of exactly three strings, preserving the original wording.

    Deduplication:

    - If the same role appears twice (e.g., a brief mentor post plus a later consolidated version), keep the most complete entry (usually the longer one containing numbered objectives).

    Example Output with the third objective of the Objective Critiquer missing:

    [
      {
        "role": "Objective Setter",
```

```
      "user_id": "U0XXXXXXXXX",
      "objectives": [
        "REPLACE - Rework company narrative ...",
        "REPLACE - Build model ...",
        "AGREE - Secure ..."
      ]
    },
    {
      "role": "Critique",
      "user_id": "U0YYYYYYYYY",
      "objectives": [
        "AGREE",
        "MODIFY - ...",
        ""
      ]
    }
  ]

  Transcript:

  {{ scenario.transcript }}
  """,
  question_name="leadcritique_objectives",
)
```

Notes: This figure shows the prompt that extracts the Lead's objectives and the Critique's response from each Slack transcript. The portion "{{ scenario.transcript }}" is a placeholder for the session's full Slack transcript. The model returns a JSON array with two objects, one for each mentor role, each containing the mentor's Slack user ID and the verbatim three-objective list. We run the prompt with Anthropic's Claude Sonnet 4.5 (claude-sonnet-4-5-20250929) at temperature zero.

Figure A2: Lead–Critique Agreement Classification Prompt

```
SYSTEM_PROMPT = """You are a precise, detail-oriented research assistant specializing in
    qualitative coding. Your role is to read the text of business objectives by two
    different startup mentors and decide the extent to which they agree or disagree with
    each other. Your expertise is in hand-coding text data for statistical analysis.

CONTEXT
You are rating the level of agreement between two business objectives set during a
    startup mentoring session at a technology accelerator program (Creative Destruction
    Lab). A lead mentor proposes an objective for the startup, and a second mentor (the
    critiquer) responds to that objective. Your task is to determine how much the
    critiquer agrees with the lead mentor's proposed objective.

RATING SCALE
Below are the five agreement levels. Read both objectives and select the option that best
     describes how much the critiquer agrees with the lead objective:

- Agrees (code 0): The critiquer explicitly agrees and proposes no substantive changes. (
```

```
    Minor punctuation/formatting allowed.)
  Example: Lead says "Validate product-market fit with 5 enterprise pilots" and critique
      says "AGREE".

- Agree with minor changes (code 1): The critiquer makes small edits (wording/clarity/
    scope tweaks) but the same business task is being done. If the critique says AGREE
    but adds a brief qualifier (e.g., "with emphasis on X", "with collection of feedback
    "), this is Agree with minor changes, not Revise -- unless the qualifier materially
    changes what work is required.
  Example: Lead says "Hire a VP of Sales" and critique says "AGREE - with focus on
      someone with enterprise SaaS experience".

- Revise (code 2): The critiquer keeps the same general business function but materially
    expands, narrows, or specifies it (e.g., adds a new constraint, metric, customer
    segment, channel, or timeline). If the critique keeps the same general business
    function (e.g., both are about stakeholder engagement) but changes who, how, or when,
     that is Revise, not Replace.
  Example: Lead says "Raise a Series A" and critique says "REVISE: Prepare a detailed
      financial model and pitch deck to approach Series A investors in Q3".

- Replace (code 3): The critiquer proposes a different business function entirely (e.g.,
    lead says "raise money", critique says "build a prototype").
  Example: Lead says "Develop a marketing strategy for North America" and critique says "
      REPLACE: Focus on building the MVP and getting first 3 paying customers before any
      marketing spend".

- Unclear (code 4): Either objective is missing/incomplete, or the critiquer's response
    does not contain a clear objective-level judgment or alternative. Also choose Unclear
     if BOTH the lead and critique objectives consist only of short tokens like "AGREE"
    or "TBD" with no substantive objective text.
  Example: Lead says "TBD" and critique says "AGREE".

COMMON EDGE CASES
- If the critique says "AGREE" followed by a brief qualifier, default to code 1 (Agree
    with minor changes) unless the qualifier fundamentally changes the business function.
- If the critique says "MODIFY" or "REVISE" but the actual objective text describes the
    same task with minor wording changes, rate as code 1 not code 2.
- If both objectives are about the same business function (e.g., both about customer
    validation) but the critique changes who, how, or when, that is Revise (code 2), not
    Replace (code 3).
- Mentors may prefix text with tokens like AGREE / MODIFY / REVISE / REPLACE / MERGE. Use
     the substantive content, not the prefix token, to determine agreement level.

Use the record_agreement tool to submit your rating. Provide your agreement_code (0-4)
    and your confidence (0.0-1.0) in this rating."""

USER_PROMPT = f"""Lead objective:
{lead}

```

```
Critique objective:
{critique}"""
```

Notes: This figure shows the prompts that rate the Critique's response on the five-point scale. The model receives the Lead-proposed objective and the Critique mentor's response, extracted via the prompt in Figure A1 or, for the 2018/19 cohort, hand-coded from CDL's session records, and returns an agreement code 0–4 along with a confidence score. The placeholders "{lead}" and "{critique}" are filled in per venture-session, with one classification call for each of the three objectives. We run the prompts with Anthropic's Claude Opus 4.6 (claude-opus-4-6) at temperature zero.

Figure A3: Plenary Disagreement Extraction Prompt

```
You are a detail-oriented research assistant reading the transcript of one venture's session at an accelerator.
     Mentors and the venture's founders are both in the room. Each comment carries a comment id and a speaker
    id.

# The roles and typical sequence of discussions:
An objective setter mentor presents their proposed objectives for the coming period, a critique mentor responds
     with any revisions, a chair moderates the follow-on discussion, and other mentors join in. The venture's
    founders may present and respond. The discussion may leave the objectives entirely and turn to other
    matters -- whether the business is viable, whether the room believes the technical results, and so on.
    Speaker ids do not say who is a founder, who is a mentor, and who is chairing; judge that from what each
    speaker says.

# Your job is to name the session's chair and identify its disagreements.

## The chair:
chair_speaker_ids: the speaker id of the person moderating this session -- the one who runs the sequence, calls
     on people, and closes topics. Judge it from what each speaker says. Always return a list: empty if no one
    is clearly moderating, and more than one id only if the role changes hands during the session.

## The disagreements. For each one, provide:
1. topic: a short label for the contested question
2. topic_description: one sentence stating the question at issue
3. brief_summary_of_why_contentious: one sentence on why the participants disagree
4. side_a: one position, what it implies the founding team would do, and its supporters
5. side_b: the opposing position, what it implies the founding team would do, and its supporters
6. disagreement_intensity: an integer from 1 to 5

# Before you answer any question, you MUST:
1. read the item definitions below.
2. never invent a position, a supporter, or an implied action that the transcript does not support.
3. respond in exactly the JSON output format below, and nothing else.

# Definition of disagreement (items 1 to 3):
## What is a disagreement:
A disagreement is a specific question about the venture -- its objectives, strategy, claims, plan, or prospects
     -- on which two or more participants state opposing positions in their own comments. Opposing positions on
     the same question count even when the speakers do not reply to each other. Report every disagreement, mild
     or strong; the disagreement_intensity item records how consequential it is. Founders are participants: a
    founder taking a side counts, whether the participant on the other side is a mentor or another founder.

## What is not a disagreement:
a proposal or revision that no one contradicts, including a critique mentor's revision of the proposed
    objectives that no one takes up; advice or a suggestion that no one stated the opposite of; two suggestions
     on different questions; a question or devil's-advocate prompt from the chair, when the chair does no more
    than prompt -- if the chair argues a position of their own, they are a participant like any other; the
    founders' presentation on its own.

## Reporting requirements:
1. Two arguments belong to the same disagreement if the same decision would resolve both; otherwise they are
    separate disagreements.
```

```
2. One stretch of argument may contain several disagreements. Judge by the decision at issue, not by whether
     the exchange ran together.
3. List the disagreements in the order in which each is first raised in the transcript.

# Sides, implied actions, and supporters (items 4 and 5):
Each side has a position (one sentence), an implied action, and its supporters.

## Implied action:
One sentence on what the founding team would spend their time doing over the next two months if the room
     settled on that side.
1. If a side means the founding team carries on with what it is already doing, say so and name what that is ("
     continue the current plan: X and Y"). That is an implied action, not an absent one.
2. Write exactly "not stated" only when the disagreement carries no implication for what the founding team does
      -- for instance a disagreement purely about a belief, where no one in the room said what should follow
     from it. Do not invent an action to fill the field.

## Supporters:
A supporter is a speaker id, that speaker's role, and the ids of the comments in which that speaker states or
     argues the position. A side may have any number of supporters. Repeat the supporter object once for every
     speaker who states or argues that side.
1. Use the speaker id exactly as printed. Do not substitute, merge, or renumber ids.
2. role is one of: founder, mentor, chair, unclear. Judge it from what the speaker says, not from the id.
3. A speaker's role holds for the whole session. Anyone you name in chair_speaker_ids has the role chair in
     every disagreement they appear in, including one where they argue a position of their own.
4. Cite a comment if that comment itself states or argues the position, whether or not it was prompted by a
     question. Do not cite procedural turns, summaries, questions, or comments that merely mention the topic,
     and do not cite a comment that only supplies a fact, only concedes, or only expresses doubt, even when it
     sits next to comments that do argue the position.
5. Speaker 0 marks a comment whose speaker was not recorded. Within one session, 0 may stand for several
     different people. Do not treat 0's comments as one speaker and do not infer anyone's role by combining them
     ; judge each 0 comment on its own content.

# Intensity (item 6):
Of the founding team's working time over the next two months, the share that would be spent differently under
     side A's implied action than under side B's:
1 = less than 5%
2 = 5 to 15%
3 = 15 to 30%
4 = 30 to 50%
5 = more than 50%
1. Estimate against the founding team's whole working time, not against the proposed objectives alone.
2. The same activity aimed at a different market, customer, or segment counts as a change of task.
3. The estimate does not depend on which side is labelled A and which is labelled B.
4. Each disagreement is estimated on its own. The shares across a session's disagreements need not sum to
     anything.
5. Where both sides' implied actions are "not stated", the disagreement moves no time and takes bin 1.

# Output format:
Return only valid JSON, no prose and no code fence, in exactly this structure. The // notes are constraints on
     each field; they are not part of the output.
{
  "chair_speaker_ids": [string],                  // ids of whoever moderates; [] if no one clearly does
  "topics": [
    {
      "topic":                            string,  // a short label for the contested question
      "topic_description":                string,  // one sentence
      "brief_summary_of_why_contentious": string,  // one sentence
      "side_a": {
        "position":       string,                 // one sentence
        "implied_action": string,                 // one sentence, or exactly "not stated"
        "supporters": [
          {
```

```
            "speaker_id":  string,                // exactly as printed in the transcript
            "role":        string,                // one of: founder | mentor | chair | unclear
            "comment_ids": [integer]              // ids of comments in this session only
          }
        ]
      },
      "side_b": {
        "position":       string,                 // one sentence
        "implied_action": string,                 // one sentence, or exactly "not stated"
        "supporters": [
          {
            "speaker_id":  string,                // exactly as printed in the transcript
            "role":        string,                // one of: founder | mentor | chair | unclear
            "comment_ids": [integer]              // ids of comments in this session only
          }
        ]
      },
      "disagreement_intensity": integer           // one of 1, 2, 3, 4, 5
    }
  ]
}
Repeat the object inside "topics" once per disagreement, and the object inside "supporters" once per speaker on
      that side. If the session has no disagreement, return an empty "topics" list; still report
     chair_speaker_ids.

Comments, one per line, in the form "- (ID <comment id>, Speaker <speaker id>): <comment>":
{conversation_text}
```

Notes: This figure shows the prompt that extracts the debates from each plenary record. The portion "{conversation_text}" is a placeholder for the session's comments, one per line, each with its comment and speaker ids. The model returns the session's chair and, for each debate, its topic, the two sides' positions, implied actions and supporters, each supporter's role, and an intensity from 1 to 5. We run the prompt with Anthropic's Claude Opus 5 (claude-opus-5) at its default settings.

Figure A4: Debate Topic Classification Prompt

```
You are a detail-oriented research assistant. You are reading one disagreement from a session at an accelerator
     , where mentors and the venture's founders discussed the venture's next objectives. The disagreement has
     already been extracted from the transcript: a topic, a one-sentence description, why it is contentious, and
      two sides, each with its position and the action it implies for the founding team. You do not see the
     transcript.

# Your job is to label what the two sides contradict each other on.

Label by the claim the two sides contradict each other on: an assessment of the venture, or a choice of what
     the venture should do. Use implied actions -- which are based on a prior extractor's interpretation -- only
      when supported by the stated positions; do not infer additional disagreement from them.

# The labels:

## value: the sides disagree about the venture's overall quality or attainable value.

Whether the venture can succeed; how much value or scale it can achieve; whether the business can be funded,
     can make money, or can grow; the venture is worth continuing or worth the room's support. A disagreement
     about whether its present capabilities support those prospects -- e.g. whether the technology works, is
     differentiated, or can be defended; the team can do what it claims; the market is real or big enough;
     customers want the product -- belongs here when it changes the assessment of the venture as a whole.
A disagreement confined to the value or feasibility of a particular path -- how attractive a market or segment
     is, how a pricing or channel choice would play out, which route yields more -- belongs to strategy.

## strategy: the sides disagree about the path the venture should take.
```

```
A choice among courses of action: which market, customer, or segment to go after; what the product should
     include; the business model, the pricing, and whether to build, partner, license, or outsource; whether,
     when, how much, and from whom to raise money; whom to hire; how the product is positioned, pitched, or
     described; what to do first and how to spend the coming weeks; how ambitious or how specific an objective
     should be; what test, evidence, or milestone to pursue next.

## other: the sides contradict on something that is neither.
A fact a document would settle, such as the terms of a deal; how the program or the session is run; what to
     call something; tone.

## unclear: the record does not say what the sides contradict on.
Use this only when the extracted text is too thin or contradicts itself. A disagreement you can read but that
     fits neither value nor strategy is other, not unclear.

# Before you answer, you MUST:
1. identify the claim the two sides contradict each other on, and state it in the rationale.
2. apply the rule for mixed cases below.
3. respond in exactly the JSON output format below, and nothing else.

# The rule for mixed cases:
When disagreement about a particular option also reflects different assessments of the venture's overall
     quality or attainable value, classify as value. When the disagreement concerns the option's value or
     feasibility without that difference in the overall assessment, classify as strategy.
"Not ready" could be value or strategy depending on context -- it can question a foundational capability and
     the venture's value today, or the feasibility and timing of a particular plan.

# The remaining fields:
1. second_kind: The most relevant different kind also contested: value, strategy, other, or none. Do not repeat
      type or count an action that merely follows from a value assessment as a separate strategy disagreement.

2. other_kind: when type or second_kind is other, say in one or two words what kind of claim the sides
     contradict on, such as "fact", "process", or "wording". Use "evidence" when the sides disagree about what
     existing observations or results establish. Otherwise write "none".
3. two_decisions: true only when the record holds two disagreements that different decisions would resolve --
     two disagreements that should have been separate records. Two claims that one decision resolves are one
     disagreement: false.

# Output format:
Return only valid JSON, no prose and no code fence, in exactly this structure. The // notes are constraints on
     each field; they are not part of the output.
{
  "type": string, // one of: value | strategy | other | unclear
  "second_kind": string, // one of: value | strategy | other | none
  "other_kind": string, // one or two words when type or second_kind is other; else "none"
  "two_decisions": boolean, // true only if different decisions would resolve two disagreements in this record
  "rationale": string // one sentence: the claim the two sides contradict on, and whether it is an assessment
     of the venture, a choice of action, or neither
}

The disagreement:
Topic: {topic}
Description: {topic_description}
Why contentious: {brief_summary_of_why_contentious}
Side A position: {side_a_position}
Side A implied action: {side_a_implied_action}
Side B position: {side_b_position}
Side B implied action: {side_b_implied_action}
```

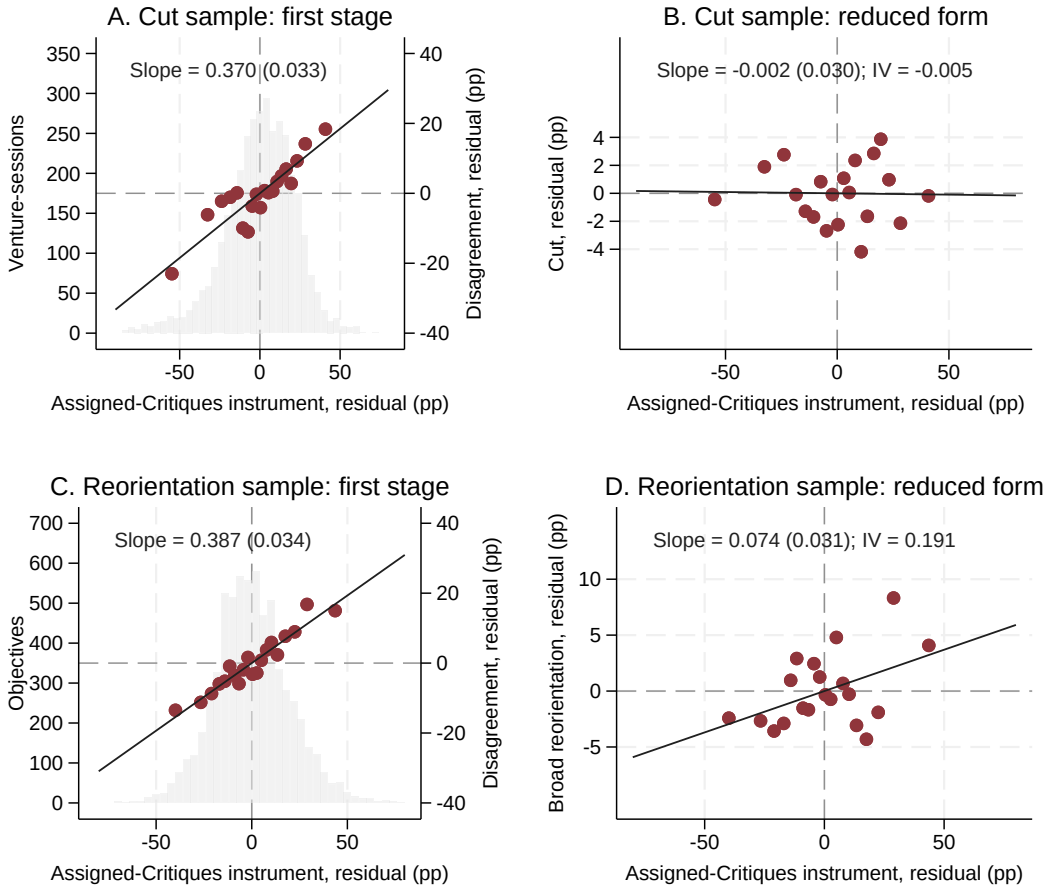
Notes: This figure shows the prompt that classifies each debate by topic. The model receives one debate's topic, description, reason for contention, and each side's position and implied action, as returned by the prompt in Figure A3; it does not see the transcript. It returns the label along with a one-sentence rationale. The placeholders, such as "{topic}", are filled in per debate, with one classification call for each debate. We run the prompt with Anthropic's Claude Opus 5 (claude-opus-5) at its default settings.

Appendix B Instrument Validity

This appendix examines the four conditions under which our IV estimates can be interpreted as treatment effects (Section 5.2), namely a strong first stage, quasi-random Critique assignment within a track-session, the exclusion restriction, and average monotonicity. We take each in turn and close with further checks of the estimates.

First stage. Figure B1, Panels A and C, plots the first stage for venture-session and objective-level outcomes. Tables 2 and 3 report Kleibergen–Paap first-stage F-statistics of 129.9 for reorientation and 126.3 for cut.

Figure B1: The First Stage and the Reduced Form



Notes: This figure shows the first stage and the reduced form of the Assigned Critiques instrument. Each panel plots twenty quantile bins of the instrument against the mean of the panel's outcome, both net of the track-session fixed effects, with the fitted line through the origin. Panels C and D add proposed-category fixed effects. In panels A and C, the gray bars are the instrument's histogram, in four-point intervals. Slopes are reported with standard errors clustered by venture.

Quasi-random assignment. This assumption could fail if organizers assign particularly disagreement-prone Critiques in response to venture-specific needs. The order in which CDL assigns roles limits how far organizers can influence the Critique assignment. We simulate how a planner who knows all ventures' best-fit mentors would assign Lead and Critique mentors. We run one simulation for each track-session of the assignment sample, with actual data on the participating ventures and mentors. The fit of venture v with mentor m is

$$q_{vm} = \tau \, \mathbf{1}\{a_v = a_m\} + s_m + \varepsilon_{vm},$$

where a_v and a_m are within-stream technology areas (in a quantum stream, for example, algorithms or hardware), s_m is the mentor's general strength, which every venture values, and ε_{vm} is the pair's "chemistry," the fit specific to the pair.

Each stream has four technology areas, which account for 40, 30, 20, and 10% of its ventures and mentors. We draw s_m and ε_{vm} from a standard normal distribution and set τ so that each of the three parts accounts for a third of the variance of fit. The planner assigns Leads first, in a serial dictatorship matching, where, in a random order, each venture takes its highest-fit mentor among those holding fewer than three roles. It then assigns Critiques, where each venture takes its highest-fit mentor other than its Lead among those holding fewer than three roles. We omit constraints from rooms and meeting times, which would lower the share of best-fit Critiques. For each venture, we record whether its Lead and its Critique are its best-fit mentors, and how much of the gap between the fit of a randomly chosen mentor and the best fit its Critique closes. We average these measures over 1,000 simulated days per track-session in Panel A and 500 in Panel B, and then across the 337 track-sessions with enough mentors to give every venture a Lead and a different Critique.

Table B1 reports the results. The planner gives a venture its best-fit Lead 86.5% of the time but its best-fit Critique only 22% of the time. The assigned Critique closes about half of the fit gap. Panel A shows how the rules produce the differences in match quality. With no constraints, every venture would receive its best-fit Critique. First, the three-role limit lowers the share of best-fit Critiques to 54%, then requiring every mentor to hold a role further lowers the share to 22%. Panel B shows that the share depends most on how much ventures agree on which mentors are strong. For illustration, the share of best-fit Critiques is 47% when general strength does not matter and 7% when general strength's weight is doubled.

To examine this assumption, Figure B2 plots the relationship between the assigned Critique's disagreement propensity and venture characteristics, including pre-program capital, team composition, the mix of proposed objectives, venture stage, and the support the venture received in the previous session. Individually, we find no discernible relationships, with the exception of a small negative correlation with venture stage. To examine these characteristics jointly, we regress the instrument on them, controlling for track-session fixed effects, and test whether their coefficients are jointly zero. For both specifications, the test does not reject joint balance. We also combine venture characteristics into measures of predicted cut and reorientation risk; neither measure is discernibly related to the instrument. Nor do we find a relationship between the assigned Critique's

Table B1: Simulated Lead and Critique Assignment

	Lead is best fit (%)	Critique is best fit (%)	Fit gap closed by Critique (%)
A. Rules added in turn			
At most three roles per mentor	86.5	53.7	83.9
+ Every mentor holds a role (baseline)	86.5	22.3	53.7
B. Baseline with one change			
No general strength	98.5	46.6	73.0
General strength half as strong	95.6	37.1	66.6
General strength twice as strong	64.8	7.0	31.1
At most two roles per mentor	75.1	15.4	46.6
At most four roles per mentor	93.0	27.2	57.4

Notes: This table shows how often a fully informed planner gives a venture its best-fit Lead and Critique under CDL's assignment order. The row with at most two roles uses the track-sessions where two roles suffice.

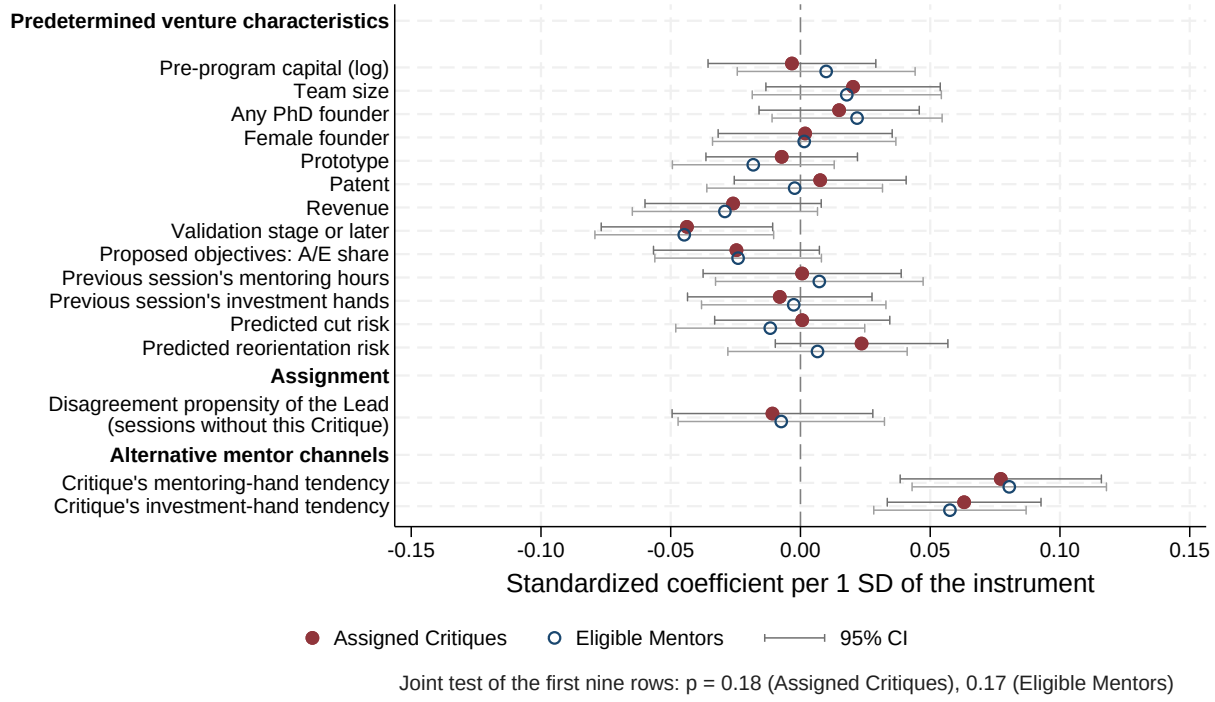
disagreement propensity and the Lead's, which is inconsistent with CDL systematically matching mentors based on their propensity to disagree.

Exclusion restriction. A particular concern is that disagreement-prone Critiques may be less willing to support ventures, so that assigning a more disagreement-prone Critique affects cut decisions directly. Figure B2 shows the opposite relationship, with more disagreement-prone Critiques more likely to offer mentoring support and express investment interest in other ventures' sessions. In Table B2, we detect no relationship between the instrument and whether the assigned Critique commits time to the venture they are assigned to. Moreover, columns 3 and 5 of Table B2 show that the estimated effects of disagreement on mentoring hours and on cut are robust, respectively, to (i) excluding the Critique's own commitment from total mentoring hours and (ii) directly controlling for the Critique's mentoring and investment tendencies.

Average monotonicity. Figure B3 shows positive first-stage coefficients across cohorts, sessions, sites, and proposed categories, consistent with average monotonicity (Chyn *et al.*, 2025).

Other checks. Figure B4 shows that the results are qualitatively similar when we vary the minimum mentor-history requirement from two to six classified sessions involving other ventures. Table B3 shows that they are also similar with two-sample IV, which adds observations without measured disagreement to the reduced form, and with instruments that leave out the focal venture-session rather than the entire venture.

Figure B2: Balance and Channel Checks for the Critique Assignment



Notes: This figure shows how the Critique's disagreement propensity relates to predetermined venture characteristics, the Lead's propensity and the Critique's support of other ventures. The unit of analysis is a venture-session in the assignment sample. Each point comes from a regression of the standardized row variable on one specification's standardized instrument, with track-session fixed effects. The A/E share is the share of proposed objectives in analysis or experimentation. The predicted risks are the linear index of the first nine rows from regressions of cut and of the session's share of broadly reoriented objectives on them with cohort and session fixed effects. The Lead's disagreement propensity is their disagreement share as Lead in their other ventures' sessions without this Critique, where they have at least four such sessions. The hand tendencies are the shares of other ventures' sessions in the cells the Critique attended, as neither Lead nor Critique, in which they raised a mentoring hand and an investment hand. Standard errors are clustered by venture.

Table B2: The Assigned Critique's Own Mentoring Hand

DV=	B2-1 Critique's Mentoring Hand	B2-2 Mentoring Hours	B2-3 Mentoring Hours, Critique's Hand Excluded	B2-4 Cut	B2-5 Cut
Disagreement	−0.063	0.784	1.036	−0.005	−0.004
	(0.079)	(0.936)	(0.881)	(0.082)	(0.084)
Instrument, per SD	−0.005	0.064	0.084	0.000	0.000
	(0.006)	(0.076)	(0.071)	(0.007)	(0.007)
N	3,573	3,573	3,573	3,573	3,573
Mean of DV	0.190	6.400	5.640	0.208	0.208
Kleibergen–Paap rk Wald *F*	126.3	126.3	126.3	126.3	120.1
Track-Session FE	X	X	X	X	X
Critique's hand tendencies					X

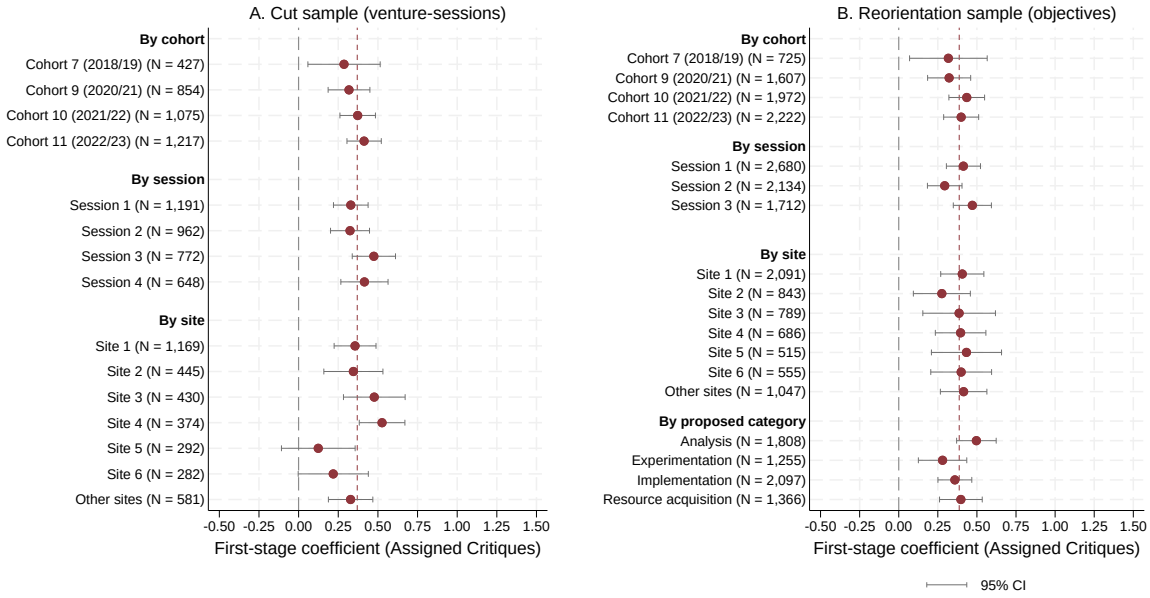
Notes: This table shows whether the Critique's own mentoring hand explains the effects of disagreement on mentoring hours and cut. The unit of analysis is a venture-session in the assignment sample of Table 3. The Disagreement row reports the IV coefficient on Disagreement, instrumented by the Assigned Critiques instrument as in panel B of Table 3. The instrument row reports the reduced form, the change in the outcome for a one-standard-deviation increase in the instrument. Column 1's outcome is whether the assigned Critique raised a mentoring hand for this venture-session. Column 3 removes the Critique's hand from the session's mentoring hours. Column 5 repeats column 4 with the Critique's two hand tendencies of Figure B2 as controls in both stages; their coefficients are not shown. Standard errors clustered by venture are in parentheses.

Table B3: Two-Sample IV and Venture-Session Leave-Out Estimates for Cut and Broad Reorientation

	Two-Sample IV	Venture-Session Leave-Out		Two-Sample IV	Venture-Session Leave-Out	
DV=	B3-1 Cut	B3-2 Cut	B3-3 Cut	B3-4 Broad Reorientation	B3-5 Broad Reorientation	B3-6 Broad Reorientation
Disagreement	0.007	−0.014	−0.007	0.182**	0.184**	0.235***
	(0.073)	(0.083)	(0.093)	(0.073)	(0.080)	(0.084)
Comparison set	Assigned Critiques	Assigned Critiques	Eligible Mentors	Assigned Critiques	Assigned Critiques	Eligible Mentors
N, first stage	3,573	3,573	3,573	6,526	6,526	6,526
N, reduced form	3,906			7,288		
Kleibergen–Paap rk Wald *F*	126.3	124.5	96.7	129.9	123.2	129.9
2-sample cond. LR *p*	0.927			0.013		
Anderson–Rubin *p*		0.864	0.942		0.025	0.006
Track-Session FE	X	X	X	X	X	X
Proposed Category FE				X	X	X

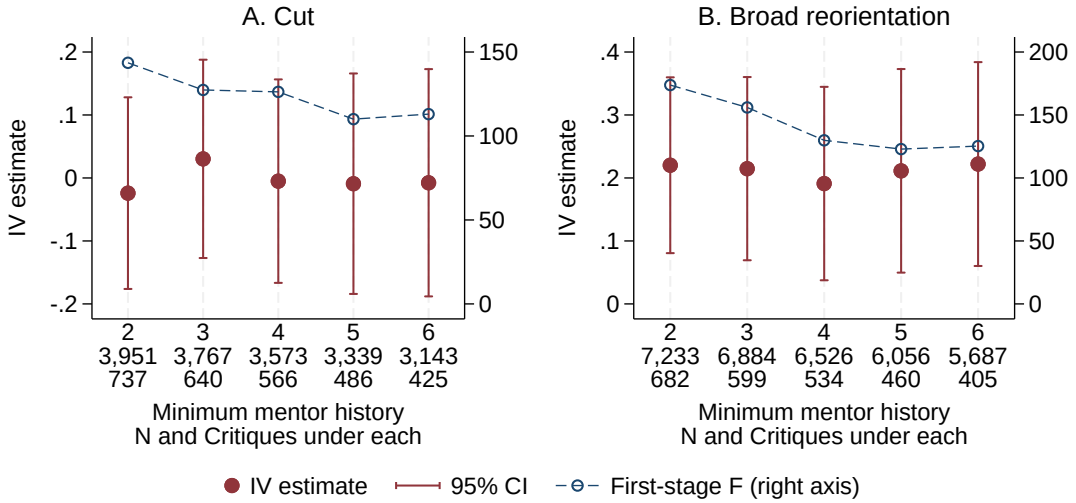
Notes: This table shows the IV estimates for cut and broad reorientation under two alternative designs. The unit of analysis is a venture-session in columns 1–3 and a finalized objective in columns 4–6, in the assignment sample. Column 1 uses the first stage of Table 3, panel B, and column 4 uses that of Table 2, column 2, both on the assignment sample. Their reduced form adds the venture-sessions outside the estimation sample whose assigned Critique is in the assignment design. The 2-sample cond. LR *p* is the *p*-value of the two-sample conditional likelihood-ratio test that the coefficient is zero. Unlike every other statistic in the table, the test is heteroskedasticity-robust and not clustered by venture. Columns 2 and 3 repeat column 1 of Table 3, panel B, and column 1 of Table C6, panel A, and columns 5 and 6 repeat column 3 of Table 2 and column 3 of Table C3, panel A, each with the venture-session, rather than the venture, left out of the Critique's propensity. Standard errors clustered by venture are in parentheses. Statistical significance is *(10%), **(5%), or ***(1%).

Figure B3: The First Stage by Subgroup



Notes: This figure shows the first stage by subgroup, a check on average monotonicity. Each point is the first-stage coefficient on the Assigned Critiques instrument within one subgroup, with track-session fixed effects. Standard errors are clustered by venture. The short-dashed line is the full-sample coefficient. Panel B adds proposed-category fixed effects. Sites are numbered in descending order of their venture-sessions in the assignment sample. Sites with fewer than 200 venture-sessions are pooled.

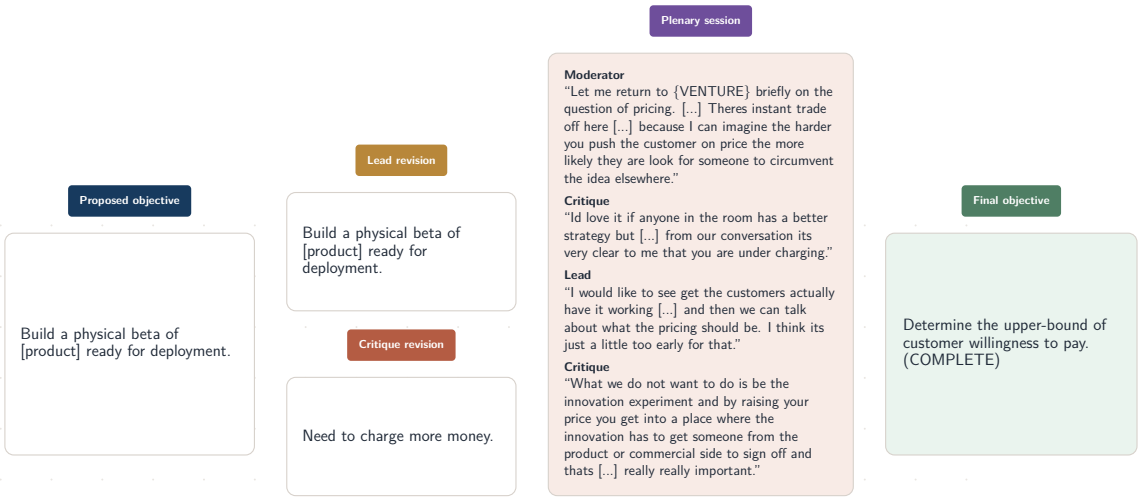
Figure B4: The Minimum-History Rule



Notes: This figure shows the IV estimates by minimum mentor history. Our main specification requires four. Each point is the IV estimate of Disagreement's effect, instrumented by the Assigned Critiques instrument, with track-session fixed effects. At each minimum, the pool of eligible mentors, the sample and the instrument are rebuilt. Panel B adds proposed-category fixed effects. Critiques is the number of assigned Critiques in each estimation. Standard errors are clustered by venture.

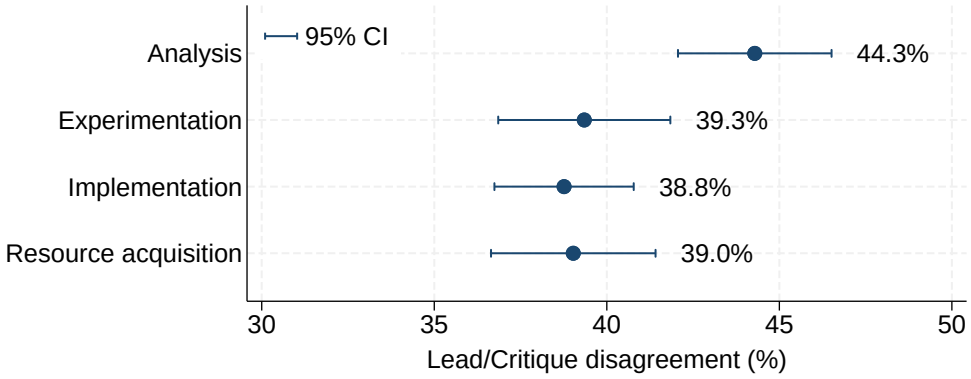
Appendix C Additional Figures and Tables

Figure C1: An Observed Case of Disagreement over Strategic Direction



Notes: This figure shows one objective's path from the founders' proposed objective to the finalized objective. The Lead retained the venture's proposed product-development objective, while the Critique prioritized pricing; the pair is classified as *Replace*. Venture and product identifiers are masked. Excerpts are shortened and otherwise verbatim.

Figure C2: Lead–Critique Disagreement by the Proposed Objective's Category



Notes: This figure shows the Lead–Critique disagreement rate by the category of the founders' proposed objective. The unit of analysis is a finalized objective of the classified sample. Standard errors are clustered by venture.

Table C1: Examples of Lead–Critique Classifications

Classification	Distinguishing feature	Example of Lead–Critique Exchange
Agree	No substantive change	*Lead:* "Close seed round." *Critique:* "AGREE."
Agree with minor changes	Same task with a limited qualification	*Lead:* "Close seed round." *Critique:* "AGREE—but consider alternative fundraising amounts or structures."
Revise	Same business function with a material change in scope, timing, or method	*Lead:* "Close seed round." *Critique:* "MODIFY: Sign term sheet in December and seek accelerated due diligence to close the seed round."
Replace	Different business function	*Lead:* "Close current round." *Critique:* "Close the POC and through the process actually work with a customer to learn what resonates, where the true pain points are, and iterate quickly based on that discovery."

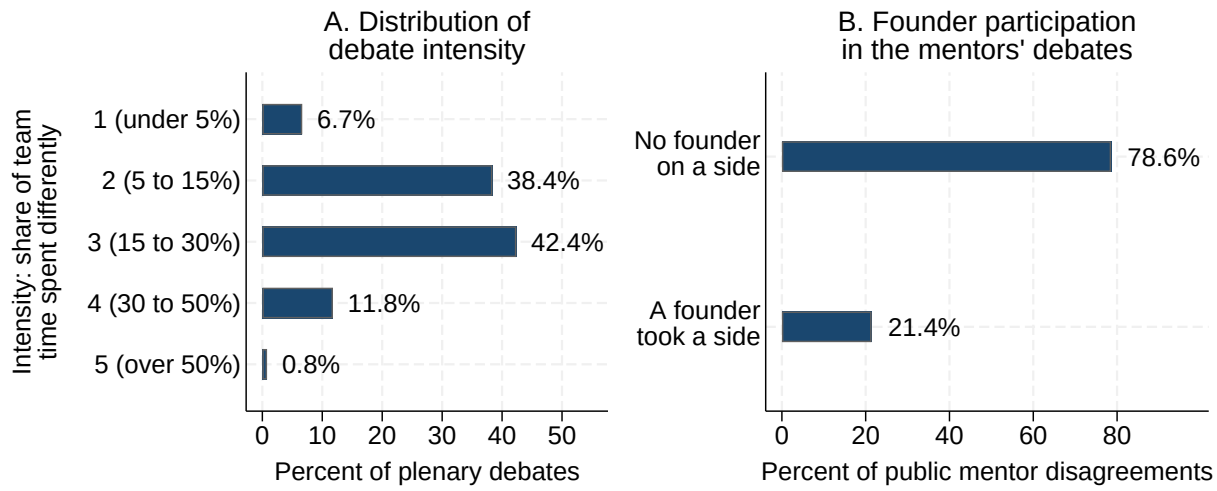
Notes: This table shows one Lead–Critique exchange for each of the four classifications.

Table C2: Examples of Public Mentor Disagreement

Topic	Coding rule	Example of Public Debate	Intensity
Value	The venture's overall viability	*Debate:* credibility of the claimed computational breakthrough. *Mentor A:* "This just strikes me as something that's either amazing or fake." *Mentor B:* "I would probably give the founder a little bit more credit . . . I would probably err on the side of amazing."	3
Strategy	Which path the venture should take	*Debate:* stay in automotive or pivot to another market. *Mentor A:* "Cars are often proprietary or very cheap. Find something else. It sounds to me like the team is fantastic but you need a different application." *Mentor B:* "I don't know. Auto could be super big. . . . Your job is to figure that out now."	5

Notes: This table shows one example of a value debate and one of a strategy debate. Each example is one debate the extraction found in a venture-session's plenary record, with one comment from a mentor on each side; the comments are shortened and otherwise verbatim. Intensity is the share of the founding team's working time over the next two months that would be spent differently under the two sides' courses of action, on the five-point scale of Appendix Figure A3: 3 is 15 to 30 percent, 5 more than 50 percent.

Figure C3: Public Debate Intensity and Founder Participation



Notes: This figure shows how intense the debates are and how often founders take part in them. Panel A covers all debates extracted by the LLM from the plenary sample's records, including those without a mentor on each side. Panel B covers the debates with a mentor on each side and reports whether a founder of the venture took a side.

Table C3: Broad Reorientation and Completion Following Lead–Critique Disagreement Alternative Instrument, Venture Fixed Effects and the Finalized Objective

	OLS	IV: Assigned Critiques	IV: Eligible Mentors		IV: Assigned Critiques	IV: Eligible Mentors
	C3-1	C3-2	C3-3	C3-4	C3-5	C3-6
Panel A: Broad Reorientation						
Disagreement	0.128***	0.255*	0.219***	0.267*	—	—
	(0.015)	(0.142)	(0.079)	(0.144)		
N	6,515	6,515	6,526	6,515		
Mean of DV	0.453	0.453	0.452	0.453		
Kleibergen–Paap rk Wald *F*		38.2	147.4	37.3		
Anderson–Rubin *p*		0.080	0.006	0.069		
Panel B: Objective Completed						
Disagreement	0.017	0.342**	0.145*	0.297**	0.129*	0.143*
	(0.014)	(0.149)	(0.077)	(0.149)	(0.077)	(0.076)
N	6,515	6,515	6,526	6,515	6,526	6,526
Mean of DV	0.619	0.619	0.619	0.619	0.619	0.619
Kleibergen–Paap rk Wald *F*		38.3	148.2	37.2	131.5	145.9
Anderson–Rubin *p*		0.012	0.057	0.031	0.093	0.058
Track-Session FE	X	X	X	X	X	X
Proposed Category FE	X	X	X	X	X	X
Venture FE	X	X		X		
Proposed Text's Scores (Panel B)	X	X	X	X	X	X
Finalized Text's Scores					X	X
Finalized Category FE					X	X

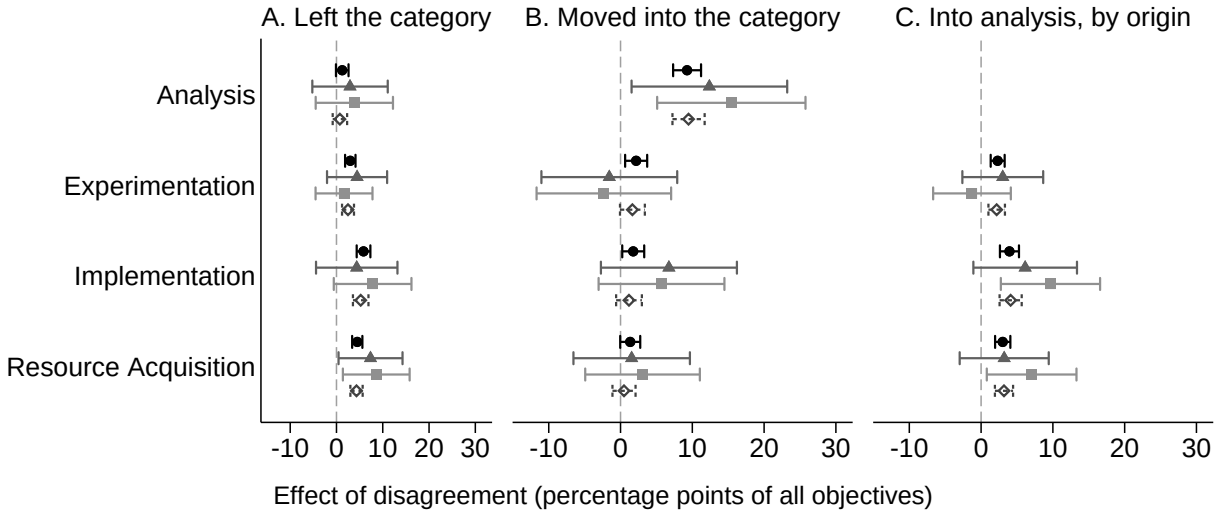
Notes: This table shows the estimates of Table 2 under alternative specifications. The unit of analysis and the sample are those of Table 2. Columns 5 and 6 add the finalized objective's two text scores and category fixed effects, which apply to completion only. Standard errors clustered by venture are in parentheses. Statistical significance is *(10%), **(5%), or ***(1%).

Table C4: Near-Verbatim Adoption Following Lead–Critique Disagreement

DV=	C4-1 Near-verbatim Adoption	C4-2 Near-verbatim Adoption	C4-3 Near-verbatim Adoption
Disagreement	−0.084***	−0.083***	−0.084***
	(0.007)	(0.007)	(0.008)
N	6,526	6,526	6,515
Mean of DV	0.093	0.093	0.092
Track-Session FE	X	X	X
Proposed Category FE		X	X
Venture FE			X

Notes: This table shows how disagreement relates to near-verbatim adoption of the founders' proposed objective. The unit of analysis and the sample are those of Table 2. Near-verbatim adoption equals one when the finalized objective repeats the text of the founders' proposed objective. Every column is ordinary least squares. Standard errors clustered by venture are in parentheses. Statistical significance is *(10%), **(5%), or ***(1%).

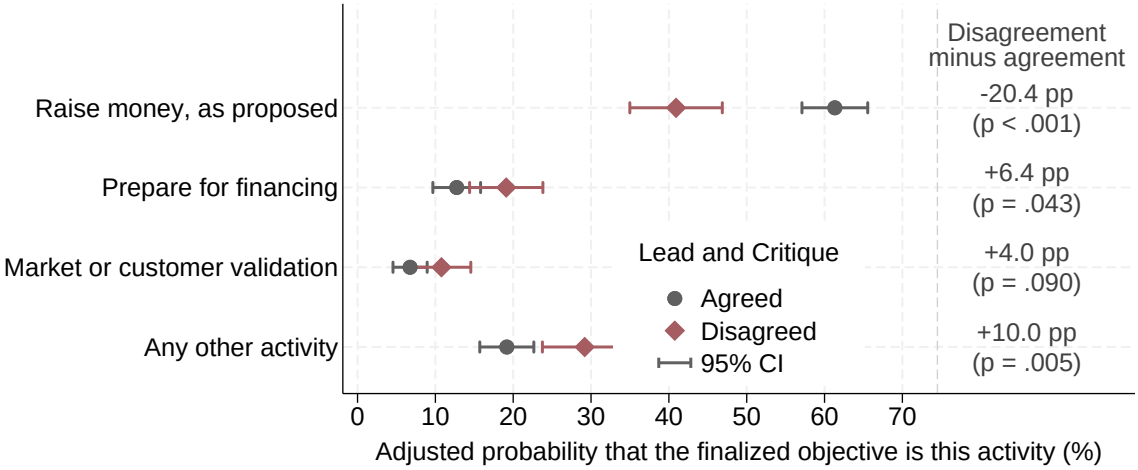
Figure C4: The Origins and Destinations of Reorientation
Both Instruments and Venture Fixed Effects



• OLS ▲ IV: Assigned Critiques ■ IV: Eligible Mentors ◊ OLS with venture fixed effects —— 95% CI

Notes: This figure shows panel B of Figure 3 under alternative specifications. The unit of analysis, the sample, the outcomes and the stacked regressions are those of Figure 3. The Assigned Critiques estimates instrument Disagreement on that comparison set, as in column 3 of Table 2. The Eligible Mentors estimates instrument it on that comparison set, as in column 3 of Table C3, panel A. The venture fixed effects are specific to each category, as in column 1 of Table C3, panel A. Standard errors are clustered by venture.

Figure C5: Where a Proposed Raise Money Objective Ends Up, by Lead–Critique Disagreement



Notes: This figure shows where a proposed Raise Money objective ends up, under agreement and under disagreement. The unit of analysis is a finalized objective of the classified sample whose proposed objective was Raise Money, in a track-session with at least one other such objective, at sessions 1–3. Market or customer validation covers market research, letter of intent, value proposition, product-market fit and market choice. Markers are adjusted probabilities from a linear probability model of the destination on disagreement with track-session fixed effects. Standard errors are clustered by venture.

Table C5: Properties of Finalized Objectives Following Disagreement

DV=	C5-1 Subjectivity Score	C5-2 Founder-Control Score	C5-3 Ambiguous Objective
Panel A: Without the Proposed Text's Scores			
Disagreement	0.179***	0.039***	0.023***
	(0.022)	(0.008)	(0.007)
Panel B: With the Proposed Text's Scores			
Disagreement	0.153***	0.027***	0.022***
	(0.021)	(0.008)	(0.007)
Proposed Subjectivity Score	0.279***	−0.004	−0.004
	(0.016)	(0.005)	(0.005)
Proposed Founder-Control Score	0.206***	0.428***	0.044***
	(0.031)	(0.013)	(0.010)
N	8,283	8,283	8,283
Mean of DV	2.460	0.673	0.084
Track-Session FE	X	X	X
Proposed Category FE	X	X	X

Notes: This table shows how disagreement relates to the finalized objective's subjectivity, founder control and ambiguity. The unit of analysis is a finalized objective of the classified sample. Ambiguous Objective is 1 where the activity labeling gives the finalized objective more than one of the 34 activity categories or cannot place it in any. Standard errors clustered by venture are in parentheses. Statistical significance is *(10%), **(5%), or ***(1%).

Table C6: Mentor Support Following Lead–Critique Disagreement
Alternative Instrument and Venture Fixed Effects

DV=	C6-1 Cut	C6-2 Mentoring Hours	C6-3 Investment Interest	C6-4 Graduated
Panel A: IV, Eligible Mentors				
Disagreement	−0.008	0.386	0.240	0.048
	(0.089)	(1.005)	(0.211)	(0.106)
N	3,573	3,573	3,573	3,573
Mean of DV	0.208	6.400	0.376	0.414
Kleibergen–Paap rk Wald *F*	110.9	110.9	110.9	90.1
Anderson–Rubin *p*	0.930	0.701	0.253	0.651
Panel B: OLS, with Venture Fixed Effects				
Disagreement	—	−0.149	−0.129*	—
		(0.203)	(0.068)	
Panel C: IV, Assigned Critiques, with Venture Fixed Effects				
Disagreement	—	0.730	0.327	—
		(1.206)	(0.280)	
Kleibergen–Paap rk Wald *F*		52.1	52.1	
Anderson–Rubin *p*		0.542	0.236	
Panel D: IV, Eligible Mentors, with Venture Fixed Effects				
Disagreement	—	0.714	0.135	—
		(1.323)	(0.308)	
Kleibergen–Paap rk Wald *F*		44.6	44.6	
Anderson–Rubin *p*		0.586	0.661	
N, panels B–D		3,042	3,042	
Mean of DV, panels B–D		7.039	0.399	
Track-Session FE	X	X	X	X

Notes: This table shows the estimates of Table 3 under alternative specifications. The unit of analysis and the sample are those of Table 3. Panels B to D use the ventures with more than one session in the sample. These panels omit Graduated, which is constant within a venture, and Cut, since nearly half of all cuts come after a venture's only session. Standard errors clustered by venture are in parentheses. Statistical significance is *(10%), **(5%), or ***(1%).

Table C7: Mentor Support by Venture Stage and Later Outcomes Following Lead–Critique Disagreement

	Support				Later Outcomes		
DV=	C7-1 Cut	C7-2 Mentoring Hours	C7-3 Investment Interest	C7-4 Graduated	C7-5 Financing Round Within 24 Months	C7-6 Any Financing Raised	C7-7 In Business June 2026
Panel A: OLS							
Disagreement	—	—	—	—	−0.048**	−0.027	−0.012
					(0.019)	(0.017)	(0.018)
N					3,247	3,247	3,247
Panel B: IV, Assigned Critiques							
Disagreement	—	—	—	—	0.035	0.006	0.053
					(0.102)	(0.093)	(0.094)
N					3,247	3,247	3,247
Kleibergen–Paap rk Wald *F*					103.7	89.7	89.7
Panel C: OLS, by venture stage							
Disagreement	0.001	0.178	0.014	−0.032	−0.039	−0.009	−0.012
	(0.025)	(0.277)	(0.086)	(0.029)	(0.031)	(0.030)	(0.029)
Validation stage or later	−0.046*	0.997***	0.141	0.027	−0.000	0.064**	0.044
	(0.027)	(0.310)	(0.086)	(0.035)	(0.037)	(0.032)	(0.034)
Disagreement ×	0.048	−0.859**	−0.224**	0.002	−0.015	−0.027	0.002
Validation stage or later	(0.031)	(0.350)	(0.103)	(0.037)	(0.038)	(0.037)	(0.037)
N	3,573	3,573	3,573	3,573	3,247	3,247	3,247
Panel D: IV, by venture stage							
Disagreement	−0.199	1.002	0.373	0.310**	0.335**	0.168	−0.006
	(0.125)	(1.429)	(0.301)	(0.140)	(0.165)	(0.146)	(0.133)
Validation stage or later	−0.241*	0.559	0.071	0.342**	0.338**	0.217	−0.035
	(0.127)	(1.459)	(0.299)	(0.147)	(0.159)	(0.141)	(0.139)
Disagreement ×	0.323*	−0.175	−0.103	−0.437**	−0.494**	−0.243	0.116
Validation stage or later	(0.179)	(2.065)	(0.425)	(0.205)	(0.224)	(0.197)	(0.194)
N	3,573	3,573	3,573	3,573	3,247	3,247	3,247
Kleibergen–Paap rk Wald *F*	31.4	31.4	31.4	29.7	29.6	29.6	29.6
Interaction *p*, Eligible Mentors	0.057	0.724	0.398	0.160	0.046	0.402	0.743
N, before validation stage	1,428	1,428	1,428	1,428	1,254	1,254	1,254
N, validation stage or later	2,145	2,145	2,145	2,145	1,993	1,993	1,993
Mean of DV	0.208	6.400	0.376	0.414	0.321	0.819	0.815
Track-Session FE	X	X	X	X	X	X	X

Notes: This table shows how disagreement relates to mentor support by stage and to its later outcomes. The unit of analysis is a venture-session in the assignment sample. Columns 5 to 7 use the venture-sessions whose three later outcomes are all observed. A financing round within 24 months is a round CDL records within 24 months after the session. In business is based on PitchBook data and whether the company's website is accessible. Graduated, any financing raised and in business weight each venture-session by the inverse of the venture's number of sessions. Panels A and B show the later outcomes with the same specifications as in Table 3. All panels include track-session fixed effects. Standard errors clustered by venture are in parentheses. Statistical significance is *(10%), **(5%), or ***(1%).