Causal Analyses on Entrepreneurial Intention, Success, and Failure

Vom Promotionsausschuss der

Technischen Universität Hamburg

zur Erlangung des akademischen Grades

Doktor der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.)

genehmigte Dissertation (Monografie)

von

Oliver Wolfgang Mork

aus

Jülich

2025

1. Gutachter: Prof. Dr. Christoph Ihl
2. Gutachter: Prof. Linus Dahlander, PhD
Tag der mündlichen Prüfung: 13.10.2025

ORCID: 0009-0001-0560-0617

doi:10.15480/882.16044

Executive Summary

In this dissertation, I embark on two interconnected journeys. One, the topical journey, follows the entrepreneurial process, tracing its evolution from the initial spark of entrepreneurial intention, through the operational challenges of scaling a venture, to the aftermath of failure and the potential for re-entry. The other, the methodological journey, navigates the empirical complexities of identifying causal relationships in entrepreneurship, leveraging recent advancements and discussions in causal inference. Together, these journeys illuminate not only the entrepreneurial life cycle but also the methodological rigor required to uncover actionable insights in the dynamic and impactful domain of entrepreneurship.

Pre-Venture Stage. Study 1 explores the formation of entrepreneurial intention and self-efficacy, foundational attitudes to initiating action. A randomized controlled experiment within a university entrepreneurship course reveals that autonomous team formation enhances entrepreneurial intention, while self-efficacy is influenced by peer interaction and only when some form of autonomy is present. These findings inform educators and policymakers on fostering entrepreneurial attitudes through structured interventions.

Venture and Growth Stage. Study 2 focuses on the operational phase, examining how external enablers like venture capitalists (VCs) influence startup outcomes. Leveraging exogenous travel shocks to design natural experiments, it identifies causal effects of on-site VC monitoring, finding that reduced travel barriers boost innovation, survival, and exit rates, while disruptions have negative impacts. The study underscores the importance of physical proximity in collaboration and informs innnovation as well as transportation policy.

Post-Venture Stage. Study 3 investigates entrepreneurial failure, the most common outcome of the entrepreneurial process, using observational data in a quasi-experimental approach and natural language processing of a large number of news articles. Results show that negative media narratives amplify stigma and deter re-entry, while balanced coverage mitigates these effects. The findings emphasize the societal role in evaluating failure and sustaining entrepreneurial activity.

Methodological Framework. Causal inference underpins this research, enabling the isolation of cause-and-effect relationships in dynamic entrepreneurial contexts. Employing randomized trials (Study 1), natural experiments (Study 2), and quasi-experimental designs (Study 3), the studies use advanced econometric techniques to ensure robust identification of causal effects. This approach strengthens the credibility of findings and offers a methodological template for future entrepreneurship research.

Contents

List of Figures

List of Tables

List of Abbrevations

ATF	Autonomous Team Formation
DOCB	EPO's master documentation database
DiD	Difference-in-Differences
EMBA	Executive Master of Business Administration
EEM	Entrepreneurial Event Model (Shapero and Sokol, 1982)
FAISS	Facebook AI Similarity Search
GMM	Generalized method of moments
HHI	Hersch-Herfindahl index
IPO	Initial public offering
IPW	Inverse probability weighting
MBA	Master of Business Administration
NAICS	North American Industry Classification System
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary least squares
RoBERTa	Robustly Optimized BERT Pre-training Approach
SD	Standard deviation
SIC	Standard Industrial Classification
SUTVA	Stable unit treatment value assumption
TBP	Theory of Planned Behavior (Ajzen, 1991)
TWFE	Two-way fixed effect regression
VC	Venture capital
WMT	Workshop/Conference on Machine Translation
%pt	Percentage points

List of Symbols and Notation

i	Unit
t	Time
D	Treatment group
G	Treatment group (timing of treatment)
W	Set of time-invariant and time-varying covariates
Z	Time-invariant covariates
X_t	Time-varying covariates
Y	Outcome
$\{Y(1), Y(0)\}$	Potential outcomes
$Y(0)$ and $Y(\infty)$	Untreated outcomes
$\widehat{Y}$	Estimated outcome
$\tilde{Y}$	Residualized outcome
τ	Treatment effect
ATE	Average Treatment Effect
ATT	Average Treatment Effect on the Treated

Chapter 1

Introduction

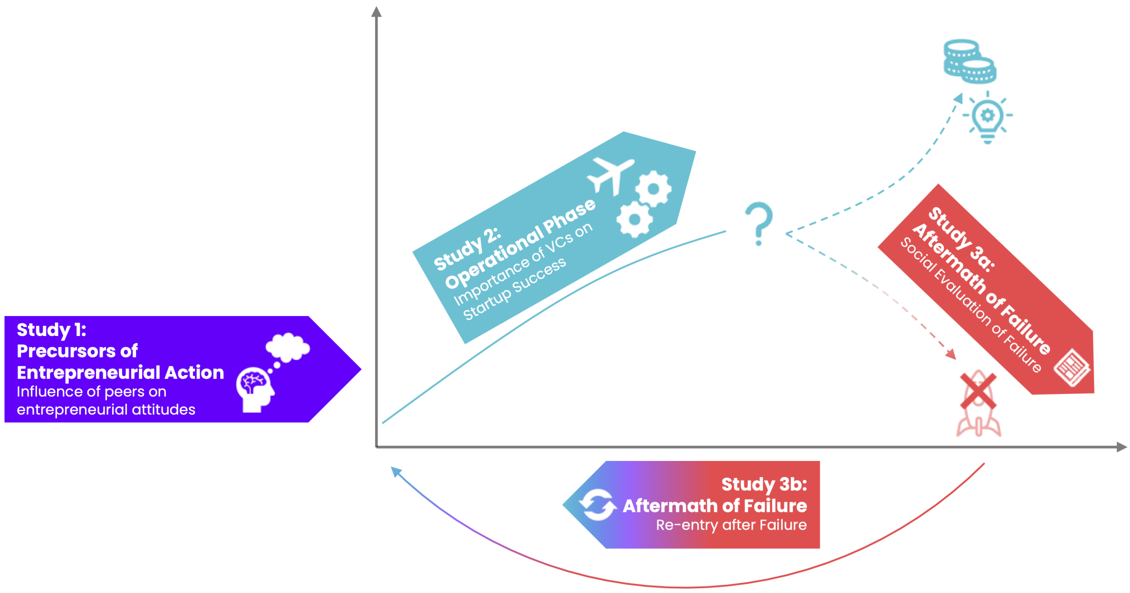
1.1 Entepreneurial Life Cycle

Entrepreneurship is a dynamic and multifaceted process that unfolds across distinct phases, from the initial spark of intention to the challenges of venture growth and, for many, the reality of failure and potential re-entry. Understanding how entrepreneurs navigate these phases - what enables success, what constrains growth, and how failure is evaluated - is essential for advancing both entrepreneurship theory and practice. Each stage of the entrepreneurial life cycle is shaped by individual agency, external enablers, and broader societal factors, making a comprehensive, contextualized approach necessary to capture the complexity of entrepreneurial activity.

This research examines three critical stages of the entrepreneurial life cycle as illustrated in Figure 1.1: pre-venture (1), venture and growth (2), and post-venture (3a and 3b). The pre-venture stage focuses on the psychological and cognitive foundations of entrepreneurship, particularly entrepreneurial intention and self-efficacy, which influence whether individuals choose to pursue entrepreneurial endeavors (Lee and Wong, 2004; Parker, 2006). The venture and growth stage investigates the conditions that determine startup success, focusing on investor relationships and external shocks such as the COVID-19 pandemic. Finally, the post-venture stage explores entrepreneurial failure, its social consequences, and the likelihood of re-entry, highlighting the long-term implications of failure on both individuals and the broader entrepreneurial ecosystem.

The motivation for studying these phases arises from the pivotal role that startups play in driving employment, economic growth, and innovation. However, for entrepreneur-

FIGURE 1.1: Entrepreneurial Life Cycle



ship to fulfill this potential, several key conditions must be met: individuals must be motivated and equipped to pursue entrepreneurial activity (study 1); newly established ventures require the necessary conditions to achieve growth and success (study 2); and when failure inevitably occurs, the social costs must be manageable to ensure that re-entry into entrepreneurship remains viable (study 3). If these conditions are not met, promising opportunities may be lost - not only at the individual level but also in terms of economic progress and societal attitudes toward entrepreneurship.

While these phases follow an intuitive sequence, the studies are connected not only sequentially but also cyclically. The final study on entrepreneurial failure and re-entry directly ties back to the first study on entrepreneurial intention and self-efficacy. Just as individuals require confidence and belief in their abilities to embark on entrepreneurship, those who experience failure must be able to rebuild these psychological foundations to consider re-entry. This study suggests that societal and institutional interventions, such as reducing failure stigma and fostering balanced media narratives, can support entrepreneurial resilience and encourage re-entry, much like structured interventions in the first study were shown to cultivate entrepreneurial intention.

Collectively, these studies provide a holistic, albeit not complete, perspective on entrepreneurship by addressing its precursors, operational challenges, and long-term implications. They shed light on the deeply personal decision to become an entrepreneur, the structural factors that influence venture outcomes, and the broader societal mechanisms that shape perceptions of failure and re-entry. By situating entrepreneurial activity within this broader life cycle framework, this research contributes to a more comprehensive understanding of the factors that enable entrepreneurship to thrive - even in the face of setbacks and failures.

1.1.1 Precursors of Entrepreneurial Action

Understanding the factors that drive the decision to become an entrepreneur serves as a natural starting point for studying the entrepreneurial process. Entrepreneurial action represents a significant decision with profound implications for an individual's financial and social circumstances. In study 1, we focus on pivotal factors in shaping the decision-making process that can ultimately lead to entrepreneurial action: entrepreneurial attitudes. As noted by Lee and Wong (2004), entrepreneurial action is unlikely in the absence of a prior intention to act entrepreneurially. Accordingly, entrepreneurial intentions and attitudes serve as a cornerstone for understanding the broader entrepreneurial process, providing critical insights into its inception.

Rooted in the theory of planned behavior (TPB) (Ajzen, 1991), attitudes have been both conceptually and empirically demonstrated to be among the most reliable predictors of behavior (Krueger and Carsrud, 1993; Bandura, 2012). In particular, we direct our attention towards the development of entrepreneurial intention and self-efficacy, both foundational attitudes that shape an individual's decision to pursue entrepreneurship. Entrepreneurial intention is the conscious commitment to starting a business, while self-efficacy presents the belief in one's ability to perform entrepreneurial tasks. Both attitudes are central to models of entrepreneurial behavior, including the TBP (Ajzen, 1991) and the Entrepreneurial Event Model (EEM) (Shapero and Sokol, 1982).

In a controlled field experiment, we examine influences of autonomous team formation (ATF), team diversity, and peer effects on entrepreneurial intention and self-efficacy. The context of this research, a university-based entrepreneurship course, provides a natural laboratory for investigating the early development of entrepreneurial attitudes. Here, individuals are still in the decision-making phase, exploring whether entrepreneurship aligns with their interests, skills, and aspirations compared to conventional employment.

By manipulating key variables such as the ability to autonomously self-select team members and autonomously choose the business idea, we disentangle the direct and indirect pathways through which autonomy shapes entrepreneurial attitudes. We further examine how peer interactions influence individual intentions and perceptions of self-efficacy.

We find that ATF positively affects entrepreneurial intention but not self-efficacy. Building on the positive effect of ATF on entrepreneurial intention, we find that this effect is partly suppressed by the lack of diversity in self-selected teams. However, peer effects only become evident on self-efficacy, and only ever, when students are allowed to autonomously select their entrepreneurial idea or enjoy some other dimension of autonomy.

By identifying the conditions that foster or inhibit these attitudes, the findings offer actionable insights for educators, policymakers, and organizations aiming to nurture the next generation of entrepreneurs. Through its experimental design, we not only advance theoretical understanding but also inform how practical interventions and educational programs can enhance entrepreneurial intent and self-efficacy, both critical precursors to venture creation.

1.1.2 Operational Phase of Startups

Building on the exploration of the transformative entrepreneurial process, we shift towards the operational phase where intentions have transitioned into actions (McMullen and Dimov, 2013). Entrepreneurs navigating this stage face a continuous stream of decisions under uncertainty, ranging from tactical choices to strategic pivots, always subject to reassessment and revision (Packard et al., 2017).

In this context, external and professional support, particularly from venture capitalists (VCs), is often regarded as pivotal. VCs are claimed to provide substantial benefits by aiding entrepreneurs in decisions related to human and financial capital, strategic guidance, oversight, and problem-solving (see e.g. Bernstein et al., 2016; Bronzini et al., 2020; Gompers et al., 2020; Hellmann and Puri, 2002). To evaluate whether such monitoring is critical for enhancing the likelihood of growth and success, study 2 examines the causal impact of VC on-site monitoring on key entrepreneurial success outcomes, including innovation activity, follow-on funding, survival, and exits.

Isolating the causal effects of VC monitoring poses significant challenges due to endogeneity and selection bias, as VCs, aiming to maximize returns, typically conduct rigorous screening processes to invest in the most promising startups. To overcome these issues, we leverage exogenous shocks to travel opportunities between lead VCs and their portfolio startups, enabling the identification of monitoring effects. By employing a natural experiment design, we estimate causal effects and provide robust insights into the role of on-site VC monitoring.

The findings generally support our hypothesis that enhanced travel opportunities have positive effects across various success measures, whereas travel disruptions lead to negative effects. However, certain outcomes exhibit no significant or even unexpected differences, which we attribute to increased efficiency within startups, advancements in communication technologies, and government interventions that mitigated some of the negative impacts of the pandemic.

The importance of physical proximity carries significant implications for startups and VCs in their location and investment strategies. Policymakers aiming to foster innovation and employment would benefit from prioritizing the development of transportation infrastructure and improving connectivity between regions to reduce geographical frictions and enhance collaboration opportunities.

In contrast to the theoretical and attitudinal focus of the first study, this research examines the tangible realities of entrepreneurial success, focusing on how external enablers, particularly VCs, influence startups during their critical growth phase. We bridge the pre-venture phase explored in the first study to the growth phase, encompassing both successful exits and potential failures, thereby contributing to a comprehensive understanding of the entrepreneurial life cycle.

1.1.3 Aftermath Following Failure and Entrepreneurial Restart

Building on study 2, which explores the pivotal phase where external enablers and operational dynamics shape startup success, we now transition to the post-venture phase, focusing on instances without successful exit. Failure is an inherent and often unavoidable aspect of the entrepreneurial journey, with most startups unable to achieve a successful exit, making it the most common outcome and a potential endpoint of the entrepreneurial process. Study 3 investigates how entrepreneurial failure is perceived and managed, with a specific focus on how media coverage of failed startups and entrepreneurs influences public narratives and affects the likelihood of re-entering entrepreneurship.

Beyond the cessation of operations, entrepreneurial failure carries significant social dimensions. Media plays a dual role, not only informing public opinion but also constructing the social narratives that frame entrepreneurship (Gamson et al., 1992; Deephouse and Heugens, 2009; McCombs and Shaw, 1972). Startups, and particularly their founders, often face considerable social costs and a loss of legitimacy in the aftermath of failure (Shepherd and Patzelt, 2015). Negative media coverage can exacerbate stigma (Cardon et al., 2011; Sutton and Callahan, 1987), discouraging founders from pursuing new ventures (Shepherd and Patzelt, 2017; Cahn et al., 2021), whereas neutral or supportive coverage may aid recovery and reintegration into the entrepreneurial ecosystem.

Using a large dataset of startup media coverage across multiple countries and applying advanced natural language processing and quasi-experimental methods, we examine how entrepreneurial failure is socially constructed and impacts founders.

The findings reveal that failure is frequently portrayed in a negative light, with significant variation influenced by startup characteristics and societal contexts. Because media narratives surrounding failed startups shape public opinion, they also have profound consequences for entrepreneurs. Founders exposed to harsh media coverage are less likely to attempt re-entry, underscoring the detrimental effects of stigma on entrepreneurial resilience.

The implications extend beyond the individual level. At a societal level, harsh treatment of failed entrepreneurs in the media risks discouraging entrepreneurial activity more broadly, potentially leading to suboptimal entrepreneurship rates and stifling innovation and economic growth. Encouraging balanced or supportive narratives around entrepreneurial failure may therefore be critical to fostering a healthy and dynamic entrepreneurial ecosystem.

By integrating this often-overlooked stage (McGrath, 1999) into the entrepreneurial life cycle, we contributes to a comprehensive understanding of entrepreneurship. While failure often marks the conclusion of a startup's life cycle, it can also serve as a springboard for new entrepreneurial endeavors. Thus, study 3 extends beyond failure's immediate impacts by examining the likelihood of entrepreneurial re-entry, conditioned on the social evaluation of failure. This exploration completes our entrepreneurial life cycle framework, linking the end of one venture to the potential beginning of another.

1.2 Causal Inference

Identifying and estimating causal effects presents significant challenges, particularly in the field of entrepreneurship, which is characterized by its dynamic nature and inherent opaqueness. The outcomes we measure, such as exit rates or news sentiment, are often influenced by the intrinsic quality of a startup, which is difficult to fully capture with observable characteristics. Similarly, entrepreneurial attitudes and decisions are highly heterogeneous, with individuals acting based on complex, underlying reasoning, resulting in many observed relationships being endogenous.

To address these challenges, we place a strong emphasis on causal inference. By employing robust identification strategies and leveraging recent advancements in econometric methods, we aim to infer causal effects from the data. This approach allows us to provide valuable insights into the entrepreneurial process and contribute to the development of more rigorous and actionable entrepreneurship research.

Causal inference is a cornerstone of modern empirical research, enabling the disentanglement of cause-and-effect relationships in both experimentla and observational data. Its significance extends across disciplines such as economics, medicine, and social sciences, where understanding causal mechanisms is essential for informing critical decisions and policies. The importance of this field was underscored in 2021 when the Nobel Prize in Economic Sciences was awarded to Joshua Angrist, Guido Imbens, and David Card for their groundbreaking methodological contributions to the analysis of causal relationships.[1]

Origins of causal reasoning can be traced back to the philosopher David Hume (1711 – 1776), who emphasized in the 18th century that causation cannot be directly observed but must be inferred from consistent associations. The concept of counterfactual reasoning, imagining alternative scenarios where a cause is absent, remains central to modern causal analysis and defines the so-called *fundamental problem of causal inference*.

1.2.1 Fundamental Problem of Causal Inference

The impossibility of observing both the presence and absence of an intervention at the same time for the same observation unit builds the starting point for all considerations abaout causality. It means that causality must always be inferred, not directly measured,

[1] `https://www.nobelprize.org/prizes/economic-sciences/2021/press-release/` [accessed: 2024-12-20]

which is why causal inference is often referred to as a essentially being a *missing data problem*.

To address this formally, the potential outcomes framework, developed by Jerzy Neyman (Splawa-Neyman et al., 1990) and later formalized by Donald Rubin (Rubin, 1974), provides a rigorous approach. It defines causal effects as the difference between observed outcomes under treatment and the unobservable outcomes that would have occurred without treatment. For each unit i, there are two potential outcomes[2], $Y_i(1)$ is the outcome if unit i receives the treatment $D_i = 1$, and $Y_i(0)$ if it does not receive the treatment, hence $D_i = 0$. Consequently, the causal effect for unit i is the difference between these two potential outcomes.

$$\tau_i = Y_i(1) - Y_i(0) \tag{1.1}$$

As $D_i \in \{0, 1\}$, we observe either $Y_i(1)$ or $Y_i(0)$, of which the unobserved outcome is called the *counterfactual outcome*. It follows that the individual treatment effect τ_i is never observable.

1.2.2 Experimental Dataa

A statistic that is better approachable is the average treatment effect (ATE), which takes an average of individual outcomes:

$$\text{ATE} = \mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]. \tag{1.2}$$

However, the missing data problem persists, as we are still not able to observe both states for a particular unit. What we can calculate is the associational difference $\mathbb{E}[Y_i(1)|D_i = 1] - \mathbb{E}[Y_i(0)|D_i = 0]$. However, there is only one setting in which we can assume that the observed associational difference equals the ATE: when the treatment is randomly assigned. This is why randomized experiments, conceptually first developed by R. A. Fisher in 1925 (Fisher, 1992), are referred to as the gold standard of causal inference.

[2] We focus on binary treatment variables. Recent reviews and discussions of continuous and multi-valued treatment variables can be found in Callaway and Sant'Anna (2021) or Lopez and Gutman (2017).

The random assignment of units into treatment groups leads to the *ignorability* or *exchangeability assumption* being fullfilled.[3] It ensures that, on average, treatment group, units with $D_i = 1$, and control group, units with $D_i = 0$, are similar in all respects except for the treatment. Any differences in the outcome can solely be attributed to the treatment. Formalized, the potential outcomes $Y(1)$ and $Y(0)$ must be independent of treatment assignment D, which eliminates any selection bias, where individuals in the treatment group differ systematically from those in the control group.

$$(Y(1), Y(0)) \perp\!\!\!\perp D \tag{1.3}$$

In other words, the observed outcome $\mathbb{E}[Y(0) \mid D = 0]$ of the control group represents an unbiased estimate of the untreated potential outcome of the treatment group, allowing the unbiased estimation of the ATE by directly comparing the observed outcomes. Expressing it with the exchangeability assumption, it guarantees that the treatment and control groups are the same and could be swapped without changing the ATE, or more formally: $\mathbb{E}[Y(1)|D_i = 1] = \mathbb{E}[Y(1)|D_i = 0]$ and $\mathbb{E}[Y(0)|D_i = 1] = \mathbb{E}[Y(0)|D_i = 0]$.

1.2.3 Observational Data

In the absence of randomized assignment to treatment groups, the credibility of the ignorability assumptions is generally compromised. Often, there are systematic differences between the treatment and control groups in terms of their characteristics W which we divide into time-invariant characteristics Z and time-varying characteristics X for the following sections. When these characteristics are also related to the outcome, they are referred to as confounders, resulting in confounding bias. In such cases, the ignorability assumption is violated, undermining the validity of causal inferences.

The obvious solution to address these challenges would be to conduct experiments. However, this is often impractical due to ethical, logistical, or financial constraints.[4] To overcome these limitations, scholars in the field of causal inference have developed various strategies to infer causal effects from observational data. These strategies fundamentally

[3]The ignorability and exchangeability assumption are mathematically equivalent but offer differing perspectives.

[4]For instance, interventions that could potentially threaten the survival of a firm would have ethically unacceptable consequences, making experimental approaches infeasible in such cases.

rely on the principle of conditional exchangeability, also referred to as unconfoundedness:

$$(Y(1), Y(0)) \perp\!\!\!\perp D \,|\, W. \tag{1.4}$$

Intuitively, it implies that ignorability only holds within sub-levels of W. When holding W constant across both groups, the groups become comparable. The calculation of the ATE is performed within each level of W, and then averaged over all levels of W.

$$\text{ATE} = \mathbb{E}_W[\mathbb{E}[Y(1)|W] - \mathbb{E}[Y(0)|W]] \tag{1.5}$$

The estimation is unbiased when we are able to control for all relevant variables W which are related to both treatment and outcome.[5]

1.2.4 Repeated Observations

A significant portion of causal inference research has focused on developing estimators that rely on the principle of conditional exchangeability in cross-sectional studies to address confounding and estimate causal effects. Common approaches include regression adjustment, matching methods, and more advanced methods, such as inverse probability weighting and doubly robust estimators.

However, in the following, we will move beyond the cross-sectional framework and focus on methods that utilize the time dimension, allowing for the relaxation of some of the previously introduced assumptions. In essence, repeatedly observing units over time allows for the elimination of time-invariant heterogeneity, reducing the need to assume unconfoundedness as required in cross-sectional studies.

In the most basic setting with two time periods, there is a pre-treatment period, $t = 0$, and a post-treatment period $t = 1$. The causal estimand now is the average treatment effect on the treated (ATT).

[5] It also requires the overlap/positivity assumption (also referred to as common support). Unfortunately, often, there are unobserved variables which are relevant. For each treated unit with particular values of W, there need to be some untreated units with the same values. Only when for some $\varepsilon > 0$, $P(D_i = 1|W) < 1 - \varepsilon$, and $E[D] > 0$, we can calculate the estimate. We also impicitly require the stable unit treatment value assumption (SUTVA), which demands no interference between units and consistency of the potential outcomes.

$$\text{ATT} = \mathbb{E}[Y(1)|D=1, t=1] - \mathbb{E}[Y(0)|D=1, t=1] \tag{1.6}$$

This quantity remains unobservable because it is impossible to observe the untreated outcome for the treated group in the post-period. However, by leveraging the evolution of outcomes in the control group, we can, under weaker assumptions than before, obtain an unbiased estimate of this quantity.

1.2.5 Difference-in-Differences

For all our studies, we observe at least one pre-treatment and one post-treatment period, allowing us to rely on the widely used method for such settings: difference-in-differences (DiD). By tracking units over time, DiD compares changes in outcomes between treated and untreated groups before and after an intervention. This approach enables us to isolate the treatment effect from confounding factors that are constant over time within groups.

The validity of DiD method hinges on the parallel trends assumption. This assumption requires that, in the absence of treatment, the treated and control groups would have followed similar trajectories over time. Formally, this ensures that any differences observed post-treatment can be attributed to the intervention rather than pre-existing trends.

$$\mathbb{E}[Y_{t=1}(0) - Y_{t=0}(0)|D=1] = \mathbb{E}[Y_{t=1}(0) - Y_{t=0}(0)|D=0] \tag{1.7}$$

$$\mathbb{E}[\Delta Y(0)|D=1] = \mathbb{E}[\Delta Y(0)|D=0]. \tag{1.8}$$

The parallel trends assumption, which underpins the validity of the difference-in-differences method, is fundamentally untestable since it requires knowledge of counterfactual trends for the treated group in the absence of treatment. Therefore, its credibility depends heavily on the specific context.

1.2.6 Identification Strategies and Estimation

This inherent uncertainty challenges us to enhance the credibility of this analyses by employing robust, convincing, and valid identification strategies. In the following, we discuss the identification stategies for the studies, summarized in Table 1.1.

TABLE 1.1: Overview of Studies with Regard to Identification and Estimation

	Level of analysis	Treatment	Outcome	Research Design
Study 1	Individual	Autonomous team formation	Entrepreneurial intent or self-efficacy	Randomized experiment
Study 2	VC-startup dyad	Air travel shock	Various startup success outcomes	Natural experiment
Study 3	Startup	Entrepreneurial Failure	Media sentiment	Quasi-experiment

To estimate the treatment effect, the estimation of the counterfactual is essential. However, the studies differ in the complexity involved in estimating it credibly. While in the experimental setting of study 1, the estimation is most straightforward, the setting in study 2 is a natural experiment with multiple treatment periods, posing threats to the identification strategy, just as the quasi-experimental setting in study 3.

Randomized Experiment

Identification. In the context of study 1, experimental data from both a pre-treatment (baseline) and post-treatment (final) period is available. Because the treatment is randomized, identification is straightforward. Randomization ensures that systematic differences in variables affecting the treatment and outcome are eliminated, allowing for unbiased estimates of causal effects. This research design adheres to the gold standard of causal inference.

Estimation. While it is technically feasible to compute the average treatment effect (ATE) by simply comparing post-treatment outcomes between groups, employing a DiD approach offers additional advantages. DiD accounts for random baseline differences between groups by effectively subtracting them out. The ATE estimate is obtained by comparing changes in outcomes along both the temporal and treatment dimensions. Specifically, the change in outcomes over time for the control group is subtracted from the change in outcomes for the treatment group, yielding a more robust estimate of the treatment effect.

$$\widehat{\tau}_{DiD}^{2\times 2} = \underbrace{(\mathbb{E}[Y|D=1,t=1] - \mathbb{E}[Y|D=1,t=0])}_{\text{Change in Treatment Group}} - \underbrace{(\mathbb{E}[Y|D=0,t=1] - \mathbb{E}[Y|D=0,t=0])}_{\text{Change in Control Group}} \tag{1.9}$$

$$= \mathbb{E}[\Delta Y|D = 1] - \mathbb{E}[\Delta Y|D = 0] \tag{1.10}$$

Instead of manually comparing means, a linear regression is typically employed to estimate the DiD estimate due to convenience, flexibility, and the ability to provide additional statistical insights. In its simplest form, the DiD model can be specified as follows:

$$Y_{it} = \beta_0 + \beta_1 \text{Post}_t + \beta_2 D_i + \tau_{DiD}^{2\times2}(D_i \times \text{Post}_t) + \epsilon_{it}. \tag{1.11}$$

Y_{it} is the outcome for unit i at time t, Post_t is a dummy variable equal to 1 if it is the post-treatment period, D_i is a dummy variable equal to 1 if the individual is in the treatment group, $D_i \times \text{Post}_t$ captures the treatment group in the post-treatment period, and ϵ_{it} is the error term. With this specification, the estimated coefficient for the treatment effect is equivalent to the one obtained through means comparison.

Besides convenience, the regression framework allows testing for statistical significance of our estimates and assessing whether the treatment effect is statistically different from zero. Regression models also enable the use of robust standard errors or clustering of standard errors, which is particularly important when outcomes can vary across individuals or clusters (for instance, student groups). Another advantage of the regression model is that it makes common time trends and initial group differences explicit through β_1 and β_2, respectively.

Natural Experiment

Identification. To estimate the causal effect of VC monitoring in study 2, in an ideal scenario, we would observe the same VC-backed startups both under active VC monitoring and in the absence of such monitoring. However, due to the fundamental problem of causal inference, it is impossible to observe a startup in both its factual and counterfactual states simultaneously.

The closest approach to addressing this issue is to conduct a randomized experiment, as exemplified in study 1. However, this would require the randomized assignment of VC monitoring to startups. Given that a VC's primary goal is to maximize returns and push startups to achieve this objective, such an experiment would necessitate significant financial incentives for VCs to refrain from monitoring startups assigned to the control group. The monetary compensation required to offset the potential loss from reduced monitoring makes this approach practically unfeasible.

Instead, the most valid and feasible strategy is to leverage natural experiments, which mimic the random assignment of a controlled trial. In this study, we exploit exogenous shocks to air travel, which naturally divide the sample into a treatment group (affected by the travel shock) and a control group (unaffected by travel shock). These exogenous shocks are unanticipated and independent of the potential outcomes for VC-startup dyads, effectively simulating the random treatment assignment of a randomized controlled trial.

Because the treatment assignment is as-good-as-random, we could reasonably assume parallel trends between control and treatment groups. However, since treatment assignment is not perfectly randomized, doubts about the validity of this untestable assumption may arise, particularly when there is an imbalance between the groups. Imbalance alone is not necessarily a cause for concern, provided the imbalanced variables are time-invariant and their effect on the outcome remains constant over time (Caetano et al., 2024). However, in practice, some of these assumptions may not hold true.

To avoid relying on this additional assumption and to enhance the plausibility of the untestable parallel trends assumption, practitioners typically adopt the conditional parallel trends assumption. This approach allows for the inclusion of time-invariant variables, such as industry or founding year, to account for systematic differences between treatment and control groups. The conditional parallel trends assumption can be expressed as

$$\mathbb{E}[\Delta Y(0)|D=1,Z] = \mathbb{E}[\Delta Y(0)|D=0,Z], \tag{1.12}$$

where Z represents time-invariant covariates. Similar to the logic of the unconfoundedness assumption in cross-sectional analyses, the conditional parallel trends assumption posits that, after controlling for the levels of Z, the parallel trends assumption holds.

Time-varying covariates. Time-varying covariates X, such as the funding stage of a startup, introduce an additional layer of complexity. When these covariates are potentially affected by the treatment, the unobserved untreated covariate does not match the observed treated covariate. This can bias the estimated counterfactual untreated outcome, as it is derived from a covariate influenced by the treatment. Formally,

$$\mathbb{E}[\Delta Y(0)|Z, X_{t=1}(0), X_{t=0}, D=1] = \mathbb{E}[\Delta Y(0)|Z, X_{t=1}(0), X_{t=0}, D=0]. \tag{1.13}$$

However, $X_{t=1}(0)$ is unobserved for the treatment group. Since some time-varying covariates could be affected by the treatment, the assumption $\mathbb{E}[X_t(0)|D=1] = \mathbb{E}[X_t(1)|D=1]$ is unlikely to hold. To address this, we need to impute $X_t(0)$ for the treated units. A detailed discussion of this issue and the imputation methods is provided in the methodological section of study 2. In essence, we use the untreated units to model the covariate as $X_t(0) = f(Z, X_{t-1}, Y_{t-1})$, where X_{t-1} and Y_{t-1} are lagged covariates and outcomes. This model is then used to estimate the untreated covariate for treated units.

Multiple periods and staggered adoption. So far, we have implicitly focused on the case of a single pre-treatment and post-treatment period. However, in our study, we observe multiple periods for both, which strengthens the identification strategy. Observing several pre-treatment periods allows us to test for parallel trends between the treatment and control groups prior to treatment, providing means to evaluate the plausibility and credibility of the conditional parallel trends assumption. The assumption generalizes to:

$$\mathbb{E}[\Delta Y_{t,t'}(0)|D=1, X_t, X_{t'}, Z] = \mathbb{E}[\Delta Y_{t,t'}(0)|D=0, X_t, X_{t'}, Z] \quad \forall t \neq t', \tag{1.14}$$

where t and t' are two different points in time.

The logic is still the same. In absence of the treatment, treatment and control group evolve the same (conditional on X and Z). Thanks to having more than one pre-period, we can perform pre-trends test. When treatment and control group evolve similarly prior to the treatment, that increases our confidence in the parallel trends assumption. If they are not moving in parallel, it is very implausible to think they would move in parallel following the treatment in absence of treatment. Note, that the assumption itself is still untestable as the parallel trends assumption refers to the post-treatment periods.

The identification strategy can also accommodate variation in treatment timing, meaning that different VC-startup dyads experience the treatment shock at different points in time. In this case, the conditional parallel trends assumption generalizes to account for staggered treatment adoption. Specifically, let there be $t = 1, ..., T$ time periods, and units adopt the treatment at time $G_i \in 1, ..., T \cup \infty$, where $G_i = \infty$ represents the never-treated group. This framework ensures that the assumption holds for comparisons across units with varying treatment timings while maintaining the validity of the identification strategy.

$$\mathbb{E}[\Delta Y_{t,t'}(\infty)|G_i = g, X_t, X_{t'}, Z] = \mathbb{E}[\Delta Y_{t,t'}(\infty)|G_i = g', X_t, X_{t'}, Z] \quad \forall t \neq t' \text{ and}, g \neq g'. \tag{1.15}$$

Estimation. A generalization of the regression specification is the two-way fixed effects (TWFE) model in Equation 1.11, which incorporates fixed effects for both units (α_i) and time periods (λ_t). This approach accounts for unobserved confounders, provided their effects are time-invariant and additive (Roth et al., 2023). The inclusion of time fixed effects captures any time-specific shocks that uniformly affect all units. As a result, time-invariant variables are excluded from the estimation because they lack within-unit variation, making their effects indistinguishable from the unit fixed effects.

$$Y_{it} = \alpha_i + \lambda_t + \tau D_{it} + \beta X_{it} + \epsilon_{it} \tag{1.16}$$

The use of two-way fixed effects (TWFE) models has been ubiquitous for estimating settings with multiple periods and staggered treatment adoption over the past decades. However, recent discussions have highlighted significant issues with TWFE in such contexts when treatment effects are heterogeneous (see e.g., Goodman-Bacon, 2021; de Chaisemartin and D'Haultfœuille, 2023). These discussions reveal that TWFE yields unbiased results only under the restrictive assumption that treatment effects are constant across groups and over time, an assumption that is rarely realistic in practice.[6] To address these limitations, several robust estimators have been proposed. In the methodological section, we examine these issues and detail the strategies employed to overcome them.

[6]Often, treatment effects recover after some time or might need some time to build up.

Quasi-Experiment

Identification. The identification strategy in study 3 presents additional challenges compared to the previous studies. Unlike study 1, there is no randomized treatment assignment, and unlike study 2, there is no natural experiment to exploit. Instead, treatment is induced by business failure, which is inherently non-random and influenced by underlying factors, such as lack of managerial expertise, a bad business model, or external forces. Consequently, we cannot assume similarity between the units in the treatment and control groups as we have to assume that startups in the treatment group are of lower quality. Again, we are not able to observe a startup in both states and a randomized experiment is hardly feasible or even imaginable.

Additionally, our data does not constitute a traditional panel that tracks the same set of startups across multiple periods, as media coverage is irregular, and startups are not reported on every quarter. Instead, the data resembles an imperfect panel or repeated cross-sections. Identification is achieved under the assumption of no compositional changes, meaning that covariates and treatment status remain invariant to the sampling period (Sant'Anna and Zhao, 2020).[7]

Although we observe units across different time periods, it is prudent to strengthen the parallel trends assumption by incorporating covariates. We do not include unit fixed effects because the panel is imperfect with irregular and few observations per startup. Therefore, we address potential biases in treatment assignment by controlling for startup quality, arguably the most critical determinant of failure. This is achieved by including quality-related variables such as founder qualifications and background, raised funding, funding stage, and other relevant metrics.

Estimation. Empirically, study 3 shares the challenge of staggered treatment timing with study 2. However, unlike study 2, where unit fixed effects are included, study 3 controls for outcome-related heterogeneity through the inclusion of covariates. Further methodological details on how these challenges are addressed are provided in the methodological section of this study.

[7] Abadie (2005) emphasizes that treatment status must be known for all units, even when only observed in the pre-treatment period. Furthermore, when using pre-treatment covariates, these must also be available for post-treatment units. In our setting, we meet these requirements but enhance robustness by restricting the sample to units observed at least once both before and after treatment.

The rest of the thesis is structured as follows: Section 2-4 present the three studies and Section 5 closes with an overall discussion and conclusion.

Chapter 2

Shaping the Startup Mindset - How Entrepreneurial Training Transforms Attitudes

2.1 Introduction

Startups play a crucial role in modern economies, driving economic growth and innovation (Audretsch, 2007). Rather than individually, more than half of all nascent entrepreneurs work in teams of two or more and those tend to be overrepresented among high performing startups (Gartner et al., 2004; Kamm et al., 1990; Timmons, 1990). Building teams is economically desirable when it enables gains from complementarities, enhancing productivity and decision-making (Lazear, 1998). Beyond economic efficiency, teaming up reflects an entrepreneur's willingness to expand their effort across a wide range of activities - including creative problem-solving, opportunity identification, and implementation tasks - which are essential for launching a new venture in a fast-paced environment (Lazear, 2005; Timmons, 1990; Vesper, 1980).

Given the prominence of team formation in entrepreneurship, it raises the question whether it constitutes an integral part of the entrepreneurial spirit, the effects on and through team diversity and how it impacts the presence or intensity of peer effects.

Empirically, studying team formation poses challenges due to its inherently endogenous nature. Without careful research design, distinguishing the causal effects of autonomous team formation and peer influence from pre-existing characteristics is difficult.

Our study addresses these challenges through a randomized experiment conducted within an entrepreneurial training course at a German university. This approach allows us to isolate the effects of autonomous team formation on entrepreneurial attitudes, providing insights into how team dynamics shape the early stages of the entrepreneurial journey, particularly the formation of entrepreneurial intention and self-efficacy, key precursors to entrepreneurial action (Krueger and Carsrud, 1993).

Findings from the entrepreneurship literature, mainly studied with regard to performance, deliver mixed results and point to different potential directions of the effect. Autonomous team formation can positively influence entrepreneurial attitudes by providing intrinsic value through self-determination and decision-making freedom (Kiessling et al., 2022; Deci and Ryan, 1985, 2000). More generally, autonomy offers procedural utility, serving as a motivational factor that enhances job satisfaction (Frey et al., 2004; Block and Koellinger, 2009; Benz and Frey, 2008b,a). This ultimately suggests that individuals may even be willing to sacrifice monetary value in exchange for autonomy (Bartling et al., 2014; Owens et al., 2014).

Beyond its intrinsic value, autonomous team formation can enhance team dynamics by fostering greater effort, coordination, and trust, as well as stronger attachment and responsibility among team members (Chen, 2017). Teams that self-select can also prioritize skill complementarity (Chen and Gong, 2018), leading to optimal team compositions compared to randomly assigned teams. This allows individuals to focus on their strengths (Büyükboyaci and Robbett, 2019) while expanding the external range of information, resources, and perspectives (Reagans et al., 2004).

However, team self-selection often occurs based on pre-existing social connections rather than skill diversity (Chen and Gong, 2018; McPherson et al., 2001). This increases internal density, which has been shown to have positive effects on coordination efforts due to familiarity among team members (Reagans et al., 2004; Cattani et al., 2013).

Despite its potential benefits, autonomous team formation also carries the risk of overvaluing familiarity, which can hinder optimal team composition (Seabright et al., 1992; Gulati and Gargiulo, 1999; Ruef et al., 2003; Ingram and Morris, 2007). The tendency for individuals to form teams with those who are similar to themselves - a phenomenon known as homophily - has been associated with negative effects on entrepreneurial performance (Hasan and Koning, 2019; Hoogendoorn et al., 2017; Steffens et al., 2012; Hansen et al., 2015). We contribute to the literature by isolating various effects of autonomous team formation on entrepreneurial attitudes.

Autonomous team formation may also influence the presence of peer effects, which can play a crucial role in shaping entrepreneurial attitudes (Portyanko et al., 2023), as the decision to start a venture is influenced by the social environment (Bosma et al., 2012; Davidsson and Honig, 2003; Engle et al., 2011). Scholars continue to debate the nuances of peer effects in social learning: does exposure to entrepreneurially experienced peers increase the likelihood of becoming an entrepreneur (Nanda and Sørensen, 2010), or does it merely increase the probability of achieving successful outcomes (Lerner and Malmendier, 2013)? We contribute to this debate by examining the role of peer influence on key precursors to entrepreneurial action.

Our findings reveal a positive effect of autonomous team formation on entrepreneurial intention. This effect consists of both a positive intrinsic effect and a negative indirect effect mediated by reduced team diversity. In terms of magnitude, the indirect effect accounts for approximately 40% of the direct effect. However, we do not observe similar results for entrepreneurial self-efficacy. But when testing for peer effects, we find them to be present only for entrepreneurial self-efficacy, and particularly in self-selected teams, while no significant peer effects emerge for entrepreneurial intention. Interestingly, autonomous ideation also enables peer effects, but both forms of autonomy do not add up.

We argue that the direct and indirect effects of autonomous team formation - namely, its intrinsic value, procedural utility, and implications for team diversity - primarily influence the perceived attractiveness of entrepreneurship and satisfaction with the entrepreneurial task. In contrast, peer effects are more likely to shape self-efficacy, as, here, individuals compare themselves with their team members. This effect may be particularly pronounced in self-selected teams, where lower perceived social distance among team members fosters mutual encouragement and motivation.

This research deepens our understanding of the interplay between autonomous team formation, team diversity, and peer effects. The findings have important implications for entrepreneurial training and educational environments, particularly in fostering entrepreneurial mindsets and behaviors.

2.2 Theory

2.2.1 Entrepreneurial Intention and Self-Efficacy

Entrepreneurial intention and self-efficacy are critical components in studying entrepreneurial behavior, serving as key predictors of entrepreneurial actions. According to the seminal Entrepreneurial Event Model (EEM) by Shapero and Sokol (1982), entrepreneurial behavior results from attitudes, specifically perceived desirability, feasibility, and the propensity to act. Krueger and Carsrud (1993) emphasize that entrepreneurial intention is the "single best predictor" of entrepreneurial action, highlighting the importance of an individual's commitment and determination in starting a new business. Subsequent models, including the Theory of Planned Behaviour (TBP) (Ajzen, 1991), introduce the roles of attitudes towards the behavior, subjective norms, and perceived behavioral control, all influencing each other and together determining intention and behavior. However, these models face critique for lacking robust empirical support (Kautonen et al., 2015; Terry et al., 1999). Despite this, there is a consensus among scholars that entrepreneurial behaviors, which often occur under conditions of uncertainty, are heavily informed by an individual's attitudes toward entrepreneurship (Bagozzi et al., 2003; Madden et al., 1992; McGee et al., 2009).

Entrepreneurial intention refers to an individual's conscious commitment and determination to engage in entrepreneurial activities. It reflects both the perceived desirability of entrepreneurship and, to some extent, the propensity to act, as outlined by Shapero and Sokol (1982). Following Ajzen (1991), intention is a direct predictor of behavior.

Entrepreneurial self-efficacy, on the other hand, reflects the belief in one's ability to successfully execute the tasks and roles necessary to start and manage a new business. This attitude corresponds to the perceived feasibility in Shapero and Sokol (1982) and is crucial for overcoming challenges, navigating the complexities of entrepreneurial tasks, and achieving entrepreneurial goals (Bandura, 1982, 2012). Ajzen (1991) refers to it as perceived behavioral control, which affects the decision to act directly but also indirectly, through increasing intention.

Therefore, both entrepreneurial intentions and self-efficacy are indispensable for understanding and predicting entrepreneurial behavior, as they encapsulate the essential motivational and capability-related aspects required for entrepreneurial action.

2.2.2 Autonomous Team Formation

In the context of entrepreneurship, autonomous team formation (ATF) refers to the ability of individuals to self-select their team members, aligning with the typical endogenous process of founder team formation. How ATF shapes entrepreneurial attitudes depends on different channels and mechanisms, working in opposing directions and with different magnitude, sometimes even with potentially ambiguous effects.

Positive Effects of Autonomous Team Formation

ATF can have a direct and positive influence by providing individuals with the freedom to make decisions, thereby shaping perceptions of the entrepreneurial activity and ultimately affecting both performance and behavioral outcomes. In a framed field experiment, Kiessling et al. (2022) find that individuals who self-select their peers performed better than those with randomly assigned peers. Their findings align with self-determination theory, which posits that autonomy is a crucial driver of intrinsic motivation by offering individuals active choices in their task environments (Deci and Ryan, 1985, 2000).

Despite the commonly low expected monetary rewards associated with entrepreneurial endeavors and self-employment, scholars have emphasized the concept of procedural utility, where the intrinsic value of autonomy in such settings serves as motivational factor and significantly contributes to job satisfaction (Frey et al., 2004; Block and Koellinger, 2009; Benz and Frey, 2008a,b). Similarly, numerous studies in the economics literature, particularly laboratory experiments, highlight the benefits of job and work autonomy (for a review see Cassar and Meier, 2018). In contrast, being subjected to exogenous decisions can diminish effort (Falk and Kosfeld, 2006; Bartling et al., 2012; Fehr et al., 2013) and further, Bartling et al. (2014) and Owens et al. (2014) suggest that individuals are even willing to forgo monetary rewards in exchange for greater autonomy.

ATF also has significant implications for team dynamics. In a laboratory experiment, Chen (2017) demonstrates that ATF increases effort and coordination, attributing this to enhanced trust, attachment, and responsibility within self-selected groups. Similarly, Chen and Gong (2018) find self-selected teams to outperform both randomly assigned teams and those formed solely based on skill complementarity. Notably, while teams formed based on skill complementarity outperformed randomly assigned teams, they did not exceed the performance of self-selected teams, suggesting benefits of both skill complementarity and

ATF.

In fact, ATF can be of instrumental value by influencing team composition and the quality of matches. From the set of all theoretically possible groups, optimal matches become feasible, enabling individuals to select teammates with complementary traits, skills, or experiences, thereby enhancing overall team effectiveness. When gains from complementaries are possible, forming teams is economically desirable (Lazear, 1998) and expands the "external range" through demographic diversity, broadening access to information, resources, and perspectives Reagans et al. (2004). Team formation based on skill complementarity also facilitates specialization, thereby enhancing team performance (Büyükboyaci and Robbett, 2019).

However, because ATF does not guarantee optimal complementarity, as individuals often self-select based on social connections rather than diverse skill sets (Chen and Gong, 2018; McPherson et al., 2001). This tendency, known as homophily, leads individuals to associate with others similar to themselves in terms of gender, ethnicity, qualifications, and other characteristics.

Nevertheless, even when familiarity is prioritized over complementarity, ATF can yield positive outcomes. In contrast to the "external range", Reagans et al. (2004) describe how team formation based on familiarity increases "internal density", which is positively associated with team performance. Self-selection can enhance team cohesion and improve task coordination, as familiarity among team members fosters smoother collaboration (Reagans et al., 2004). In the context of software industry, shared work experience enhances team performance (Huckman et al., 2009), while prior collaboration reduces conflicts and improves coordination in the entertainment industry (Cattani et al., 2013). Familiarity with peers' work styles and personalities facilitates communication and task management, while attachment and responsibility toward the group are often stronger in self-selected teams (Chen and Gong, 2018). Notably, Gompers et al. (2017) found that homophily can have partially positive effects by elevating low-performing teams to mid-performing levels, though it does not necessarily propel teams to top performance.

Research of the direct effect of ATF on attitudes is limited, as most literature focuses rather on performance outcomes. Nevertheless, it provides a basis for hypothesis development. The intrinsic value, improved coordiation and increased effectiveness through ATF likekly foster more positive attitudes toward entrepreneurship and enhance both the desire to engage in entrepreneurial activities and the confidence in one's ability to succeed. Thus, we hypothesize:

Hypothesis 1a. Autonomous team formation increases entrepreneurial intent.

Hypothesis 1b. Autonomous team formation increases entrepreneurial self-efficacy.

Negative Effects of Autonomous Team Formation

However, ATF is not necessarily performed with performance or behavior outcome maximization in mind (Chen and Gong, 2018) but familiarity and personal connections are valued instead. The reluctance to build new relationships, known as relational inertia, can result in suboptimal team compositions (Seabright et al., 1992; Gulati and Gargiulo, 1999). Rather than being driven by beneficial differences and optimal complementarities, Ruef et al. (2003) find that ATF for entrepreneurial teams primarily bases on homophily and strong network ties. Ingram and Morris (2007) demonstrate that even highly motivated EMBA[1] students struggle to form connections outside their comfort zones, typically creating networks that do not extend beyond their existing social circles, while Kaiser and Mueller (2013) document founder teams to be more homogeneous at founding and only later become more heterogeneous.

In fact, most studies with regard to homophily and entrepreneurial outcomes reveal negative effects. In a startup bootcamp context, Hasan and Koning (2019) observe that entrepreneurs with many prior ties limit their new interactions, resulting in lower startup performance compared to those with fewer prior ties. Hoogendoorn et al. (2017) find a more nuanced result by showing that teams with intermediate levels of ability dispersion outperform those with very low or very high dispersion and interestingly, the average ability is not related to the outcome. Among Swedish startups, more homogeneous startup teams tend to perform worse in the long-term when considering the operational status and profitability (Steffens et al., 2012).

More generally, Hansen et al. (2015) show that individuals in male-dominant groups learn less. In the context of open-source software projects, Heath et al. (2023) document the diversity of contributors to be extremely low, and being trapped in this "homophily trap" results in adverse effects on continuity and popularity of the projects. In the setting of a garment plant, even with ability held constant, heterogeneity in ability increases

[1]EMBA refers Excutive Master of Business Administration

productivity (Hamilton et al., 2003).

Given these findings, we propose a mediated effect of ATF on entrepreneurial attitudes, with team diversity acting as the mediator. The first path, where individuals self-select into teams based on similarity, is well-established. The second path, focusing on the impact of team diversity and heterogeneity on entrepreneurial attitudes, is less straightforward but crucial. Emphasizing the positive impact of team diversity on entrepreneurial phenomena, we hypothesize that while ATF leads to the formation of more homogeneous teams, this similarity can negatively impact entrepreneurial attitudes, such as intention and self-efficacy.

Hypothesis 2a. The positive effect of autonomous team formation on entrepreneurial intention is attenuated by reduced team diversity.

Hypothesis 2b. The positive effect of autonomous team formation on entrepreneurial self-efficacy is attenuated by reduced team diversity.

Further Effects of Autonomous Team Formation

One particular relevant mechanism that ATF can have an impact on are peer effects. As the decision to become an entrepreneur is a conscious and intentional one, it is also influenced by the social environment (Bosma et al., 2012; Davidsson and Honig, 2003; Engle et al., 2011). The influence of peers, who are defined as individuals within an immediate social context sharing a similar rank and attributes (Tartari et al., 2014), plays a significant role in shaping entrepreneurial intentions and behaviors (Portyanko et al., 2023). In the behavioral model introduced by Ajzen (1991), subject norms are an indicator of behavior, suggesting that social expectations and the willingness to comply with them influence the decision-making process. Interaction with others forms and shapes attitudes, as individuals perceive social pressures and expectations, ultimately affecting the decision to act.

In social sciences, peer effects are understood as the propensity of an individual's behavior to align with that of their group (Manski, 1993). Through social learning from their reference group, individuals observe and emulate the behaviors of those around them, a process ubiquitous in management and entrepreneurship (McDonald and Eisenhardt, 2020; Portyanko et al., 2023) and particularly relevant in educational settings, where peer interactions can significantly impact students' attitudes and behavior.

Estimating peer effects, however, poses theoretical and empirical challenges. Per Manski (1993), similar behavior within a group can be explained by three distinct effects:

(a) *Endogenous effects*: Individuals' behavior is influenced by the behavior of their peers within the group.

(b) *Exogenous (contextual) effects*: Individuals' behavior is influenced by the observable characteristics of their peers (e.g., age, prior achievement), rather than by their behavior.

(c) *Correlated effects*: Similar behavior within a group arises due to shared individual characteristics or common environmental factors, rather than direct peer influence.

Endogenous effects (a) and, occasionally, exogenous effects (b) are of theoretical significance, while correlated effects (c) cannot be considered causal peer effects. Nevertheless, separating these effects in practice remains complex.

Due to the endogenous formation of social ties, as homophily, or the tendency of individuals to associate with similar others (McPherson et al., 2001), complicates the isolation of genuine peer effects from pre-existing similarities (Manski, 1993). When teams are not composed randomly, peers could be similar even without interaction due to assortative matching, which means they were similar to begin with (correlated group effects) and individual outcomes are highly correlated with group average outcomes (Angrist, 2014). Following Lerner and Malmendier (2013), this selection bias can be ruled out by randomly assigning teams. Interestingly, Feld and Zölitz (2017) show that peer effect estimates tend to be biased upwards without random assignment and biased towards zero with random assignment. Additionally, controlling for common factors could help to eliminate the correlated effects.

Another identification challenge in the context of peer effects is the reflection problem that concerns the endogenous effects. It is a form of simultaneity bias because an individual's own attitude and peers' attitudes are determined simultaneously and, in any model, would occur both on the left- and right-hand side. According to Ahern et al. (2014), using peer attitudes determined before the interaction as a proxy for current attitudes addresses the reflection problem.

Researchers have addresses the empirical challenge through natural experiments, randomized controlled trials, and instrumental variable approaches to establish causal relationships. For instance, Lerner and Malmendier (2013), exploiting close-to-random assignment to MBA[2] sections, reveal that while a higher share of entrepreneurial peers might decrease overall entrepreneurship rates, the ventures that do emerge tend to be more successful, indicating a more educated decision.

Given the importance of peers in shaping entrepreneurial attitudes, we hypothesize that peer interactions influence students' entrepreneurial intentions and self-efficacy by aligning them with those of their teammates. However, we expect these effects to differ between teams with ATF and randomly assigned teams. Based on prior arguments suggesting lower team diversity in ATF teams due to homophily, we anticipate that peer effects will be stronger in randomly assigned teams, where greater diversity and a lack of self-selection may enhance exposure to differing perspectives.

Hypothesis 3a. Peer effects with regard to entrepreneurial intention are attenuated in teams with autonomous team formation compared to randomly assigned teams.

Hypothesis 3b. Peer effects with regard to entrepreneurial self-efficacy are attenuated in teams with autonomous team formation compared to randomly assigned teams.

2.2.3 Autonomous Ideation

Beyond autonomy in team formation, autonomy in the ideation process may also yield important benefits, particularly in relation to peer effects. When individuals are assigned tasks - especially those requiring innovation and creativity - their motivation may be lower. In contrast, granting individuals the freedom to choose their own ideas can significantly enhance intrinsic motivation (Lovas and Ghoshal, 2000) and increase engagement (Thompson, 2018). Autonomous ideation allows individuals to align their projects with their interests, backgrounds, and skills, fostering a greater sense of ownership and personal investment. Furthermore, autonomy in ideation may influence self-efficacy by reinforcing individuals' confidence in their ability to develop and execute entrepreneurial ideas (Portyanko et al., 2023).

Given these mechanisms, we expect that autonomous ideation will strengthen the relationship between peer attitudes and an individual's own attitudes. Students who select their own ideas may exhibit higher levels of commitment and engagement, which could lead to deeper social learning processes and stronger peer influence effects. Consequently,

[2] MBA refers to "Master of Business Administration".

we hypothesize that autonomous ideation positively moderates the extent to which peer effects shape entrepreneurial intention and self-efficacy.

Hypothesis 4a. Peer effects on entrepreneurial intention are amplified in teams with autonomous ideation.

Hypothesis 4b. Peer effects on entrepreneurial self-efficacy are amplified in teams with autonomous ideation.

2.3 Data

2.3.1 Sample

We conducted a field experiment in a real-world setting at a public university in Germany. The study duration was 11 weeks, during which undergraduate students were required to take a Business and Entrepreneurship course. The course structure consisted of both lectures and tutorial sessions. During the tutorial sessions, the participants worked in small groups and engaged in weekly, 90-minute team-based sessions focused on developing a business concept. Participants were supported by an experienced entrepreneur and a teaching assistant, both referred to as mentors, who were blind to the experiment. Upon completing the course, the participants presented their final business proposal to business angels, entrepreneurs, and venture capitalists, including a pitch and pitch deck. In addition, we asked all participants to complete two surveys. First, the baseline survey, with a nearly complete response rate, collected information on demographics, entrepreneurial experience, preferred team composition, and current skills. Second, the endline survey, asked participants about their teamwork, satisfaction with the team and idea, newly developed entrepreneurial skills, and overall learning, achieving a 97.5% response rate. Importantly for our paper, we asked participants about entrepreneurship-related attitudes in both surveys. This design enables us to observe changes in entrepreneurial intention and efficacy from the survey administered before and after the experiment. The experiment was conducted for three cohorts in the years 2016 and 2017 and involved 939 participants in 310 teams, randomly assigned to one of four treatments in a two-by-two experimental design: (i) choose both their team and idea, (ii) choose their team members but not the idea, (iii) choose their idea but not their team, or (iv) choose neither team nor idea (see Boss et al., 2023: for elaboration and analysis at the team-level). We excluded 25 participants who discontinued the class before completion, leaving us with 914 participants in teams of approximately three members.

2.3.2 Variables

Dependent Variables

Our dependent variables are self-reported levels of entrepreneurial intention and self-efficacy after the experiment. The level of intention is inferred from a single question asking participants about their likelihood of starting a business or becoming self-employed, with values ranging from 0 to 100 and averaging 32.8 at baseline. The level of self-efficacy is determined by the answers to nine questions and can range from 1 to 7. In our sample, the lowest value is 1.44, while the highest is 7. The average value is 4.4 at baseline.

Independent and Mediator Variables

The independent variable(s) depend on the hypothesis being tested. To test H1-H2, the primary independent and treatment variable indicates whether an individual was part of a team that endogenously formed (ATF = 1) or was randomly assigned (ATF = 0). For H2, we also add various mediators capturing team diversity from the individual perspective, representing how diverse the team members j are from the perspective of focal student i:

- M_1: social diversity. Students were able to name a few students they wished to be in the team with. Social diversity is defined as the relative share of "new faces", i.e. the fraction of previously unknown team members. Being in a randomly assigned team does not necessarily mean that no wishes at all are fulfilled as they could be fulfilled by chance. But students who were allowed to self-select their team members are naturally more likely to have their wishes fulfilled. The list of wished for team members of student i is W_i, and the number of members in i's group are N_i. For every individual i, we calculate how many of the team members j were wished for.

$$M_{1i} = \frac{1}{(N_i - 1)} \sum_{j \neq i} (j \in W_i)$$

- M_2: demographic diversity. To calculate demographic diversity, we include the characteristics age, gender, and nationality/foreigner.[3] For each of these variable p, we measure how distant the focal student i is to the teammates j by taking the absolute distance.

$$d_{pi} = \frac{1}{(N_i - 1)} \sum_{i \neq j} |p_i - p_j|$$

Then, because we want each characteristic to have same scale, we divide by the average of all sample distances $\overline{d_p}$:

$$\tilde{d}_{pi} = \frac{d_{pi}}{\overline{d_p}}$$

Finally, to combine all characteristics, we average:

$$M_{2i} = \frac{1}{P} \sum_{p=1}^{P} \tilde{d}_{pi}$$

- M_3: entrepreneurial diversity. We follow the same procedure as for M_2 in calculating the measure for entrepreneurial attitudes and experience prior to the entrepreneurial training. More precisely, we include entrepreneurial intention, self-efficacy and exposure at baseline.

- M_4: educational diversity. Again, as for M_2 and M_3, we count the students with other study majors than oneself in a team. This measure could comprise several unobservable factors that played a role in choosing the study major. Students who were able to self-select are more likely to end up with teammates that undertake the same course of study and might share more similarities.

For convenience, as a final step, we multiply all mediator variables by 100 so that their average value equals 100.

To check for peer effects in H3-H4, we use the other team members' average initial self-efficacy or entrepreneurial intention levels as the main independent variable, alongside participants' own initial levels and autonomous ideation status.

[3] Note that gender and nationality/foreigner are binary variables: gender is coded as 1 for male and 0 for female, while nationality is coded as 1 for German and 0 for non-German.

Control Variables

In examining the effect of ATF on entrepreneurial attitudes and peer effects, we account for mentor and cohort as fixed effects. Additionally, we control for key individual characteristics, including age, gender, nationality, study major, and prior entrepreneurial exposure.

For analyses involving the mediated effect through team diversity, we apply the same set of control variables. However, for each specification of the mediator, we exclude the variables that were used in its construction to prevent collinearity. Furthermore, all analyses incorporate autonomous ideation as a control variable.

2.4 Estimation Strategy

2.4.1 Difference-in-Differences

The value of interest is obtained by running a linear regression on the interaction of the time Post_t and treatment dimension ATF_i. Both variables are also included as a main effect and in some specifications, we control for a set of control variables represented by Z.

$$Y_{it} = \beta_0 + \beta_1 \text{Post}_t + \beta_2 \text{ATF}_i + \tau(\text{ATF}_i \times \text{Post}_t) + Z\gamma + \epsilon_{it} \tag{2.1}$$

2.4.2 Causal Mediation Analysis

To investigate whether the effect of ATF is suppressed by the tendency of individuals to associate with similar others, we run a causal mediation analysis. Essentially, we want to disentangle the direct and indirect effect related to ATF. Therefore, our model has the following specification:

$$D = \epsilon_D \tag{2.2}$$

$$M = \alpha_0 + \alpha_1 D + \epsilon_M \tag{2.3}$$

$$Y = \beta_0 + \beta_1 D + \beta_2 M + Z\gamma + \epsilon_Y. \tag{2.4}$$

The treatment variable D represents whether an individual chose their own team members (autonomy) or was randomly assigned, with treatment being exogenous. The treatment D affects the mediator M, which we operationalize as measure of team diversity. Both the mediator M and the treatment D directly affect the outcome variable Y, meaning that D influences Y both directly and indirectly through M. Additionally, we include a set of control variables, represented in matrix notation Z.

Using maximum-likelihood estimation, the model's parameters are estimated to identify those that best fit the data. From the estimated coefficients, we can derive

$$
\begin{aligned}
&\text{the total effect} && (2.5)\\
TE &= DE + IE && (2.6)\\
&= \mathbb{E}[Y(D=1)] - \mathbb{E}[Y(D=0)] && (2.7)\\
&= \beta_1 + \beta_2 * \alpha_1, && (2.8)\\
&\text{the direct effect} && (2.9)\\
DE &= \mathbb{E}[Y(D=1,m)] - \mathbb{E}[Y(D=0,m)] && (2.10)\\
&= \beta_1, && (2.11)\\
&\text{and the indirect effect} && (2.12)\\
IE &= \mathbb{E}[Y(D=1,M(D=1))] - \mathbb{E}[Y(D=0,M(D=0))] && (2.13)\\
&= \beta_2 * \alpha_1. && (2.14)
\end{aligned}
$$

The indirect effect IE captures the influence of ATF mediated through M (e.g., team diversity), while the direct effect DE measures the effect independent of M. The total effect TE is the sum of the direct and indirect effects, reflecting the overall influence of ATF on the outcome.

For the identification of DE and IE, we rely on the sequential ignorability assumption. This assumption requires that the treatment is effectively randomized - which is a reasonable expectation given our experimental design - and that there are no unobserved confounders influencing both the mediator and the outcome. Because in our setting, team diversity is solely influenced by ATF in our setting, we consider the sequential ignorability assumption to be plausible.

2.4.3 Linear-In-Means Model

Specification

A common technique to test for evidence of peer effects, which makes relatively weak assumptions about the mechanism of peer influence, and is mostly used in the literature is the so-called linear-in-means regression (e.g. Ahern et al., 2014; Hacamo and Kleiner, 2018; Graham and Hahn, 2005; Lerner and Malmendier, 2013). It assumes that peer influence can be measured by including the average level of attitude of the other team members in the regression, and thus, it treats all peers as equally influential.

Our formal specification of the linear-in-means model, which we run for both outcomes and values of ATF, is

$$Y_{i1} = \beta_0 + \beta_1 Y_{i0} + \beta_2 \overline{Y}_{-i,0} + Z\gamma + \epsilon_i,$$

where Y_{i1} denotes the individual's posterior level of a particular attitude (either entrepreneurial self-efficacy or entrepreneurial intent) and Y_{i0} denotes the individual's prior level. $\overline{Y}_{-i,0}$ denotes the "leave-one-out" average composed of the average of all team members' prior level except for i. According to H3, we expect the coefficient of β_2 to be generally positive, and particular when teams are randomly composed, as it indicates that an individual's attitude learns and adapts to the peer group. Again, we include a set of control variables, represented in matrix notation Z.

About half of the teams were allowed to autonomously select their business idea that they worked on during the class, while the other teams were assigned pre-defined ideas. Our formal specification to test the moderating role as hypothesized in H4 is the following

$$Y_{i1} = \beta_0 + \beta_1 Y_{i0} + \beta_2 \overline{Y}_{-i,0} + \beta_3 D_{i,(\text{Aut. Ideation})} + \beta_4 \overline{Y}_{-i,0} * D_{i,(\text{Aut. Ideation})} + Z\gamma + \epsilon_i,$$

where $D_{i,(\text{Aut. Ideation})}$ denotes whether the team that individual i was assigned to was able to choose an own business idea or was assigned an idea (0 = assigned idea, 1 = autonomous ideation). Both specifications are tested with different sets of control variables. Standard errors are clustered at the team and mentor levels, while team-level fixed effects are excluded to avoid inducing high multicollinearity and redundancy in peer effect measurements.

Problem of Identification

Generally, peer effects could occur in all teams that went through the experiment. Interacting with other team members could lead to updating one's attitudes according to their peers' attitudes. Because we include prior behavior on the right side of the equation, we do not suffer from the reflection problem that describes the simultaneous observation of all individual and group outcomes.

However, due to the discussed identification problems with regard to selection and correlated group effects, our claim to identify causal peer effects is stronger in randomly assigned teams. For teams with ATF, we need to assume that we can control for selection by the inclusion of predetermined attitudes, which is a stronger assumption than for randomly assigned teams. In both specifications, we include mentor and cohort fixed effects to account for shared environmental factors.

However, even with our experimental setting and model specification, we can still not claim whether the peer effect is caused by an individual's attitude being influenced by peer attitudes (endogenous effect) or peers' exogenous characteristics (contextual effect) (Sacerdote, 2011).

2.5 Results

Table 2.1 provides summary statistics of our sample, segmented by the treatment condition ATF. Of the total participants, 449 students were randomly assigned to teams, while 465 students had the opportunity to self-select their team members. The pre-experiment variables are generally well-balanced across both groups, reflecting the randomized assignment. However, we observe a minor discrepancy in the prior levels of entrepreneurial intention between the treatment and control group, which we attribute to random variation but account for due to observing students at base- and endline.[4]

Other imbalances refer to variables used for mediation analysis and determined after treatment assignment and team formation but prior to the endline observation. It shows, as expected, that ATF leads to teams being less diverse, with particular strong effects on the social diversity (measured in fulfilled team wishes as announced in the baseline

[4]Generally, the use of balance test is discouraged for experimental data (Mutz et al., 2019).

TABLE 2.1: Descriptive Statistics

	Random (N=449)		ATF (N=465)		
	Mean	SD	Mean	SD	Diff.
DV					
Intention (post)	34.1	27.6	33.8	27.1	−0.3
Self-Efficacy (post)	4.7	1.1	4.7	1.0	0.0
Prior Values					
Intention (prior)	34.2	24.9	31.4	22.8	−2.8+
Self-Efficacy (prior)	4.5	1.1	4.4	1.0	−0.1
Ent. Exposure (prior)	1.8	1.0	1.7	1.0	−0.1
Diversity					
Social Diversity	108.5	7.0	91.8	29.4	−16.7***
Demographic Diversity	101.6	67.2	98.4	68.8	−3.2
Entrepreneurial Diversity	103.6	43.6	96.5	40.4	−7.0+
Educational Diversity	120.8	45.5	80.0	62.3	−40.8***
Other					
Autonomous Ideation	0.5	0.5	0.5	0.5	0.0
Team Wishes fulfilled	0.0	0.1	0.3	0.5	0.3***
Characteristics					
Age	21.3	1.9	21.5	2.3	0.2
Semester	3.3	1.7	3.3	1.8	0.0
Gender: Male	0.7	0.4	0.7	0.4	0.0
Foreigner	0.1	0.3	0.1	0.4	0.0
Study Major	*N*	*%*	*N*	*%*	
Mechanical Engineering	68	15.1	74	15.9	
Computer Science	59	13.1	68	14.6	
Construction Engineering	65	14.5	58	12.5	
Logistics & Mobility	56	12.5	56	12.0	
General Engineering Science	44	9.8	47	10.1	
Computer Engineering	39	8.7	37	8.0	
Environmental Engineering	25	5.6	29	6.2	
Electrical Engineering	14	3.1	27	5.8	
Chemical Process Engineering	22	4.9	17	3.7	
Bio Process Engineering	16	3.6	14	3.0	
Naval Architecture	15	3.3	15	3.2	
Techno Mathematics	17	3.8	10	2.2	
Mechatronics	9	2.0	13	2.8	

Notes. This table presents summary statistics for the control group (random team) and the treatment group (ATF). Means and standard deviations are reported for continuous and binary variables, while frequencies and percentages are provided for categorical variables. Last column shows a balance test between the groups.

survey) and educational diversity (measured in overlap of study majors). Differences in diversity with regard to the entrepreneurial attitudes and experiences (value distances to team members) are limited and for demographic diversity, we do not observe a meaningful difference, potentially because the sample is rather homogenous. The diversity values present an individual score, with distances and overlap being computed as comparison from the focal individual to the team members. By construction, the value is 100, on average.

2.5.1 Direct Effect of Autonomous Team Formation

Table 2.2 presents the results of our tests for H1a and H1b which predict a positive direc teffect of ATF on entrepreneurial attitudes. We find evidence supporting a positive effect

TABLE 2.2: ATF on Entrepreneurial Attitudes

	Intention			Self-Efficacy		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
ATF	−2.63 +	−2.43 +	−2.45 +	−0.05	−0.03	−0.04
	(1.51)	(1.41)	(1.41)	(0.07)	(0.07)	(0.07)
Post	−0.06	−0.06	−0.06	0.25 ***	0.25 ***	0.25 ***
	(1.04)	(1.04)	(1.04)	(0.07)	(0.07)	(0.07)
Autonomous Ideation			−1.90			−0.05
			(1.49)			(0.06)
ATF x Post	2.43 +	2.43 +	2.43 +	0.05	0.05	0.05
	(1.40)	(1.40)	(1.40)	(0.09)	(0.09)	(0.09)
Num.Obs.	1.828	1.828	1.828	1.828	1.828	1.828
R^2	0.050	0.188	0.052	0.049	0.077	0.049
Mean DV	33.37	33.37	33.37	4.57	4.57	4.57
Mentor FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls		Yes	Yes		Yes	Yes

Notes. This table presents results from regression analyses on entrepreneurial intention and self-efficacy, where *ATF* represents groups with autonomous team formation. Standard errors, clustered at the student group level, are reported in parentheses, and significance levels are indicated as follows: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

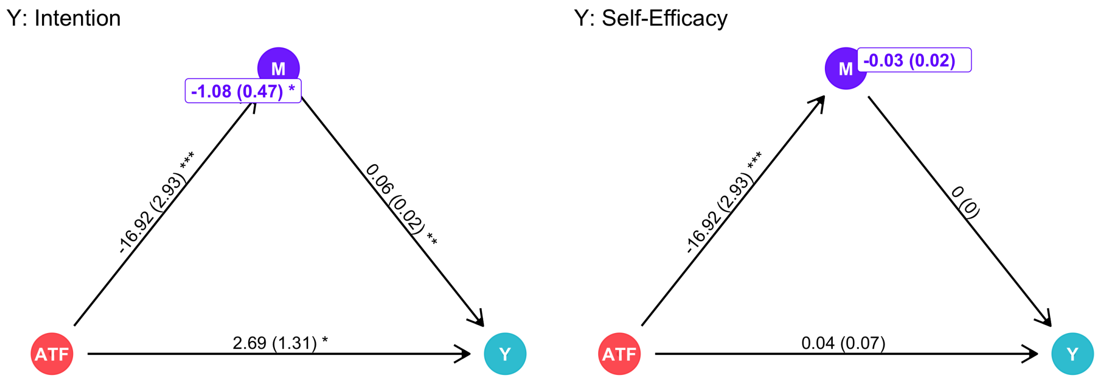
of ATF on entrepreneurial intent. In terms of magnitude, this effect is moderately small, approximately 10% of a standard deviation. As previously mentioned, there was an initial difference in entrepreneurial intention between the treatment and control groups, which balanced out after the experiment. Consequently, the coefficient for ATF is positive - an outcome that, while uncommon in a randomized experiment, can occur due to random variation. Given the significant treatment effect of ATF, we conclude that H1a is supported.

Regarding H1b, which examines entrepreneurial self-efficacy, we do not find any significant effect of ATF. The only statistically significant effect is the *Post* variable, suggesting that, on average, students exhibit higher levels self-efficacy as a result of the entrepreneurial training. This suggests that while the training itself enhances self-efficacy, ATF has no discernible effect. Consequently, we do not confirm H1b.

2.5.2 Mediated Effect through Team Diversity

We further examine the effect of ATF by testing for a mediated effect. H2a and H2b predict that, due to homophily, ATF leads to a decrease in team diversity with negative effects on entrepreneurial attitudes. We present the results in Table 2.3 and Figure 2.1. First, it shows that the path α_1 from ATF to the mediators is as expected: ATF decreases the diversity measured in social, demographic, entrepreneurial, and educational aspects. The second path β_2, from team diversity to the respective attitude, is large and significant only for entrepreneurial intent and social diversity. The aggregated diversity mediator,

FIGURE 2.1: Direct and Indirect Effect of ATF



Notes. This graphs illustrates the direct and indirect effects of ATF, with each path explicitly reported. Treatment variable ATF represents autonomous team formation, which decreases team diversity M_{1-4}, measured as the aggregate of mediators M_1, M_2, M_3, and M_4. Y is the outcome variable, either entrepreneurial intention (left panel) or self-efficacy (right panel). Stars indicate significance level: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

which is an average of the other mediators is also negative and significant. We observe a suppressing effect[5]: while the direct effect of ATF is positive, as expected and confirmed in the previous analysis, there is a negative effect through reduced team diversity. This reduces the positive total effect of ATF by approximately 40%, leaving the total effect positive but no longer statistically significant.[6]

Generally, all analyzed mediations point in the expected direction with $\alpha_1 < 0$ and $\beta_2 > 0$. However, only social and the aggregated diversity show a strong effect and only for entrepreneurial intent. Based on these findings, we find limited support for H2a, but reject H2b.

[5]Suppressing effects are also refered to as inconsistent effects.

[6]Note that the total effect of ATF differs from that in Table 2.2 due to a slightly different model specification.

TABLE 2.3: Causal Mediation Analysis of ATF

	Post Intention					Post Self-Efficacy				
	(M_1)	(M_2)	(M_3)	(M_4)	(M_{1-4})	(M_1)	(M_2)	(M_3)	(M_4)	(M_{1-4})
Direct Effect	2.22	1.63	1.29	1.74	2.69 *	0.03	0.00	0.01	0.02	0.04
	(1.36)	(1.24)	(1.24)	(1.35)	(1.31)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Indirect Effects:										
M_1 (Social Diversity)	−0.89 +					−0.02				
	(0.54)					(0.02)				
M_2 (Demographic Diversity)		−0.09					0.00			
		(0.20)					(0.00)			
M_3 (Entrepreneurial Diversity)			−0.04					−0.01		
			(0.12)					(0.01)		
M_4 (Educational Diversity)				−0.27					0.00	
				(0.50)					(0.03)	
M_{1-4} (Aggregated)					−1.08 *					−0.03
					(0.47)					(0.02)
Total Effect	1.34	1.54	1.26	1.47	1.61	0.00	0.00	0.00	0.02	0.01
	(1.24)	(1.26)	(1.24)	(1.26)	(1.28)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Other Effects:										
Baseline Level	0.71 ***	0.73 ***	0.72 ***	0.72 ***	0.75 ***	0.16 ***	0.16 ***	0.18 ***	0.15 ***	0.16 ***
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Ideation Autonomy	−1.10	−0.96	−1.02	−1.02	−1.17	0.05	0.06	0.06	0.05	0.05
	(1.36)	(1.37)	(1.37)	(1.37)	(1.40)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Partial Effects:										
α_1 ($ATF \rightarrow M$)	−16.67 ***	−3.16	−7.05 +	−40.79 ***	−16.92 ***	−16.67 ***	−3.16	−7.05 +	−40.79 ***	−16.92 ***
	(1.92)	(6.94)	(4.23)	(5.44)	(2.93)	(1.92)	(6.94)	(4.23)	(5.44)	(2.93)
β_2 ($M \rightarrow Y$)	0.05 +	0.03 **	0.01	0.01	0.06 **	0.00	0.00	0.00	0.00	0.00
	(0.03)	(0.01)	(0.02)	(0.01)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Num.Obs.	914	914	914	914	914	914	914	914	914	914
Mean DV	33.96	33.96	33.96	33.96	33.96	4.71	4.71	4.71	4.71	4.71
Mentor FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. This table presents the results of the causal mediation analysis. All variables included in the mediation estimation formula are shown. *Baseline Level* refers to the outcome variable. Control variables include all controls which were not used in creating the respective mediator. Standard errors, clustered at the student group level, are reported in parentheses, with significance levels indicated as follows: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

2.5.3 Peer Effects

H3a and H3b predict that peer effects influence entrepreneurial attitudes, meaning that team members adapt to one another in terms of entrepreneurial intention and self-efficacy. The results are presented in Table 2.4, differentiated by outcome and sample - either teams that were randomly assigned (ATF = 0) or those that were autonomously formed (ATF = 1). We test multiple model specifications, primarily varying the inclusion of interaction effects.

For entrepreneurial intention, we find no evidence of peer effects across any specification. However, for entrepreneurial self-efficacy, we observe a positive peer effect in autonomously formed teams, with coefficients ranging from .15 to .16 in the main models (a) and (b), corresponding to approximately 15% of a standard deviation of the outcome. In contrast, for randomly assigned teams, the coefficient is negative but not statistically significant.

Thus, we do not find support for H3a or H3b as hypothesized. Specifically, there is no evidence of peer effects on entrepreneurial intention (H3a), and for entrepreneurial self-efficacy (H3b), the effect is observed only in autonomously formed teams but appears in the opposite direction for random teams. While we do not confirm H3a or H3b, the findings for H3b reveal an interesting divergence in how peer effects manifest depending on team formation.

Examining the moderation effects of autonomous ideation, as hypothesized in H4a and H4b, we find a positive moderation effect for entrepreneurial self-efficacy in randomly assigned teams but a negative effect in autonomously formed teams. The marginal effects are illustrated in Figure 2.2. Peer effects are strongest under ATF, but they are also positive under autonomous ideation or a combination of both. Notably, under neither ATF nor autonomous ideation, we observe negative peer effects.

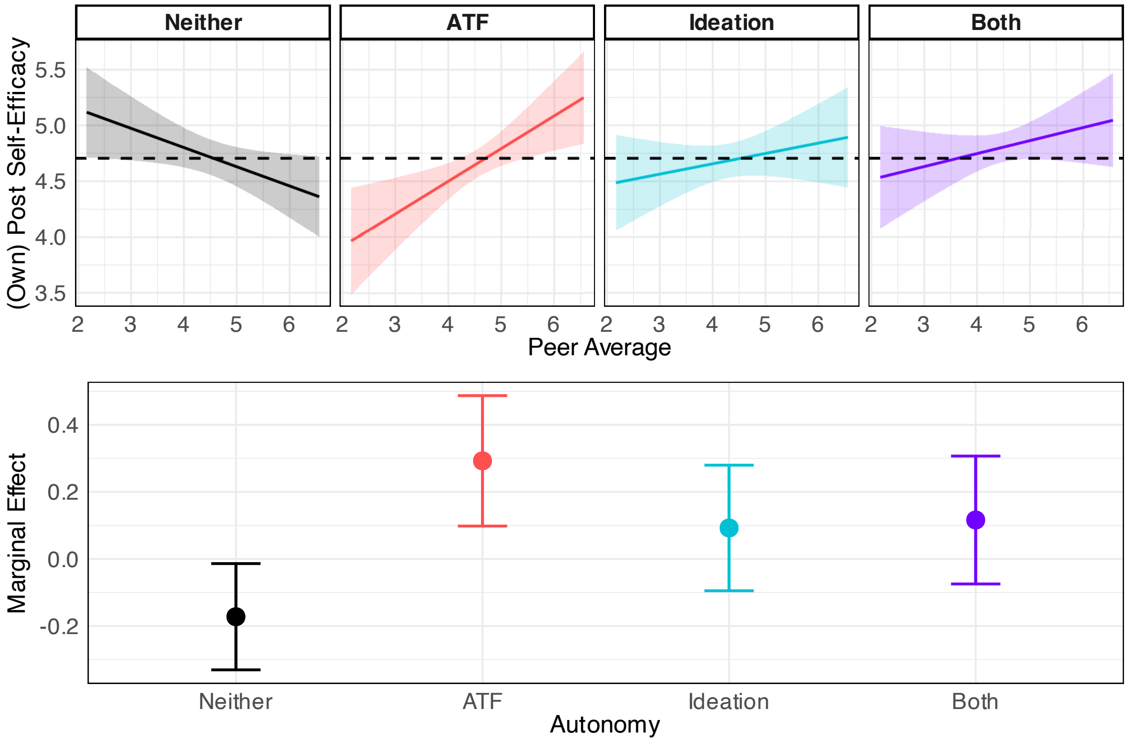
Consequently, we do not find evidence supporting H4a, while for H4b, we observe limited support in randomly assigned teams but an opposing effect in autonomously formed teams. Other tested specifications, such as interacting the peer effect with an individual's baseline level or including a squared peer effect term, also do not produce any significant effects.

TABLE 2.4: Linear-in-Means Model to Test for Peer Effects

Panel A: Post Intention	ATF = 0					ATF = 1				
	(1a)	(1b)	(1c)	(1d)	(1e)	(2a)	(2b)	(2c)	(2d)	(2e)
Baseline Level	0.75 ***	0.72 ***	0.72 ***	0.76 ***	0.72 ***	0.75 ***	0.71 ***	0.71 ***	0.69 ***	0.71 ***
	(0.05)	(0.05)	(0.05)	(0.09)	(0.05)	(0.05)	(0.06)	(0.06)	(0.08)	(0.05)
Peer Average	0.03	0.04	0.01	0.08	−0.12	0.03	0.05	0.03	0.03	0.01
	(0.06)	(0.05)	(0.07)	(0.08)	(0.19)	(0.06)	(0.06)	(0.10)	(0.10)	(0.19)
X Ideation Autonomy			0.08					0.05		
			(0.11)					(0.12)		
X Baseline Level				0.00					0.00	
				(0.00)					(0.00)	
X Peer Average					0.00					0.00
					(0.00)					(0.00)
Num.Obs.	449	449	449	449	449	465	465	465	465	465
R^2	0.512	0.529	0.529	0.529	0.530	0.476	0.488	0.488	0.488	0.488
Mean DV	33.96	33.96	33.96	33.96	33.96	33.96	33.96	33.96	33.96	33.96
Mentor + Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Panel B: Post Self-Efficacy	**ATF = 0**					**ATF = 1**				
	(3a)	(3b)	(3c)	(3d)	(3e)	(4a)	(4b)	(4c)	(4d)	(4e)
Baseline Level	0.17 **	0.11 +	0.11 +	0.18	0.11 +	0.18 **	0.14 *	0.14 *	0.23	0.14 *
	(0.06)	(0.06)	(0.06)	(0.35)	(0.06)	(0.06)	(0.06)	(0.06)	(0.32)	(0.06)
Peer Average	−0.08	−0.08	−0.18 *	−0.01	−0.01	0.16 *	0.15 *	0.25 *	0.24	0.45
	(0.07)	(0.07)	(0.09)	(0.35)	(0.53)	(0.07)	(0.07)	(0.12)	(0.30)	(0.64)
X Ideation Autonomy			0.22 +					−0.17		
			(0.13)					(0.15)		
X Baseline Level				−0.02					−0.02	
				(0.08)					(0.07)	
X Peer Average					−0.01					−0.03
					(0.06)					(0.08)
Num.Obs.	449	449	449	449	449	465	465	465	465	465
R^2	0.127	0.182	0.186	0.182	0.182	0.172	0.193	0.196	0.193	0.193
Mean DV	4.71	4.71	4.71	4.71	4.71	4.71	4.71	4.71	4.71	4.71
Mentor + Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes

Notes. This table presents the results of linear-in-mean models used to test for peer effects. *Peer Average* refers to the leave-one-out average and "X" interaction with it. Panels show different outcomes and are divided by the value of ATF. Standard errors, clustered at the student group level, are reported in parentheses, with significance levels indicated as follows: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

FIGURE 2.2: Peer Effect Moderation by Autonomy Dimension



Notes. This figure illustrates the peer effect on entrepreneurial self-efficacy, moderated by the level of autonomy. Individuals in the sample vary in their degree of autonomy: some were allowed to choose both their team and idea, others only one of these, and some had no autonomy in either dimension. The dashed line in the top row represents the average post-treatment level of self-efficacy. Confidence intervals are depicted at the 95% level.

2.6 Discussion

Our results indicate that ATF positively influences entrepreneurial intention but has no significant impact on entrepreneurial self-efficacy. Autonomy in team selection - a crucial aspect of the entrepreneurial process - likely enhances students' motivation and enjoyment during training. However, while this increased enjoyment may arise from the inherently autonomous nature of entrepreneurial work, it does not necessarily lead to a higher perception of one's own abilities. A possible explanation is that autonomy may introduce greater uncertainty, which could dampen students' confidence in their skillset and ultimately neutralize the overall effect. In conclusion, while ATF fosters entrepreneurial intention by reinforcing the motivation to start a business, it does not strengthen individuals' belief in their ability to succeed.

Building on the positive effect of ATF on entrepreneurial intention, we find that this effect is partially offset by the lack of diversity in ATF teams. These teams tend to be more homogeneous across various social, entrepreneurial, and educational dimensions, which negatively impacts entrepreneurial intention. This suggests that while students benefit from ATF by directly strengthening their entrepreneurial intention, they simultaneously face limitations due to reduced team diversity. A possible explanation for this suppressing effect is the limited exposure to diverse perspectives and problem-solving approaches (e.g., students from different academic disciplines), which might otherwise challenge assumptions, expand thinking, and foster greater entrepreneurial motivation.

However, while we also observe an overall indirect negative effect of aggregated team diversity, our analysis reveals that the primary driver of this effect is social diversity, specifically the existence of fulfilled team preferences. A likely explanation for this pattern is the relative homogeneity of the sample across other diversity dimensions. Given that participants in our study share similar educational backgrounds, career aspirations, and levels of entrepreneurial exposure, variations in age, gender, study major, or prior entrepreneurial experience may be too minor to meaningfully influence team dynamics or outcomes.

We extend our analysis by examining whether team members adapt their entrepreneurial attitudes to one another following entrepreneurial training. Our findings provide partial support for the existence of peer effects, revealing nuanced relationships between ATF, autonomous ideation, and the specific attitudes susceptible to peer influence.

Compared to entrepreneurial intention, self-efficacy appears to be more socially driven and, therefore, more susceptible to peer influence. As self-efficacy relates to an individual's perception of their own abilities, it can be shaped by interactions with teammates. In contrast, entrepreneurial intention seems to be more individually determined, and the social interactions within teams may not be strong enough to significantly alter this foundational belief over the duration of the study.

However, even the presence of peer effects on entrepreneurial self-efficacy is not consistent across all conditions. Peer effects emerge only in self-selected teams and are absent in randomly assigned teams, suggesting that individuals in self-selected teams may feel more connected, experience less social distance, and engage in mutual reinforcement of confidence, leading to stronger self-efficacy effects. Conversely, in randomly assigned teams, weaker social bonds and lower familiarity may limit the extent to which individuals influence one another's self-perceptions.

When incorporating autonomous ideation as a second dimension of autonomy, we uncover an important finding: peer effects require some level of autonomy to emerge but do not accumulate when multiple autonomy dimensions are introduced. They are strongest when only ATF is present, while any additional autonomy (autonomous ideation or a combination of both) leads to only a moderate effect. Notably, neither form of autonomy present results in negative peer effects. This suggests that the benefits of autonomy are not necessarily additive and that different autonomy dimensions shape team dynamics and entrepreneurial attitudes in complex ways.

2.6.1 Empirical Contributions

We contribute to the ATF literature by shifting the focus from performance outcomes, which have been the primary focus of prior studies (see e.g. Boss et al., 2023; Bailyn, 1985; McPherson et al., 2001), to the examination of entrepreneurial attitudes. We also complement existing research on team diversity (Seabright et al., 1992; Gulati and Gargiulo, 1999; Ruef et al., 2003; Ingram and Morris, 2007; Huckman et al., 2009; Cattani et al., 2013; Hoogendoorn et al., 2017; Gompers et al., 2017; Steffens et al., 2012).

By shifting the focus to entrepreneurial intention and self-efficacy, we build on established models that explain the decision to pursue entrepreneurship, such as the EEM (Shapero and Sokol, 1982) and the TBP (Ajzen, 1991). We conceptualize entrepreneurial intention and self-efficacy as key attitudes leading to entrepreneurial action (Krueger and Carsrud, 1993), shaped in part by autonomy, team characteristics, and the social environment.

This also allows us to contribute to the literature on peer effects in the entrepreneurial context (Nanda and Sørensen, 2010; Lerner and Malmendier, 2013) and to address the call by Portyanko et al. (2023) to explore whether peer effects in the pre-formation phase, among non-entrepreneurial peers, can foster entrepreneurial intentions.

2.6.2 Practical Implications

Our findings are particularly relevant in settings where individuals are in the decision-making phase regarding entrepreneurial entry, but can be applied across various contexts to enhance both entrepreneurial intention and self-efficacy.

In the specific context of entrepreneurial training, such as the university course where our experiment was conducted, program designers could utilize our findings by incorporating ATF to boost entrepreneurial intention among participants. Allowing students to select their own teams can increase motivation and engagement in entrepreneurial activities, creating a more supportive environment for cultivating entrepreneurial mindsets.

For accelerators and incubators, our results highlight the importance of carefully managing peer dynamics. We find that peer effects, particularly on self-efficacy, emerge most strongly in self-selected teams, suggesting that the ability to choose teammates fosters an environment where individuals influence each other's confidence and perceived abilities. This implies that allowing ATF in such settings can enhance entrepreneurial behavior by leveraging peer interactions more effectively.

However, our findings also highlight the suppressing effect of reduced team diversity on entrepreneurial intent. This calls for a balance between autonomy and diversity. Program designers in educational or professional settings should encourage diverse team formation while still providing autonomy in team selection, as diversity is crucial for exposing individuals to different perspectives and problem-solving approaches, which in turn fosters greater entrepreneurial motivation.

By understanding the interplay between ATF, peer influence, and autonomous ideation, educators and program designers can create environments that more effectively nurture entrepreneurial mindsets and behaviors. These findings are not only applicable in educational settings but can also be extended to organizational contexts, such as fostering entrepreneurial intention through intrapreneurship programs. In organizations, allowing employees to autonomously form teams and select projects while promoting diversity could drive both innovation and entrepreneurial behavior internally.

Policymakers and program designers for entrepreneurial incubators and accelerators should consider how to grant autonomy while simultaneously promoting team diversity to maximize entrepreneurial intention. These insights can be leveraged to shape environments that are more conducive to entrepreneurship, both in academic settings and within organizations aiming to foster an entrepreneurial spirit.

2.6.3 Limitations and Future Research

One limitation of this study is the relative homogeneity of the sample, as participants were university students, which may limit the generalizability of the findings to more diverse populations and partially explain the modest effect sizes observed. Future research in more heterogeneous contexts could provide additional insights.

Additionally, the short time frame for team formation (students had 60 minutes but, on average, formed teams in just 13 minutes) may have constrained the ability to fully realize the benefits of ATF. In real-world entrepreneurial settings, team formation likely takes longer, with individuals giving more thoughtful consideration to their choices. Future studies could simulate more complex team formation processes to better capture these dynamics.

While this study reveals the direct and indirect effects of ATF, it does not provide specific recommendations for designing teams to optimally foster entrepreneurial attitudes. Future research should explore the ideal balance between team diversity and autonomy to fully understand the potential of autonomy in entrepreneurial settings.

Chapter 3

Being Close to Succeed? Venture Capital Monitoring and Startup Success

3.1 Introduction

Does venture capital (VC) play a significant role in fostering startup growth and innovation beyond providing financing? VC-backed startups are often observed to grow rapidly, demonstrate higher levels of innovation, and secure larger funding amounts compared to their non-VC-backed counterparts (see e.g. Ando, 2024). As of December 2024, all five of the most valuable public companies by market capitalization (Apple, NVIDIA, Microsoft, Amazon, and Alphabet (Google)) received VC funding during their early stages.

It is frequently asserted that VCs do not contribute only financial resources but also managerial and technical expertise, commonly referred to as "monitoring". Effective VC monitoring can provide startups with essential resources, guidance, and oversight, enabling them to overcome early-stage challenges and achieve successful outcomes. Anecdotal evidence, however, does not constitute empirical evidence supporting a causal claim.

Isolating the causal impact of VC monitoring presents a considerable challenge, as VCs follow a rigorous screening processes to invest in startups with the largest growth potential. It might be the case, that VCs merely select companies that are already well-positioned to innovate and succeed, regardless of their active involvement. Simple comparisons between VC-backed and non-VC-backed startups are therefore prone to overestimating the effects of VC involvement.

Various methodological approaches have been used in existing research to examine the causal impact of VC involvement on entrepreneurial success. Most studies document positive effects, with many identifying VC reputation as a key driver of these outcomes. Some compare VC-backed and non-VC-backed startups by matching them on observable characteristics (Hellmann and Puri, 2002; Chemmanur et al., 2011; Puri and Zarutskie, 2012; Bertoni et al., 2011), while others apply more sophisticated estimators that, despite their complexity, rely on similar assumptions (Sørensen, 2007; Quas et al., 2021).

Building on a growing body of research that leverages natural experiments to address endogeneity concerns and strengthen causal identification (Bernstein et al., 2016; Kortum and Lerner, 2000; Bronzini et al., 2020), this study exploits exogenous shocks to travel accessibility between lead VCs and their portfolio startups. Specifically, we examine whether VC monitoring, particularly on-site monitoring, plays a crucial role in driving startup innovation and success.[1]

Our approach most closely follows Bernstein et al. (2016), but unlike previous studies, we leverage two distinct sources of exogenous variation: (i) the introduction of direct flight routes between investors and startups, representing a positive shock, and (ii) the widespread cancellation of flight routes during the COVID-19 pandemic, representing a negative shock. These shocks provide a unique opportunity to analyze the effects of both improved and restricted physical travel, enabling us to estimate the causal impact of on-site VC monitoring on key entrepreneurial outcomes, including innovation activity (measured by patent applications), follow-on funding, survival, and exit rates.

Furthermore, our sample period (2010–2023) spans a time of rapid advancements in communication technologies, allowing us to assess the continued relevance of on-site VC monitoring despite the increasing availability of remote collaboration tools. This contributes to the broader debate on whether physical proximity remains essential in an era of purportedly diminishing geographic constraints.

Our analysis reveals that flight introductions generally have positive effects, while flight cancellations exhibit negative effects. However, these effects do not perfectly mirror each other, likely due to the differing nature of the shocks. Flight introductions represent

[1]Exogenous travel shocks have also been used to study scientific collaboration (Catalini et al., 2020), firms' learning at conferences (Baruffaldi and Poege, 2024), and plant-level investment decisions by headquarters (Giroud, 2013).

or are perceived to be a permanent improvement in travel accessibility, whereas flight cancellations during the pandemic were temporary disruptions.

Consistent with prior research, we find that flight introductions positively influence exit rates and innovation activity (Bernstein et al., 2016). Additionally, we observe a significant positive effect on survival rates (Puri and Zarutskie, 2012). In contrast, travel disruptions caused by flight cancellations primarily result in fewer follow-on funding rounds and a lower exit rate. No significant effects are observed for innovation activity or survival rates. This lack of significance for innovation activity may be explained by a time delay and the limited analysis horizon, while the resilience in survival rates may be attributed to strong governmental support programs for startups during the pandemic, which helped many sustain operations despite disruptions.

Unlike prior studies, we do not find VC reputation to be a key driver of the observed effects, potentially suggesting that signaling benefits of renowned VCs do not depend on physical proximity. However, we find that VC-startup dyads with longer relationships are, on the one hand, better positioned to leverage the benefits of improved travel options and, on the other hand, more resilient to sudden geographical frictions.

This study offers several practical implications. For VCs, the findings underscore the importance of considering geographical proximity in their investment decisions to optimize the potential for effective monitoring. Startups, on the other hand, may benefit from strategically locating in areas with better travel accessibility to attract and obtain most efficient VC investment. For policymakers, the results highlight the value of designing entrepreneurial ecosystems that facilitate face-to-face interactions, such as enhancing transportation infrastructure or fostering innovation hubs around well-connected regions. These efforts can help amplify the benefits of physical proximity in promoting startup success and innovation.

3.2 Theory

3.2.1 Disentangling VC Monitoring from VC Screening

The impact of VC investment on startup performance remains a central question in entrepreneurship and entrepreneurial finance research. While VC-backed firms generally outperform non-VC-backed counterparts[2], establishing causal treatment effects is challenging due to endogeneity concerns. VCs actively screen and select startups with better growth prospects, meaning that simple comparisons between VC-backed and non-VC-backed firms are likely biased. To address this, researchers have employed various methodologies to disentangle screening and monitoring effects.

For instance, Puri and Zarutskie (2012) examine the lifecycle dynamics of VC-backed and non-VC-backed firms using a matched sample and document lower failure rates among VC-backed firms. These differences are most pronounced in the initial years following investment. VC-backed firms tend to achieve greater growth in size but not necessarily in profitability, and performance gaps narrow in the post-internet bubble years (1998–2000) without completely disappearing.

Chemmanur et al. (2011) employ a range of advanced econometric techniques, including switching regressions with endogenous switching, regression discontinuity, and propensity score matching, to estimate both screening and monitoring effects. Their findings show that VC-backed startups exhibit higher pre-investment efficiency and faster efficiency growth post-investment, ultimately leading to higher probabilities of successful exits. These gains are attributed to improvements in sales and, for startups backed by high-reputation VCs, reductions in production costs.

Highlighting the role of VC experience, Sørensen (2007) uses a two-stage estimation approach to model the sorting process. He finds that startups supported by experienced VCs are significantly more likely to go public, with screening accounting for two-thirds of the observed effect and monitoring contributing the remaining one-third. In contrast, using a sample of Italian startups and generalized method of moments (GMM) estimators, Bertoni et al. (2011) identify monitoring as the primary driver of growth among VC-backed companies, with screening playing a negligible role.

Focusing on the non-financial contributions of VCs, Quas et al. (2021) compare Spanish firms that received either VC financing or participative loans from government agencies. Using a difference-in-differences approach to account for unobserved heterogeneity, they find that firms backed by experienced VCs experience substantial yearly growth, including

[2] Ando (2024) show that VC-backed firms contributed 10.6% of net employment growth and 15.8% of net payroll growth between 1990 and 2019, despite representing only 0.2% of all firms. Furthermore, 22% of VC-backed firms established from 1980 to 2010 achieved high-growth status (annual average employment growth $>$ 20% in the first five years and surpassing 100 employees within 20 years), compared to just 1% of non-VC-backed firms. VC-backed firms also exhibit 14% higher R&D expenditure relative to revenue and secure larger funding.

a 12.86% increase in employment, a 38.13% increase in total assets, and a 54.03% increase in sales, effects attributed to the non-financial resources provided by VCs.

Further enhancing the validity of causal claims, Kortum and Lerner (2000) leverage a natural experiment by examining the impact of a 1979 policy shift that spurred VC fundraising. They find that increased VC activity significantly boosts patenting rates. Despite accounting for only 3% of total R&D expenditures, VC-backed firms contribute an estimated 8% of industrial innovations. Similarly, Bronzini et al. (2020) use confidential VC data to compare trajectories of VC-backed startups with those rejected late in the screening process. They find that both groups are initially similar, but VC-backed startups achieve greater growth in employment and innovation activity, albeit at the expense of profitability.

Bernstein et al. (2016) also leverage a natural experiment, exploiting exogenous variation in travel times resulting from new airline routes connecting VCs with their portfolio startups. Their findings demonstrate that increased on-site VC involvement enhances both the scale and quality of innovation while also improving the likelihood of successful exits. Among existing studies, their approach is the most closely aligned with ours.

3.2.2 Facets of VC Monitoring

In surveys, venture capitalists themselve assess selection to be most important when compared to sourcing (pre-selection) and monitoring (Gompers et al., 2020). However, VC employees also stress the importance and variety of monitoring tasks they undertake. According to a survey by Gorman and Sahlman (1989), venture capitalists report spending approximately half of their time monitoring portfolio startups. For each company on whose board they serve, VCs typically dedicate around 80 hours of on-site time and 30 hours of phone time per year. In the first six months after investment, more than 85% of VCs interact with portfolio startups several times a week (Gompers et al., 2020). These tasks include assisting with additional fundraising, conducting strategic analysis, and recruiting management (Gorman and Sahlman, 1989). In line with these findings, a survey by Bottazzi et al. (2008) shows that VCs actively engage in recruiting managers and directors and aid in fundraising, thereby positively influencing the success of their portfolio startups. According to Gompers et al. (2020), VCs report offering a wide range of services to their portfolio startups after investment, including strategic guidance (87%), investor connections (72%), customer connections (69%), operational guidance (65%), board member recruitment (58%), and employee recruitment (46%). Interestingly, they find little variation in the level of involvement across their 444 respondents. These insights are consistent with the responses from current and former VC representatives in a survey by Bernstein et al. (2016).

In empirical analyses, Hellmann and Puri (2002) demonstrate that VC investment is associated with increased professionalization, providing evidence that VCs play a role beyond that of a traditional financial intermediary. The measures implemented by VCs include instituting human resource policies, adopting stock option plans, hiring a marketing VP, or even replacing the CEO. Additionally, Hellmann and Puri (2000) show that VCs influence product market strategies, with VC investment significantly reducing the time required to bring a product to market, particularly for innovative firms. The influence of VCs through the recruitment of managers and directors is highlighted by Amornsiripanitch et al. (2019), with evidence showing that, in most IPOs, VC-backed companies have typically replaced the founder as CEO (Baker and Gompers, 2000). Ewens and Marx (2018) also show VCs to improve performance by replacing founders.

3.2.3 Effectiveness of VC Monitoring

Proximity and Face-to-Face Communication

The rapid advancement of communication technologies has fueled predictions of the "death of distance" (Cairncross, 2001), as digital tools, particularly virtual meetings, offer the potential to replace in-person interactions. While these tools provide flexibility, cost savings, and instant connectivity, they fall short in replicating several critical aspects of face-to-face engagement. In-person interactions facilitate stronger interpersonal connections, enable non-verbal communication, and promote spontaneous idea generation - particularly for complex or nuanced discussions, which are often best conveyed in direct, physical settings (Daft and Lengel, 1986).

Storper and Venables (2004) emphasizes the indispensable role of face-to-face contact in communication, incentive alignment, socialization, and learning, as well as in providing psychological motivation. These aspects are particularly crucial in environments characterized by imperfect, rapidly evolving, and uncodified information - key features of creative and knowledge-intensive activities. Moreover, innovation-related problem-solving frequently occurs in geographically "sticky" locations, where transferring tacit knowledge remains costly and challenging (Von Hippel, 1994). Consequently, telecommunications technology is better understood as a complement to, rather than a substitute for, in-person interactions (Gaspar and Glaeser, 1998).

Neither does empirical support the notion of the "death of distance." For instance, urban areas, rather than rural ones, are more likely to adopt new technologies (Forman et al., 2005), and internet users tend to prefer websites from geographically closer countries, even after controlling for language (Blum and Goldfarb, 2006). Similarly, physical distance continues to constrain online trade, and crowdfunding projects attract disproportionately more local funders (Agrawal et al., 2015). Paradoxically, advancements in communication technology have, for instance, in context of scientific knowledge production, reinforced the importance of co-location (Agrawal and Goldfarb, 2008).

Studies on investment and collaboration further highlight the continued significance of geographical proximity. Giroud (2013) shows that reducing travel time via new airline routes increases plant-level investment and productivity. In the context of remote work, Yang et al. (2021) demonstrate that firm-wide remote setups cause collaboration networks to become more static and siloed, limiting the exchange of novel ideas. Remote work also reduces synchronous communication while increasing asynchronous exchanges, making it more difficult for employees to acquire and share new information across the organization. Similarly, Catalini et al. (2020) provide theoretical and empirical evidence of the significant role travel costs play in shaping scientific collaboration decisions. Their quasi-experimental study, leveraging the introduction of new airline routes, finds that reduced travel costs increase collaborations by 0.3 to 1.1 times and shift the nature of these partnerships toward higher-quality and more novel projects.

In the context of venture capital (VC) investments, exerting influence and monitoring is more efficient and cost-effective when portfolio startups are geographically closer. As a result, VCs are more likely to take board seats at nearby firms (Lerner, 1995) and tend to cluster in regions with high levels of VC-backed investment activity (Chen et al., 2010).

A large-scale survey by Bernstein et al. (2016) underscores the importance of travel costs in VC monitoring. Respondents report that direct flights facilitate greater interaction with portfolio startups, improve advisory effectiveness, enhance value creation, and provide deeper insights into company challenges. Additionally, VCs highlight that direct flights increase travel frequency and flexibility, strengthen communication, and foster more effective long-term relationships.

Building on the discussed findings, we base our first hypothesis on the idea that ease of travel plays a critical role in a VC's ability to positively impact a startup's prospects and growth. When travel costs and times are reduced, VCs can perform more frequent on-site monitoring, which likely enhances their capacity to provide effective oversight and strategic input. Thus, the introduction of new airline routes is expected to intensify monitoring efforts, resulting in improved success prospects for the portfolio startup.

Hypothesis 1a. The introduction of a new airline route for a VC increases the success prospects of its portfolio startup.

Conversely, when travel costs and times increase, on-site monitoring becomes more challenging, potentially leading to adverse effects on the portfolio startup.

Hypothesis 1b. The cancellation of an airline route for a VC decreases the success prospects of its portfolio startup.

VC Reputation

The reputation of a VC is widely recognized as a critical factor influencing the outcomes of portfolio startups. Prior research underscores the heterogeneity of VC monitoring and its impact, with VC reputation often being central source.

High-reputation VCs are often associated with superior monitoring capabilities and better selection of investment opportunities. For instance, Sørensen (2007) demonstrates that experienced VCs are more likely to select investments that succeed, with their influence extending beyond selection to include enhanced monitoring, access to extensive networks (Hochberg et al., 2007), and stronger market signaling. These signals, such as backing by a reputable VC, may indicate unobserved qualities of the startup to other investors (Pollock et al., 2010). For an exit via IPO, VC plays a certification role to reduce information asymmetries according to Megginson and Weiss (1991). Also, the networks of high-reputation VCs have been shown to improve the survival and exit rates of their portfolio startups (Hochberg et al., 2007).

Further evidence suggests that high-reputation VCs deliver significant non-financial value to their portfolio startups or, at the very least, are widely perceived as doing so. For example, Hsu (2004) finds that entrepreneurs are willing to accept discounted equity offers (10-14%) from reputable VCs, expecting them to deliver greater "extra-financial" value.

This value includes facilitating affiliations with reputable partners, recruiting high-caliber talent, and enhancing the startup's market credibility. Similarly, Nahata (2008) shows that startups backed by more reputable VCs have a higher likelihood of successful exits and faster access to public markets, reflecting these VCs' expertise in both selection and monitoring. Similarly, Clarysse et al. (2013) report higher trade sales for portfolio startups of VCs experienced in trade sales, either by own experience or experience learned from syndicate partners.

The benefits of VC reputation also extend to operational improvements. Chemmanur et al. (2011) report a larger increase in total factor productivity for startups supported by high-reputation VCs, attributing this to superior monitoring abilities that constrain production costs. Interestingly, their findings suggest that startups selected by low-reputation VCs are initially more efficient, further highlighting the transformative impact of high-reputation VCs.

Empirical evidence also suggests that reputation serves as a proxy for the non-financial resources VCs provide. For example, Quas et al. (2021) find that only the most experienced VCs consistently deliver valuable non-financial resources, reinforcing the idea that reputation signals both expertise and access to critical networks. This aligns with findings by Bottazzi et al. (2008), who argue that the size and age of a VC often serve as proxies for its quality and reputation.

While some studies, such as Puri and Zarutskie (2012), fail to find significant differences based on VC reputation, the broader literature overwhelmingly supports the hypothesis that high-reputation VCs provide unique advantages. These advantages stem not only from their financial contributions but also from their ability to add operational, strategic, and network-based value.

Based on this body of evidence, we hypothesize that VC reputation positively impacts the effects of flight introductions on success outcomes.

Hypothesis 2a. VC reputation strengthens the positive effect of the introduction of a new airline route for a VC on the success prospects of its portfolio startup.

Similarly, for travel disruptions we expecte a more negative impact, as portfolio startups may suffer more acutely from the absence of high-quality monitoring provided by reputable VCs.

Hypothesis 2b. VC reputation strengthens the negative effect of the cancellation of a new airline route for a VC on the success prospects of its portfolio startup.

VC Monitoring Tenure

The benefits of VC monitoring tend to operate most strongly in the immediate phase after the initial VC investment. For example, Puri and Zarutskie (2012) find that the reduction in failure rates associated with VC involvement primarily occurs immediately following the initial VC investment and diminishes over time. Similarly, Chemmanur et al. (2011) document efficiency gains that are most pronounced during the first two rounds of VC financing, and Bertoni et al. (2011) attribute the majority of VC treatment effects to the early post-investment phase.

Given this temporal dynamic, we hypothesize that the VC monitoring tenure moderates the impact of flight introductions and cancellations on follow-on funding.[3] VCs and startups that have been involved longer likely experience a reduced effect from disruptions in on-site monitoring, as the most critical phase for hands-on guidance and monitoring may have already passed. Conversely, dyads of VCs and startups in earlier stages, where monitoring is more crucial, may be more significantly affected by changes in travel conditions.

Hypothesis 3a. VC monitoring tenure attenuates the positive impact of flight introductions on success prospects of the portfolio startup.

Hypothesis 3b. VC monitoring tenure attenuates the negative impact of flight cancellations on the success prospects of the portfolio startup.

[3] One could argue that the age of the portfolio startup is a more relevant indicator of its development stage. However, we contend that the date of VC investment provides a more precise benchmark. While some startups may have existed for a period before securing funding, they may not have fully transitioned into operational ventures. VC investment, by contrast, marks a critical inflection point where all funded startups reach a comparable level of formalization and resource availability, making it a more consistent measure of their development stage.

3.3 Data and Variables

3.3.1 Data

Startups and VCs

For our startup database, we merged startups from PitchBook[4] and Crunchbase[5] with at least one equity invest, and end up with a total number of 163,269 startups operating in the years 2006 to 2023. About half of them (81,685) receive VC funding from 32,645 VC investors in their funding history, while the other half does not (81,904). As we compare VC-backed startups with different degrees of VC monitoring, we remove the non-VC-backed startups, which would bias results due to introducing a selection effect into the estimation. Because for our analysis, geographical information is required, we drop all companies and investors without addresses, which leaves us with dyads between 80,296 startups and 29,808 investors. The availability of air traffic data necessitates conditioning the sample on dyads where at least one member is located in the United States. After this restriction, the sample includes 63,703 startups and 24,110 investors, with the majority of dyads consisting of both members located within the United States.[6]

Lead VC

Startups typically receive funding through multiple rounds and from various investors. Our analysis focuses on the active influence exerted by a lead VC, more specifically, the investor with the most significant involvement and the earliest and most frequent contributor. To identify this lead VC, we use a multi-step approach that combines three indicators: (a) a self-calculated measure of the VC that invested earliest and most frequently, and the lead VC information provided by (b) PitchBook and (c) Crunchbase.[7]

First, we filter for VC-startup dyads where the lead VC can be uniquely identified based on all three indicator columns (37,347). Next, we prioritize VCs that invested earliest and were most active; if a unique match is found, we designate that VC as the lead (10,432). Finally, for remaining startups where ambiguity persists, we identify the lead VC as the

[4] `https://pitchbook.com/`

[5] `https://www.crunchbase.com/`

[6] In our data, startups are either from the United States, United Kingdom, Germany, or Canada. Investors can be from any location.

[7] We could only rely on PitchBook and Crunchbase, however, the information does not agree and often reports multiple lead VCs. Therefore, we prefer our multi-step approach.

one that invested earliest, is the most active, and is also recognized as the lead by either PitchBook or Crunchbase (4,270). For 11,654 startups we do not unambiguously identify the lead VC and, consequently, drop them from the sample.

Further, we set the investment date to be later than 1st January 2006 and while VCs might invest in later stages and take on a leading role, we consider only lead VCs that have invested at latest in the Series D funding. Ultimately, we end up with dyads of 49,998 startups and 14,162 lead VCs, with each startup having only one lead VC.[8]

Distances

To study the effect of travel shocks on on-site VC monitoring, we identify dyads of VCs and portfolio startups that depend on air travel by analyzing their geographical distance and the travel time required by car. For actors located within close proximity, travel shocks are unlikely to impact on-site engagement, as car travel would typically be used instead. We set a maximum car travel time of three hours, an admittedly arbitrary threshold, but test the robustness of our results using both smaller and larger values. To calculate distance and travel time, we employ a routing service[9] based on OpenStreetMap data.[10] The average distance between dyads is 1,753 km (standard deviation of 2,269 km) and the travel duration is 1,191 minutes (standard deviation of 1539 minutes). For 2,391 dyads, there is no viable travel option by car. By our definition of three hours of travel, 38.4% of the dyads are too close to be affected by flight routes.

Similarly, we use this service to calculate the car travel distance and duration between VCs or portfolio startups and nearby airports. Airports are considered relevant only if they are within a 45-minute drive of both the VC and the portfolio startup. We include a very comprehensive list of airports, which is why we find on average 3.79 and 4.74 airports for startups and investors, respectively.

[8] We also exclude dyads involving Russian investors that would otherwise qualify for treatment due to the cessation of air traffic between Russia and the United States following the Russian invasion of Ukraine. Since startups in these dyads may have been subject to sanctions due to their ownership structure, isolating the treatment effect would not be feasible, potentially confounding the analysis.

[9] `http://project-osrm.org/`

[10] `https://www.openstreetmap.org/about`

Air Traffic

We use monthly aggregated flight data from the Air Carrier Statistics (Form 41 Traffic) - All Carriers database[11], commonly referred to as the T-100 data bank. Data has been compiled by the Office of Airline Information within the Bureau of Transportation Statistics since 1990. More specifically, we fetch international and domestic segment data which encompasses both international and domestic flight markets, including various parameters such as carrier information, origin, destination and number of passengers. T-100 data is very extensive and reliable due to the legal requirement for airline carriers operating in the United States to submit the Form 41. However, the data only includes flights involving either an origin or destination within the United States. Therefore, we focus our analysis on at least one member of a dyad being located in the United States.

Patents

We obtain data on innovation activity by retrieving patent applications for the period from 2010 to 2022 from PATSTAT[12], a comprehensive database containing patent information from leading industrialized and developing countries. Since neither our startup database nor PATSTAT contains perfectly standardized company names, we perform several steps to match the two datasets effectively.

First, we clean and preprocess all company names from the startup database (including both "name" and "legal name" fields) and all corresponding fields in PATSTAT ("person name," "person name in original language", and names standardized by DOCDB, PATSTAT, and OECD procedures). The preprocessing includes converting strings to lowercase, replacing or removing special characters, and handling whitespaces to create a consistent format across datasets.

Next, we conduct a matching process across all combinations of names. A match is deemed valid if one of the following conditions is met: (i) the match is unique, meaning no patent is assigned to multiple startups, or (ii) the match is unique after further conditioning on city or postal code.

Using this approach, we successfully identify matches for 18.25% of startups in our database within PATSTAT. The vast majority of matches (99.6%) are based on condition

[11] https://www.transtats.bts.gov/DatabaseInfo.asp?QO_VQ=EEE&YvOx=D
[12] https://www.epo.org/en/searching-for-patents/business/patstat

(i).

3.3.2 Sample and Variables

We study both the improvement and deterioration of travel opportunities for VC-startup dyads, which is why we create two samples. With the first sample, the route introduction sample, we exploit an exogeneous positive shock to the ease of travel, while conversely, for the second sample, the route cancellation sample, we exploit a negative shock to the travel possibilities. Due to the different nature of the shocks, but also the correct construction of control groups, we essentially perform two separate albeit related studies which we explain in the following.

Treatment

For each distant dyad and their respective nearby airports, we analyze all direct flight connections and calculate the actual monthly passenger volume, averaged in quarters, traveling from the VC location to the startup location to assess travel possibilities. We assume that routes with very few passengers indicate the absence of a commercial route. Specifically, we classify a route as commercial only when it records more than 120 monthly passengers. We set the analysis period from the first quarter of 2010 to the last quarter of 2023, as we have full data coverage with regards to the outcome variables starting in 2010.

We consider a positive shock when, for a dyad, the volume increases from consistently zero monthly passengers to consistenly more than 120 monthly passengers. To avoid minor disruptions and idiosyncrasies, we define "consistent" as at least six quarters within a two-year window for the treatment and as at least 75% of the periods preceding the treatment and following the initial shock.

For route cancellations, we adopt a similar methodology, identifying a negative shock as a reduction in passenger volume of at least 99% compared to the reference year 2019 for at least six quarters within a two-year window. Again, a flight connection is considered as commercial when it carries more than 120 monthly passengers. We only consider dyads with lead VC investment latest in first quarter of 2020 for the cancellation sample, just before the pandemic unfolded.

It is important to note that we do not directly measure the introduction or cancellation of routes but instead rely on passenger volume. This approach is motivated by two considerations: (i) using passenger volume rather than the number of flights allows us to account for scenarios where flights operated but passenger entry was restricted, and (ii) it ensures comparability between samples.

The primary phase of flight cancellations occurred between 2020 and 2022 at the peak of the Covid-19 pandemic, which is just one year before the final period in our sample. For all periods following the travel shock, the treatment indicator remains active (=1), even if air traffic partially recovers, because we consider potential lagged effects to be both relevant and of interest. Essentially, from an econometrical perspective, we treat both shocks as persistent even though one of them is of temporary nature.

Control Group

We distinguish between several types of dyads that do not experience a shock. These include three primary groups: (1) dyads that were never connected by a non-stop route throughout the entire sample period, referred to as "never-flyers"; (2) dyads that remained continuously connected without any interruptions, referred to as "always-flyers"; and (3) dyads that are not reliant on air travel due to their short geographical distance. We do not select (3) as a control group due to potential risk of differences in characteristics. It might be possible that VCs tend to favor geographically close investments.

For the route introduction analysis, the control group consists of the "never-flyers" (1). We adopt a strict definition, ensuring that the number of passengers for these dyads remains zero throughout the entire observation period. In the cancellation sample, the relevant control group comprises the "always-flyers" (2), To ensure consistency, we apply a stringent criterion that these dyads maintain a minimum of 1,000 monthly passengers throughout the observation period, even during the pandemic.

Thus, for both samples, we adopt conservative definitions for the control groups to account for potential fuzziness in the construction of the treatment groups. This approach minimizes the risk of assigning dyads to the control group that could, under slightly modified treatment assignment rules, qualify as part of the treatment group. It follows that some geographically distant dyads are excluded from both the treatment and control group because they cannot be unambiguously assigned to either group.

While we rely on never-flyers and always-flyers for the introduction and cancellation samples, respectively, it is theoretically possible for either groups to serve as valid control group, provided no other external factors influence success outcomes within these groups and no unobserved prior heterogeneous treatment effects are present.

Outcome

We focus on a set of outcomes that provide a nuanced understanding. First, we measure the scale of a startup's innovation activity, which has been widely gauged based on patent-based measures (see e.g. Bernstein et al., 2016; Bronzini et al., 2020; Kortum and Lerner, 2000; Lerner et al., 2011; Aghion et al., 2013; Seru, 2014). Because the number of patents a startup applies for is very unevenly distributed and often zero, we add one and take the logarithm. Other than Bernstein et al. (2016) who also count the number of citations, we refrain from doing so as the time window to count these citations is too short for our most recent observations.

Second, we measure follow-on funding, a widely recognized metric in entrepreneurial finance for assessing success and growth (Kerr et al., 2014; Kaplan and Schoar, 2005; Hochberg et al., 2014). Follow-on funding is evaluated along two dimensions: at the extensive margin, by counting the number of successful follow-on funding rounds in the periods following the shock, and at the intensive margin, by summing up the total funding amount raised during these rounds and logarithmizing.

However, the raised amounts in a funding round are not always disclosed, often for strategic reasons. Therefore, we defer the analysis of follow-on funding amounts to a robustness check, restricting the sample to startups that (i) consistently report their raised amounts and (ii) have imputable values where appropriate.

Third, we observe the operational outcomes, namely survival and exit rates, which have been studied in other studies as well (see e.g. Bernstein et al., 2016; Hochberg et al., 2007; Sørensen, 2007; Puri and Zarutskie, 2012).[13] In a robustness test, we differentiate between acquisitions and IPOs as two distinct forms of exits.

Fixed Effects and Time-Varying Controls

Following common methodology, we inlude fixed effects at the VC-startup dyad level and fixed effects for quarters. For most specification, we also interact period and startup

[13] Exit includes acquisitions and IPOs. Bernstein et al. (2016) notes that not all acquisistions are positive outcomes and sets an arbitrary cutoff at $25m to distinguish between negative and positive outcomes. We count every acquisition for two reasons: (i) the threshold is fairly arbitrary and (ii) often, the acquisition value is not published.

country to control for time-varying shocks that could have affected specific countries. This way, we account for any unobserved, time-invariant, dyad-specific heterogeneity, which is constant over time (dyad fixed effects), and for any time-varying shocks that affect all dyads (time fixed effects), potentially only in particular countries.

Generally, covariates expected to correlate with startup success (Cinelli et al., 2024) should be included in the model. However, when employing a DiD strategy with dyad fixed effects, time-invariant covariates at the dyad, startup, or investor level automatically drop from the estimation due to their lack of variation within the dyad. In this setting, only time-varying covariates influence the estimates.

We include a time-varying control variable for the round type (Min: Seed - Max: Series I+) a startup is in, as this variable is likely to affect expected success outcomes. In robustness checks, we also include the logarithm of the total raised funding amount up to the specific observation period.[14] In section 3.4.2, we discuss the challenges of including such time-varying variables that may themselves be affected by the treatment.

Specifically for the analysis of the pandemic-induced flight disruptions, we leverage data from the "Oxford Coronavirus Government Response Tracker" (OxCGRT) project (Hale et al., 2021), which collects response metrics across various dimensions and provides a composite measure reflecting the stringency of government responses to the pandemic.

We incorporate time-varying metrics of government response stringency and workplace closures at the startup location into the estimation. The granularity of this data varies by country. For example, in the United States and Canada, the data is available at the state level; in the United Kingdom, it differentiates between England, Northern Ireland, Scotland, and Wales; and in Germany, it is aggregated at the national level.

By including these measures, we aim to control for factors related to government responses to the pandemic that could adversely affect startup operations but are not directly caused by flight interruptions. This approach ensures that the estimated effects of flight disruptions are not confounded by broader pandemic-related policies.

Moderators

VC reputation has been measured with various measures in the literature, with all referring to some kind of previous VC success and the ability to raise funds, a symbol of previous

[14] This control is included only in robustness checks due to the aforementioned issues with incomplete reporting of funding amounts.

success. Measures range from basing on market share of amount of funds raised in a five-year window (Chemmanur et al., 2011), the number of previous investments (Sørensen, 2007) in general or within an industry (Hsu, 2004; Nahata, 2008), to number of successfully exited investments in previous five years (Quas et al., 2021). We follow the literature and measure VC reputation as the number of investments a VC has done within five years leading up to the treatment. VCs higher than the median at that particular year are classified as VCs with high reputation.

VC monitoring tenure is simply measured as the number of years since the investment partnership was initiated at the treatment date.

3.4 Estimation Strategy

3.4.1 Identification

Identification Challenges

To evaluate the validity of an identification strategy, it is helpful to consider the ideal, albeit hypothetical setting. Our objective is to estimate the effect of VC monitoring. In an ideal scenario, we would observe the same VC-backed startups both under active VC monitoring and in the absence of such monitoring. However, due to the fundamental problem of causal inference, it is impossible to observe a startup in both its factual and counterfactual state simultaneously.

The closest approach to addressing the fundamental problem of causal inference, and often referred to as the gold standard, is a randomized experiment. In the context of our study, this would require the randomized assignment of VC monitoring to startups. However, since a VC's primary objective is to maximize returns and exert influence on startups to achieve this goal, VCs would need to be financially incentivized to refrain from monitoring startups assigned to the control group. The monetary compensation required to offset the potential loss from reduced monitoring makes such an experiment practically infeasible.

Thus, in this context, the lack of experimental data forces scholars to rely on observational data and employ more complex identification strategies. The most straightforward approach in prior research has been to compare VC-backed startups to non-VC-backed startups, balance the groups based on observable characteristics, and attribute outcome differences to the causal effect of VC influence (Puri and Zarutskie, 2012; Chemmanur et al., 2011; Quas et al., 2021). However, such strategies face significant limitations. Even with carefully executed matching and balancing procedures, the risk of unobserved factors influencing the outcomes remains. This challenge is particularly pronounced in the opaque startup ecosystem, where limited public information makes it difficult to account for all relevant variables. Consequently, these approaches carry a high risk of producing biased estimates.

The direction of this bias is likely upward. VCs employ rigorous screening processes to identify the most promising startups, and these startups may achieve better outcomes even in the absence of VC backing. Since matching on observable characteristics cannot fully replicate the VCs' screening process, the treatment group (VC-backed startups) would appear to perform better, leading to an overestimation of the true causal effect of VC monitoring.

Identification Approach

We, therefore, adopt an approach that closely simulates a randomized treatment assignment in a natural setting. Specifically, we exploit exogenous shocks caused by route introductions and travel restrictions, which effectively divide the sample into a treatment group (exposed to a travel shock) and a control group (unaffected by changes in travel conditions). These exogenous shocks are unanticipated and independent of startups' potential outcomes, meaning that VCs and startups are unlikely to have foreseen whether they would experience a travel shock. As such, this setup mimics a random treatment assignment, similar to that of a controlled experiment.

The factors driving airlines to introduce new routes between a VC and a portfolio startup are unlikely to be related to the specific activities or characteristics of individual startups. Similarly, the sudden and severe travel restrictions that complicated or entirely prevented travel during the COVID-19 pandemic occurred quasi-randomly, affecting some VC-startup dyads while leaving others unaffected. These shocks, therefore, provide a unique opportunity to causally identify the effects of VC monitoring.

It is important to note, however, that this approach allows us to identify only the on-site monitoring effect. With the increasing adoption and continuous development of communication technologies, remote collaboration has become more feasible, potentially substituting parts of traditional in-person monitoring. Thus, our estimates capture only a part of the overall monitoring effect. Nevertheless, by investigating shocks to on-site VC monitoring, we can causally relate VC monitoring to different outcomes and different contexts in order to shed light on important facets of heterogeneous effects.This methodology aligns closely with the approach employed by Bernstein et al. (2016), who similarly leverage exogenous variation in travel conditions to study VC involvement.

3.4.2 Estimation

Estimation Challenges

We employ a difference-in-differences (DiD) approach, a widely used method in policy evaluation that estimates treatment effects by comparing changes in outcomes over time between treatment and control groups. The key assumption for unbiased estimation is not the perfect similarity of groups but rather that, in the absence of treatment, the difference in average outcomes between the two groups would have remained constant over time. In the canonical two-period, two-group setting, this parallel trends assumption is expressed as

$$\mathbb{E}[Y_{t=1}(0) - Y_{t=0}(0)|D = 1] = \mathbb{E}[Y_{t=1}(0) - Y_{t=0}(0)|D = 0] \tag{3.1}$$

$$\mathbb{E}[\Delta Y(0)|D = 1] = \mathbb{E}[\Delta Y(0)|D = 0], \tag{3.2}$$

where $D = 1$ indicates units in the treatment group, and $D = 0$ represents the control group. This assumption generalizes to settings with multiple time periods and treatment cohorts as follows:

$$\mathbb{E}[\Delta Y_{t,t'}(\infty)|G_i = g] = \mathbb{E}[\Delta Y_{t,t'}(\infty)|G_i = g'] \quad \forall t \neq t' \text{ and}, g \neq g'. \tag{3.3}$$

where $Y_{it}(\infty)$ denotes the potential untreated outcome for unit i at time t, t and t' are any two time periods, G_i indicates the treatment cohort for unit i, and g and g' represent distinct treatment cohorts.

To enhance the plausibility of this untestable parallel trends assumption in observational settings, researchers often include covariates, relaxing the assumption toward a conditional parallel trends framework.

The two-way fixed effects (TWFE) estimator, historically the standard approach for such settings, is specified as:

$$Y_{it} = \alpha_i + \lambda_t + \tau_{\text{TWFE}} D_{it} + \beta X_{it} + \varepsilon_{it}, \tag{3.4}$$

where α_i represents unit fixed effects, λ_t captures time fixed effects, D_{it} is the treatment indicator (equal to 1 in post-treatment periods for treated units), X_{it} are time-varying covariates, and ε_{it} is the error term. Because unit fixed effects absorb time-invariant covariates Z_i, only time-varying covariates X_{it} can be included to strengthen the plausibility of the parallel trends assumption. We interpret the estimated coefficient for τ_{TWFE} as the average treatment effect on the treated ATT.

However, while estimation with TWFE is valid in both two-period and multiple-period settings, challenges arise when there is variation in treatment timing and treatment effects are heterogeneous (see e.g., de Chaisemartin and D'Haultfœuille, 2023; Goodman-Bacon, 2021; Borusyak et al., 2024). Moreover, the inclusion of time-varying covariates poses considerable risk of bias, when it itself affected by the treatment (Caetano et al., 2024; Caetano and Callaway, 2024). We discuss these issues in the following and present solutions.

(i) Variation in treatment timing. The core issue with TWFE estimation arises from the comparisons it implicitly makes across groups with different treatment timings. The ATT estimate combines valid and invalid comparisons. While it is valid to compare treated groups to never-treated groups or to groups that have not yet been treated, TWFE also compares treated groups to already-treated groups. Once a group has been treated, there is no longer variation in the treatment indicator for that group, making these comparisons problematic when treatment effects are heterogeneous.

The validity of such comparisons hinges on the assumption of homogeneous treatment effects, where the treatment effect is constant over time (e.g., it shifts at the treatment date but neither grows nor recovers afterward). However, this assumption is highly restrictive and rarely holds in practice. In most real-world settings, treatment effects evolve over time. For example, in our context, the effect of on-site VC monitoring might be larger in the initial years following a route introduction but could diminish in subsequent years. This heterogeneity in treatment effects would lead to biased estimates if TWFE is used without accounting for such dynamics.

(ii) Time-varying covariates. A second issue arises from the inclusion of time-varying covariates in the estimation in order to strengthen the parallel trends assumption, effectively resulting in the conditional parallel trends assumption. For simplicity, we refer to the canonical two-period, two-group setting, expressed as

$$\mathbb{E}[\Delta Y_t(0)|Z, X_t(0), X_{t-1}, D = 1] = \mathbb{E}[\Delta Y_t(0)|Z, X_t(0), X_{t-1}, D = 0]. \tag{3.5}$$

Unit fixed effects absorb time-invariant characteristics Z, such as industry or country, which are therefore automatically excluded from the estimation due to the lack of variation within a startup-VC dyad over time. As a result, any observed or unobserved heterogeneity that remains constant across time for a given dyad is effectively controlled for. Similarly, time fixed effects account for common shocks or trends affecting all units simultaneously, ensuring that time-varying factors shared across groups are also captured. Consequently, variables only affect the estimation when they are time-varying or have a time-varying effect on the outcome (Roth et al., 2023).

In our setting, we identify two critical time-varying characteristics that are likely to affect various outcomes: the stage a company is in and, when available, the funding it has raised. Ignoring these variables or fixing their values at a pre-treatment period could bias the estimates. Fixing covariate values at pre-treatment periods is a common approach but comes with inherent pitfalls. Specifically, for never-treated units in the multi-period setting, there is no meaningful pre-treatment period by definition, which undermines the validity of this approach.

If the covariates are unaffected by the treatment, fixing them at pre-treatment values is valid. However, in our case, we expect the treatment to influence funding-related outcomes, which implies that the funding covariates themselves may also be affected. Fixing these covariates at pre-treatment values would artificially halt the funding trajectory only for treated units, leading to incorrect estimates of the treatment effect.

On the other hand, including time-varying covariates without caution can introduce "bad controls" (Cinelli et al., 2024) or post-treatment bias (Montgomery et al., 2018; Acharya et al., 2016). This occurs when covariates are influenced by the treatment, as their inclusion can absorb part of the treatment effect, leading to distorted estimates. Thus, careful consideration is necessary to balance the need for controlling relevant covariates while avoiding biases caused by post-treatment variables.

However, $X_t(0)$ is unobserved for the treatment group with $D = 1$. Since these time-varying covariates are likely affected by the treatment, the condition $\mathbb{E}[X_t(0) \mid D = 1] = \mathbb{E}[X_t(1) \mid D = 1]$ is unlikely to hold. Therefore, we must find a method to impute $X_t(0)$.

Estimation Approach

With our estimation approach, we address the issues associated with variations in treatment timing and time-varying covariates that are affected by the treatment. Particularly, the challenges associated with variation in treatment timing have recently gained significant attention in the literature (Goodman-Bacon, 2021; Borusyak et al., 2024; de Chaisemartin and D'Haultfœuille, 2023; Sun and Abraham, 2020; Wooldridge, 2021; Callaway and Sant'Anna, 2021), leading to the development of several robust methods. We adopt the estimator proposed by Gardner (2022) and implemented in R by Butts and Gardner (2022).

At its core, estimating a valid treatment effect involves approximating the untreated, unobserved outcome, commonly referred to as the counterfactual outcome. The proposed estimator achieves this through a two-stage procedure. In the first stage, outcomes for untreated observations are regressed on fixed effects and covariates to impute the counterfactual outcome for treated units. In the second stage, the observed outcomes are residualized by subtracting the estimated counterfactual outcome. The residualized outcome thus isolates the portion of the variation attributable to the treatment, enabling a valid estimation of the treatment effect.

First stage. Using the subsample of untreated observations $D_{itr} = 0$, we estimate a first stage that regresses the outcome $Y(0)$ on fixed effects for dyads α_i and country-periods λ_{tr} of country r at period t, and a vector of time-varying covariates X_{it}. ϵ_{itr} denotes the error term.

$$Y_{itr}(0) = \alpha_i + \gamma_{tr} + \beta X_{it} + \varepsilon_{it} \tag{3.6}$$

Second stage. Using the fitted model, we impute $\widehat{Y}_{itr}(0)$, which is the counterfactual outcome for the treated units $D_{itr} = 1$, and subtract it from the observed outcomes. We obtain a residualized outcome $\tilde{Y}_{itr} = Y_{itr} - \widehat{Y}_{itr}(0)$, which contains what cannot be explained by the first stage model. Regressing the adjusted outcome on the treatment indicator yields the ATT. Treatment effect heterogeneity can be analyzed by adding interaction effects.

$$\tilde{Y}_{itr} = \tau D_{it} + \epsilon_{itr} \tag{3.7}$$

This approach relies on the following assumptions to be satisfied.

(1) (Conditional) parallel trends assumption: in absence of treatment, treated startups follow a similar trend as untreated startups (conditional on covariates).

(2) No anticipation: the failure of a startup is not anticipated prior to the failure.

(3) Correct specification of the first stage: the parametric model for $Y(0)$ needs to be correctly specified to impute the counterfactual outcomes.

We can assess the plausibility of our assumptions by examining pre-treatment periods for zero treatment effects (1). This helps validate the absence of confounding trends prior to treatment. However, the assumption itself remains fundamentally untestable. Additionally, by analyzing the periods leading up to the treatment, we can evaluate whether the treatment was anticipated. In our case, this seems unlikely due to the exogeneity of the shock, particularly in the context of flight disruptions. For assumption (3), we are confident in the correct specification of the model, which includes dyad fixed effects, country-time fixed effects, and additional controls for time-varying variables, such as the rounds (and funding) a startup has completed. These elements are designed to appropriately capture the dynamics of the counterfactual outcome and enhance the robustness of our estimation.

To address the second issue, which has been recognized for quite some time (Montgomery et al., 2018; Acharya et al., 2016), but only recently has been directly linked to the

DiD setting (Caetano et al., 2024; Caetano and Callaway, 2024), our implementation shares similarities with the imputation approach for the untreated potential outcome. However, in this case, we impute the untreated and unobserved covariate instead.

Given that time-varying covariates are suspected to be influenced by the treatment and are also expected to support the plausibility of parallel trends for the outcome variable, we cannot fix the covariate values at a pre-treatment date or use the observed covariate values directly.

To address this, we adapt the approach suggested by Caetano et al. (2024). Specifically, we first fit a regression of the covariate on lagged covariates $X_{i,t-1,r}$ and lagged outcomes $Y_{i,t-1,r}$ using untreated observations. This allows us to model the expected values of the covariate in the absence of treatment. Unit and country-period fixed effects are included, as well.

$$X_{itr}(0) = \mu_i + \eta_{tr} + \rho X_{i,t-1,r} + \phi Y_{i,t-1,r} + \nu_{itr} \tag{3.8}$$

Then, we estimate the counterfactual untreated covariate for the treated observations and impute these values for our estimation. By obtaining the estimated covariate values as they would be in the absence of treatment, we can proceed to estimate the potential outcome in the absence of treatment, as described in the previous paragraph. This two-step process ensures that both the covariate and the outcome reflect the counterfactual scenario, addressing potential biases introduced by time-varying covariates affected by the treatment.[15]

3.5 Results

3.5.1 Startups, VCs and Dyads

Table 3.1 and Table 3.2 present the summary statistics for the introduction and cancellation samples, respectively. As we are rather strict about the assignment to the treatment groups and due to the nature of the shocks, they are considerably smaller than the control groups.

Startups in treated dyads generally exhibit similar characteristics to their counterparts

[15] To model the funding round evolution in the untreated state, we predict follow-on funding and increase the round level by 1 if follow-on funding $> .5$

TABLE 3.1: Descriptive Statistics - Flights Introduction

	Control				Treatment			
	Pre (N=111,814)		Post (N=0)		Pre (N=6,218)		Post (N=4,067)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
DV								
Log. Number of Patents	0.03	0.23	-	-	0.04	0.29	0.03	0.18
Number of Patents	0.10	1.03	-	-	0.20	2.55	0.06	0.63
Follow-on Funding (x100%)	11.76	32.21	-	-	15.13	35.84	6.44	24.55
Log. Follow-on Funding Amt.	1.45	4.33	-	-	2.05	5.08	0.75	3.34
Follow-on Funding Amt.	0.76	7.96	-	-	1.27	11.70	0.77	8.85
Survival (x100%)	99.77	4.76	-	-	99.68	5.66	99.85	3.84
Exit (x100%)	0.20	4.52	-	-	0.51	7.16	0.61	7.82
Acquisition (x100%)	0.18	4.27	-	-	0.48	6.93	0.59	7.66
IPO (x100%)	0.02	1.50	-	-	0.03	1.79	0.02	1.57
Covariates								
Log. Raised Amt.	12.02	0.98	-	-	11.76	1.15	12.04	0.90
Raised Amt. ($M)	0.22	0.14	-	-	0.20	0.15	0.23	0.16
Round Type	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Seed	43,508	38.9	-	-	2,104	33.8	988	24.3
Series A	23,283	20.8	-	-	1,511	24.3	769	18.9
Series B	15,440	13.8	-	-	987	15.9	642	15.8
Series C	10,292	9.2	-	-	548	8.8	563	13.8
Series D	6,507	5.8	-	-	367	5.9	391	9.6
Series E	3,965	3.5	-	-	208	3.3	274	6.7
Series F	2,140	1.9	-	-	110	1.8	157	3.9
Series G	1,275	1.1	-	-	47	0.8	122	3.0
Series H	865	0.8	-	-	50	0.8	61	1.5
Series I+	1,249	1.1	-	-	73	1.2	100	2.5
Startup								
Num. of Founders	2.68	4.92	-	-	2.81	1.95	3.14	2.41
Num. of Female	0.40	1.23	-	-	0.36	0.82	0.37	0.73
Num. of Bachelor	1.85	3.94	-	-	2.11	2.12	2.45	2.40
Num. of Master	1.47	4.51	-	-	1.51	2.14	1.88	3.33
Num. of PhD	0.52	1.28	-	-	0.63	1.22	0.79	1.37
Num. of Experienced	0.46	1.65	-	-	0.30	0.85	0.40	1.03
Country	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
United States	91,762	82.1	-	-	5,578	89.7	3,714	91.3
United Kingdom	13,342	11.9	-	-	283	4.6	124	3.0
Canada	3,765	3.4	-	-	303	4.9	152	3.7
Germany	2,945	2.6	-	-	54	0.9	77	1.9
Dyad								
Distance (km)	3,313	2,850	-	-	2,493	1,495	3,091	1,915
Duration (minutes)	2,277	1,972	-	-	1,703	999.2	2,100	1,294
Startup Age at Lead VC Investment	2.47	2.53	-	-	2.02	2.21	2.00	2.31
Lead VC Investment Round Type	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Seed	72,442	64.8	-	-	4,277	68.8	3,009	74.0
Series A	20,943	18.7	-	-	1,205	19.4	572	14.1
Series B	10,019	9.0	-	-	369	5.9	302	7.4
Series C	5,323	4.8	-	-	266	4.3	143	3.5
Series D	3,087	2.8	-	-	101	1.6	41	1.0
Countries	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
United States	38,793	34.7	-	-	2,956	47.5	1,660	40.8
United States and Other	73,021	65.3	-	-	3,262	52.5	2,407	59.2

Notes. This table shows summary statistics for treated and control group in the introduction sample. For continous and discrete variables, we show mean and standard deviation and for categorical variables, we show counts and frequencies. When we use logarithm values (+1), we also present the level in italics. Amount (Amt.) is in millions and startup age in years.

TABLE 3.2: Descriptive Statistics - Flights Cancellation

	Control				Treatment			
	Pre (N=173,870)		Post (N=0)		Pre (N=11,624)		Post (N=7,559)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
DV								
Log. Number of Patents	0.05	0.29	-	-	0.06	0.29	0.02	0.17
Number of Patents	0.16	1.56	-	-	0.15	1.08	0.05	0.52
Follow-on Funding (x100%)	11.44	31.83	-	-	14.11	34.81	5.45	22.70
Log. Follow-on Funding Amount	1.52	4.55	-	-	1.83	4.88	0.69	3.27
Follow-on Funding Amount ($M)	1.50	15.17	-	-	1.34	10.71	1.63	16.83
Survival (x100%)	99.69	5.53	-	-	99.73	5.16	99.74	5.14
Exit (x100%)	0.34	5.85	-	-	0.28	5.32	0.24	4.87
Acquisition (x100%)	0.32	5.63	-	-	0.25	4.99	0.20	4.45
IPO (x100%)	0.03	1.61	-	-	0.03	1.85	0.04	1.99
Covariates								
Log. Raised Amt.	11.93	1.01	-	-	11.86	0.94	12.06	0.66
Raised Amt. ($M)	0.22	0.15	-	-	0.20	0.14	0.21	0.11
Covid-19 Measures Stringency	13.43	22.49	-	-	1.76	7.96	32.92	26.36
Workplace Closures	0.33	0.69	-	-	0.04	0.22	0.86	0.95
Round Type	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Seed	56,761	32.6	-	-	4,130	35.5	1,836	24.3
Series A	36,564	21.0	-	-	2,841	24.4	1,634	21.6
Series B	26,695	15.4	-	-	1,708	14.7	1,187	15.7
Series C	18,519	10.7	-	-	1,165	10.0	947	12.5
Series D	12,571	7.2	-	-	613	5.3	763	10.1
Series E	7,676	4.4	-	-	367	3.2	447	5.9
Series F	4,752	2.7	-	-	192	1.7	267	3.5
Series G	2,917	1.7	-	-	134	1.2	190	2.5
Series H	1,537	0.9	-	-	41	0.4	91	1.2
Series I+	2,128	1.2	-	-	54	0.5	197	2.6
Startup								
Number of Founders	2.71	1.72	-	-	2.76	1.97	2.68	1.83
Number of Female	0.36	0.72	-	-	0.39	0.78	0.43	0.79
Number of Bachelor	1.98	1.95	-	-	2.05	2.23	1.84	1.88
Number of Master	1.52	1.96	-	-	1.56	2.08	1.44	1.92
Number of PhD	0.56	1.26	-	-	0.56	1.21	0.51	1.09
Number of Experienced	0.50	1.80	-	-	0.48	1.97	0.50	1.64
Country	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
United States	169,975	97.8	-	-	9,117	78.4	5,620	74.3
United Kingdom	2,397	1.4	-	-	987	8.5	947	12.5
Canada	1,407	0.8	-	-	961	8.3	635	8.4
Germany	91	0.1	-	-	559	4.8	357	4.7
Dyad								
Distance (km)	2,521	1,759	-	-	4,838	4,275	5,108	4,306
Duration (minutes)	1,692	1,148	-	-	3,341	2,974	3,526	2,999
Startup Age at Lead VC Investment	2.11	2.22	-	-	1.91	2.07	2.33	2.37
Lead VC Round Type	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Seed	122,398	70.4	-	-	7,772	66.9	4,602	60.9
Series A	29,001	16.7	-	-	2,235	19.2	1,737	23.0
Series B	12,339	7.1	-	-	846	7.3	530	7.0
Series C	6,379	3.7	-	-	492	4.2	453	6.0
Series D	3,753	2.2	-	-	279	2.4	237	3.1
Countries	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
United States	162,150	93.3	-	-	2,993	25.7	1,599	21.2
United States and Other	11,720	6.7	-	-	8,631	74.3	5,960	78.8

Notes. This table shows summary statistics for treated and control group in the cancellation sample. For continous and discrete variables, we show mean and standard deviation and for categorical variables, we show counts and frequencies. When we use logarithm values (+1), we also present the level in italics. Amount (Amt.) is in millions and startup age in years.

in the control dyads. Founder team composition is notably consistent across groups, with about 2.7 – 2.9 founders per team, a tendency to be male-dominated, and educational backgrounds that include more than two thirds of the team with a bachelor's degrees, more than half with a master's degree, and, on average, in half of the startups, there is a founder with at least one PhD. However, founder teams in the control group tend to be slightly more experienced, with the number experienced founders being .46 versus .3 in the introduction sample and .5 versus .48 in the cancellation sample.

There is a minor deviation in terms of raised funding, with treated startups receiving more funding than their control counterparts. However, due to the mentioned publishing issues of funding amounts, it has to be noted that it is based only on a subsample of dyads and therefore, we use this metric only for robustness tests. Across both samples, we observe that treatment dyads tend to be at slightly more advanced funding stages.

Treated dyads in the introduction sample are slightly less geographically distant than control dyads, partly due to the exclusion of close dyads in the initial selection process. In contrast, treated dyads in the cancellation sample tend to be more geographically distant than their control counterparts, indicating that longer routes were more frequently interrupted than shorter ones. Across all samples and groups, key characteristics such as the age of the startup and stage at the time of investment are broadly comparable.

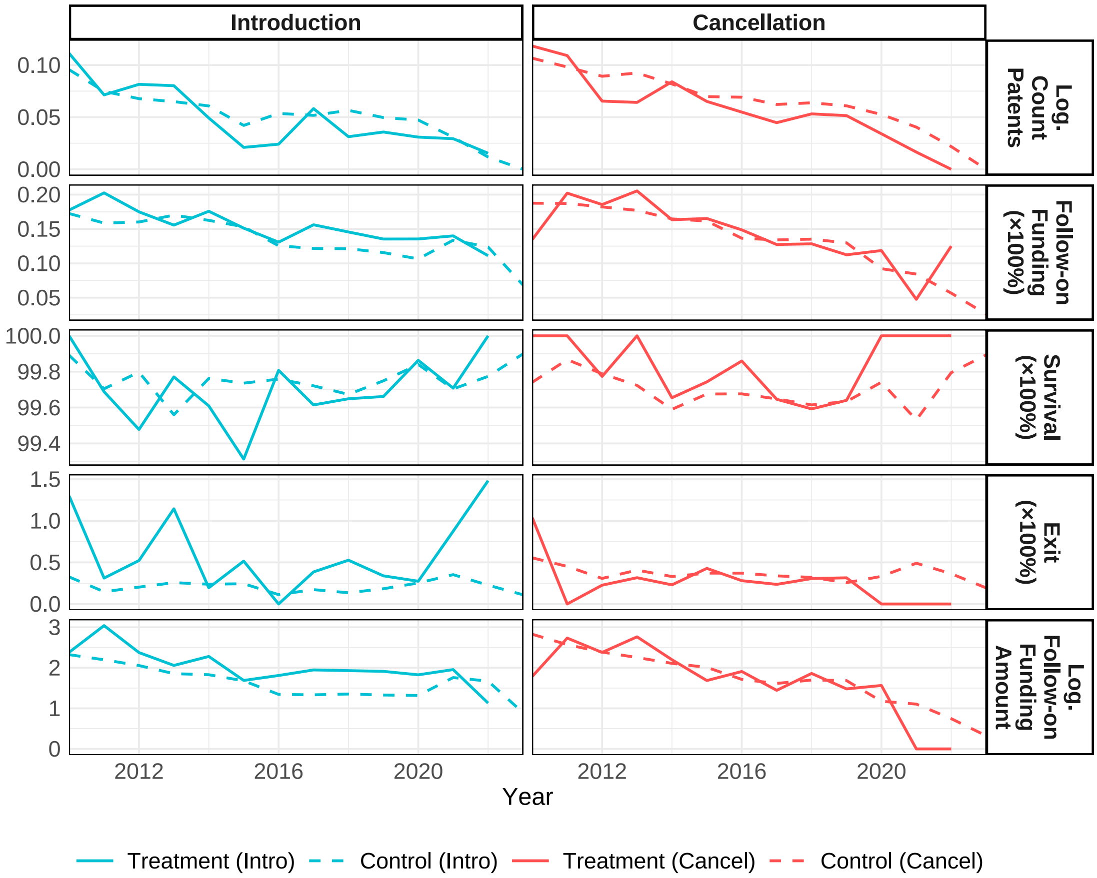
It is important to emphasize that our estimation strategy does not rely on the treated and control groups being perfectly matched in their characteristics. Instead, it hinges on the conditional parallel trends assumption and our ability to model the untreated outcome, as described in the previous section. Specifically, we assume that, in the absence of treatment, the groups would evolve similarly over time, conditional on time-varying covariates and multi-dimensional fixed effects. Nonetheless, the comparability of the groups enhances the credibility of our identification strategy.

Moreover, in Figure 3.1, we see that there is no systematic deviation in the dependent variables over time. However there is some variation over time in the levels for both groups, which we will control for with time fixed effects.

3.5.2 Traffic and Treatment Adoption

Figure 3.2 illustrates the locations of startups and VCs alongside air traffic data for 2019 and 2020, measured as the passenger volume traveling from the VC's location to the

FIGURE 3.1: Evolution of Outcome Variable on Untreated Dyads

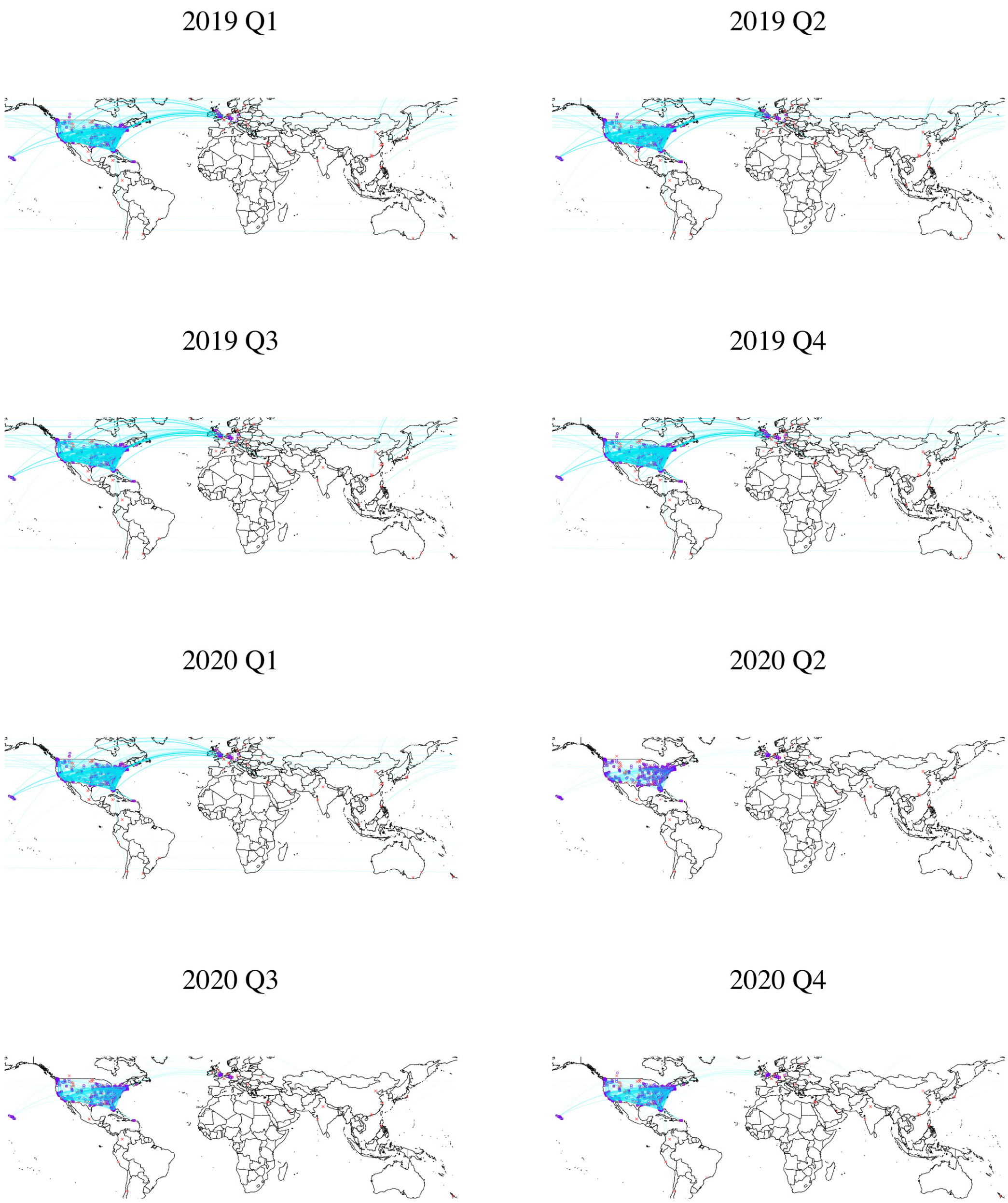


Notes. This figure compares the outcomes over time for not-yet-treated dyads (solid lines) and never-treated dyads (dashed lines). The introduction sample is depicted on the left (blue), and the cancellation sample on the right (red). It is important to note that changes after first quarter of 2020 in the cancellation not-yet-treated group are based on a limited number of observations, as the majority of dyads were already treated by this point. The outcome variable, *Log. Follow-on Funding Amount*, includes only dyads that consistently reported raised amounts.

startup's location. This direction reflects the nature of VC on-site monitoring. Between the first quarter of 2019 and the first quarter of 2020, air traffic shows little variation. However, as the COVID-19 pandemic began to escalate, passenger volumes experienced a drastic decline, globally but also within the United States were most VCs and startups are located. Although air traffic did not cease entirely, the number of passengers and active routes decreased substantially, with the most significant drop occurring in the second quarter of 2020. The figures support the credibility of our identification strategy by illustrating that some VC-startup dyads are abruptly disconnected, while others remain unaffected.

Figure 3.3 provides insights into the differential timing of treatment events. Almost all route interruptions occurred in the second quarter of 2020, with only a few starting one quarter earlier or later. In contrast, route introductions were more evenly distributed

FIGURE 3.2: Passenger Volume between VCs and Startups from 2019 to 2020



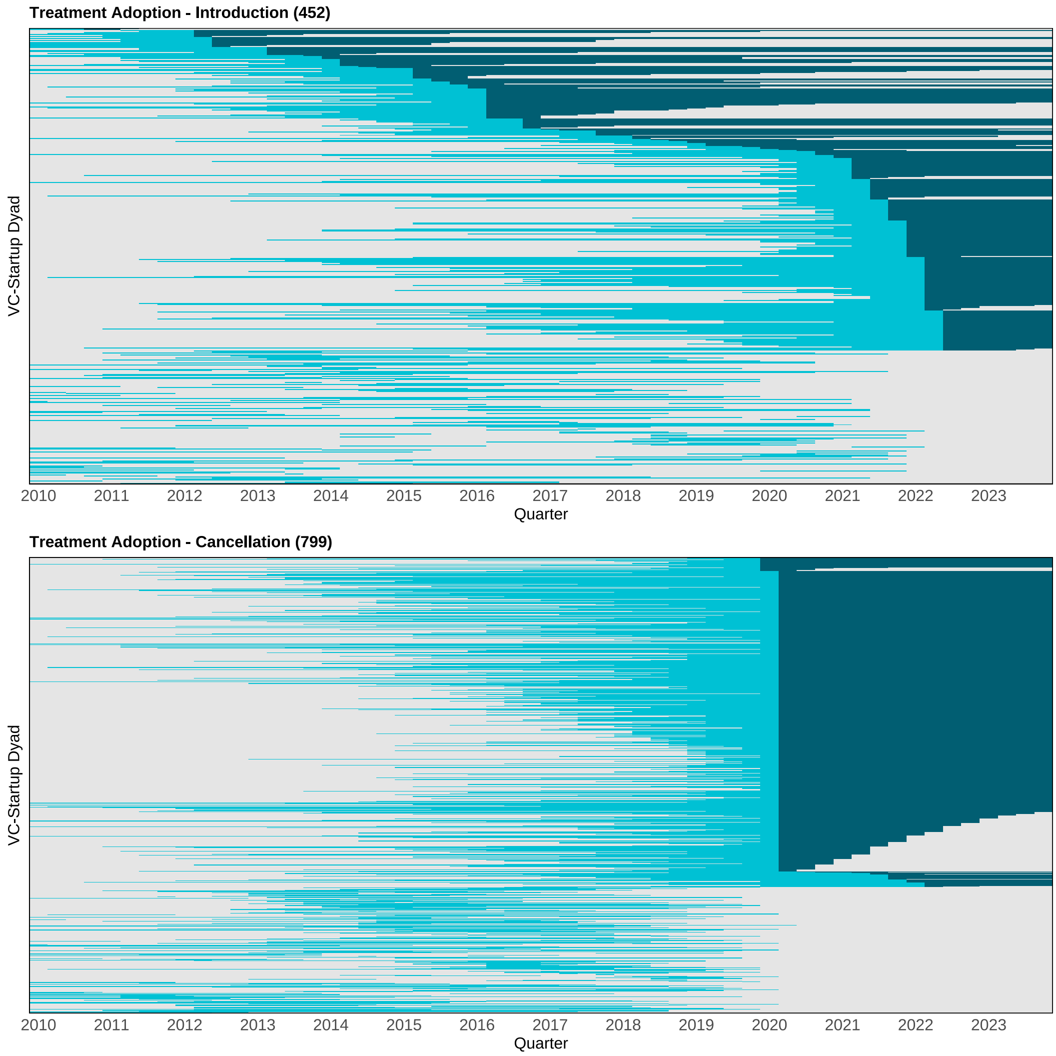
Notes. This figure shows actual passenger volumes from lead VCs to startups and covers the periods of treatment during Covid-19. Each route between airports is only printed once, even if it is used for several dyads. Opacity of line indicates the passenger volume. Startup locations are shown in purple and VC locations in red.

throughout the analysis period, with a slight concentration between 2015 and 2019 and after 2021.[16] When startups are sold or closed, they exit the sample, and the earliest observed period for each startup corresponds to its period of lead VC investment or the

[16]As previously described, a dyad is only conisdered treated when, for the whole lifespan of the dyad, there was no traffic before the treatment.

start of our analysis period.

FIGURE 3.3: Staggered Treatment Adoption



Notes. This figure shows the treatment adoption of treated VC-startup dyads for both types of analysis, flight introductions and cancellations. A light blue bar indicates observations before treatment, a dark blue bar observations after treatment. Periods in grey are not observed, e.g. because the VC was not invested in the startup or the startup was closed or sold. The total number of treated dyads is shown in parentheses. Units that we do not observe in the post-treatment period were sold or closed before.

It can be seen that some dyads, due to exit or closure, do not actually experience the shock for the minimum amount of quarters or not at all. Thus, prior to the treatment, we stop observing these dyads. However, we choose to include these dyads in the analysis for the following reasons: (i) ignoring that these dyads would have experienced the treatment and assigning them to the control group leads to higher closure and exit rates in the control group, biasing the estimated treatment coefficients on those outcomes; (ii) removing all closed or sold dyads would make the estimation of exit and survival rates impossible and leave out valuable information; (iii) assigning these dyads neither to treatment nor control group would still keep the closure and exit rates too low compared to the control group. Therefore, we decide to assign the dyads, which would have experienced treatment to the treated group. Of course, they only enter the sample with pre-treatment periods.

3.5.3 Main Results

Table 3.3, Figure 3.4, and Figure 3.5 present the results of testing H1a and H1b, which hypothesize a positive impact of route introductions and a negative impact of route cancellations on various success measures of portfolio startups due to shocks in on-site VC monitoring.

Innovation activity. Following flight introductions, we observe a moderate increase of approximately 1.5% in patent applications for treated dyads, an effect that is not mirrored by corresponding negative effects after flight cancellations.[17] The dynamic treatment effects indicate that the impact on patent applications becomes more pronounced after a few years. For flight cancellations, however, the limited time horizon due to the data availability restricts our ability to examine effects beyond three years.

Follow-on funding. For follow-on funding at the extensive margin, the effects align with the expected direction only for the cancellation sample. In response to flight cancellations, the quarterly probability of follow-on funding decreases by approximately 3.2%pt (percentage points), equivalent to about 7.5% of a standard deviation, as illustrated in Figure 3.5, and exhibits a monotonically decreasing negative trend over time, as shown in Figure 3.4. In the specifications not controlling for startup stage, the effect is twice as large.

For the introduction sample, we also observe negative effects, though only about half the magnitude. Statistical significance emerges only in the specification that controls for startup stage. Interestingly, follow-on funding shows a short-term increase immediately following the treatment, but then declines steadily in subsequent periods.

[17] In this section, we primarily focus on the column (c) specification for clarity and readability. However, we explicitly highlight instances where the results differ significantly across specifications.

TABLE 3.3: Testing of Main Hypothesis

Panel A: Flight Introduction

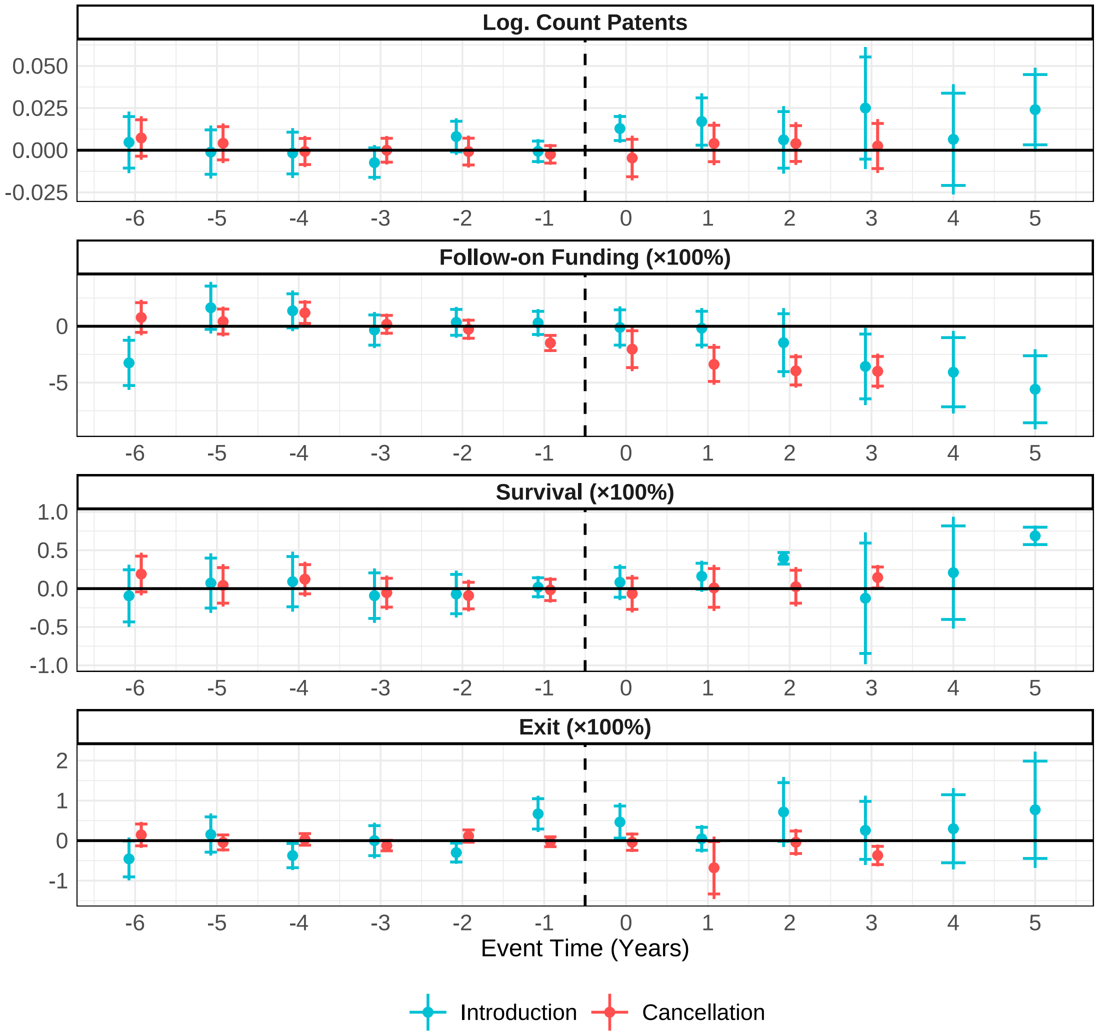
	Log. Count Patents			Follow-on Funding (×100%)			Survival (×100%)			Exit (×100%)		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)	(3a)	(3b)	(3c)	(4a)	(4b)	(4c)
Post X Treat	0.015 *	0.014 *	0.015 *	-1.35	-1.57	-1.64 +	0.25 ***	0.25 ***	0.26 ***	0.27 *	0.26 *	0.27 *
	(0.006)	(0.006)	(0.006)	(1.20)	(1.21)	(0.86)	(0.07)	(0.07)	(0.07)	(0.13)	(0.13)	(0.13)
Num. Obs.	111,467	111,467	111,467	122,099	122,099	122,099	122,099	122,099	122,099	122,099	122,099	122,099
Mean DV	0.035	0.035	0.035	11.8	11.8	11.8	99.8	99.8	99.8	0.234	0.234	0.234
Num. Dyads	6,860	6,860	6,860	7,071	7,071	7,071	7,071	7,071	7,071	7,071	7,071	7,071
Num. Treated Dyads	452	452	452	452	452	452	452	452	452	452	452	452
Dyad FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No	No
Country x Period FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Round Type	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Panel B: Flight Cancellation

	Log. Count Patents			Follow-on Funding (×100%)			Survival (×100%)			Exit (×100%)		
	(5a)	(5b)	(5c)	(6a)	(6b)	(6c)	(7a)	(7b)	(7c)	(8a)	(8b)	(8c)
Post X Treat	0.002	-0.002	0.001	-6.27 ***	-6.21 ***	-3.22 ***	0.05	-0.00	0.01	-0.21 **	-0.31 *	-0.27 *
	(0.006)	(0.006)	(0.006)	(1.03)	(1.18)	(0.70)	(0.07)	(0.07)	(0.07)	(0.07)	(0.15)	(0.14)
Num. Obs.	184,196	184,196	183,892	193,053	193,053	192,705	193,053	193,053	192,705	193,053	193,053	192,705
Mean DV	0.0534	0.0534	0.0534	11.4	11.4	11.4	99.7	99.7	99.7	0.336	0.336	0.336
Num. Dyads	8,305	8,305	8,305	8,305	8,305	8,305	8,305	8,305	8,305	8,305	8,305	8,305
Num. Treated Dyads	799	799	799	799	799	799	799	799	799	799	799	799
Dyad FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No	No
Country x Period FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Round Type	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Covid Measures	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Notes. Robust standard errors, clustered at the dyad level, are shown in parentheses. Patent data is only available up to the spring 2023, resulting in the exclusion of periods after this date from the analysis and a lower number of observations. Stars indicate significance level: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

FIGURE 3.4: Yearly Averaged Dynamic Treatment Effects

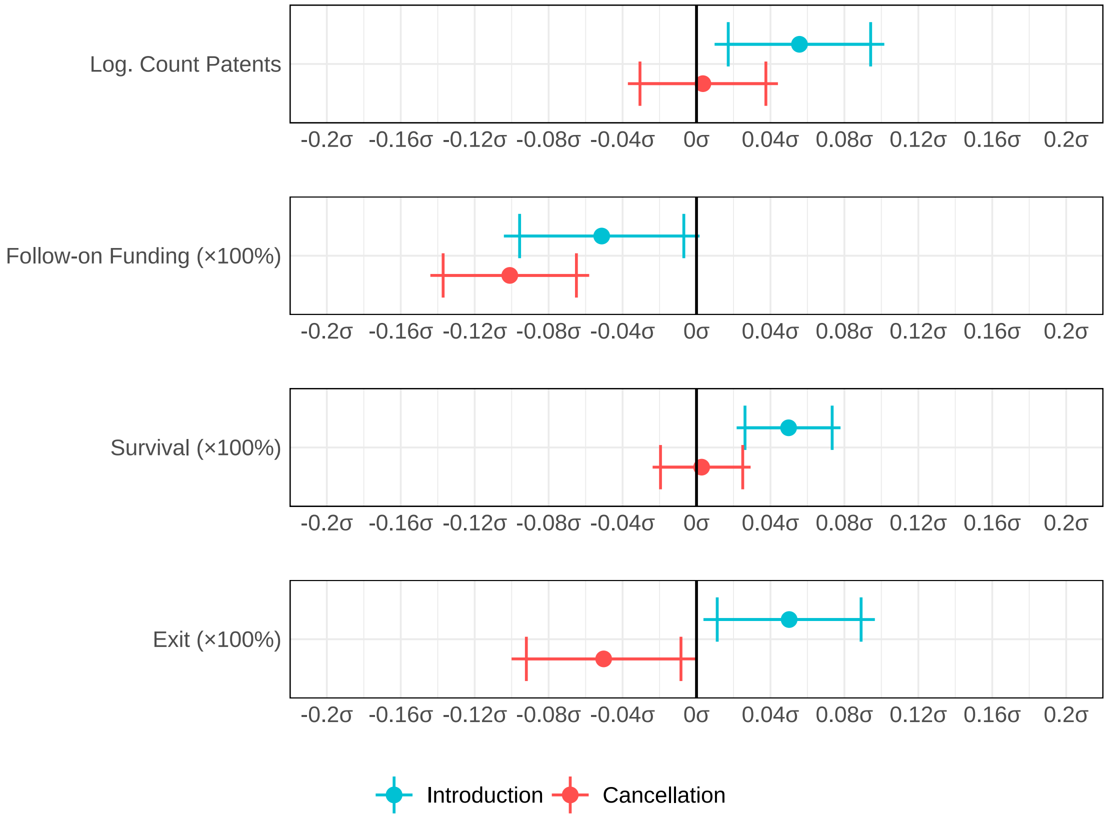


Notes. This figures shows yearly averaged event study estimates of flight introductions and cancellations on various outcomes. Six years prior and sixx years since treatment time are shown with 90% (line) and 95% (bar) confidence intervals. Coefficients are estimated using specification (c), controlling for dyad- and country-period-shocks and the stage a portfolio startup is in.

Operational outcomes. Startups in dyads benefiting from easier travel experience an increase in the survival rate of approximately 0.26%pt. Conversely, for startups in dyads affected by travel disruptions, we find no statistically significant change. However, the effect on survival rates following flight introductions becomes significant only after four years, a time horizon unavailable for analysis in the case of flight disruptions.

For exit rates, we observe an increasing effect following flight introductions and a decreasing effect following flight cancellations. The effects are of same magnitude, with quarterly changes in exit rates of 0.27%pt and −0.27%pt for the introduction and cancellation

FIGURE 3.5: Comparison of Treatment Effects for Introduction and Cancellation



Notes. The graphs puts the treatment effects from flight introductions (blue) in comparison to treatment effects from flight cancellations (red). For comparability across outcomes, the x-axis represents standard deviations for each outcome, measured across the whole sample. 90% and 95% confidence intervals are shown with errorbars and lines, respectively.

samples in the most robust specification, respectively.

These findings provide partial support for H1a, indicating that route introductions have a more pronounced impact on innovation activity, survival rates, and exit outcomes, while their effect on follow-on funding is limited. For H1b, we also find some support, particularly in the observed negative effects on follow-on funding rounds and exit rates. However, the lack of consistent contrary effects across all outcomes suggests that the mechanisms at play differ by the context of the shock, which we explore further in the discussion part of this study.

3.5.4 Effect Heterogeneity

In addition to analyzing the main effects, we extend our predictions to examine moderating effects. The results of the interaction analyses are presented in Table 3.4.

TABLE 3.4: Moderation of Treatment Effects of Travel Shocks

Panel A: Flight Introduction

	Log. Count Patents			Follow-on Funding (×100%)			Survival (×100%)			Exit (×100%)		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)	(3a)	(3b)	(3c)	(4a)	(4b)	(4c)
Post X Treat	0.011 +	0.019 +	0.017	-2.18	-5.13 ***	-8.60 ***	0.34 ***	0.10	0.07	0.46	0.17	0.34
	(0.006)	(0.011)	(0.013)	(1.48)	(1.38)	(2.08)	(0.04)	(0.12)	(0.09)	(0.37)	(0.19)	(0.43)
X VC Reputation	0.004		0.002	0.63		3.45 *	-0.10		0.02	-0.22		-0.16
	(0.009)		(0.009)	(1.74)		(1.70)	(0.09)		(0.08)	(0.39)		(0.40)
X VC Monitoring Tenure		-0.001	-0.001		1.08 ***	1.23 ***		0.05 **	0.05 ***		0.03	0.02
		(0.002)	(0.002)		(0.26)	(0.28)		(0.02)	(0.02)		(0.05)	(0.05)
Num. Obs.	111,467	111,467	111,467	122,099	122,099	122,099	122,099	122,099	122,099	122,099	122,099	122,099
Mean DV	0.035	0.035	0.035	11.8	11.8	11.8	99.8	99.8	99.8	0.234	0.234	0.234
Num. Dyads	6,860	6,860	6,860	7,071	7,071	7,071	7,071	7,071	7,071	7,071	7,071	7,071
Num. Treated Dyads	452	452	452	452	452	452	452	452	452	452	452	452
Dyad FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	No	No	No	No	No	No	No	No	No	No	No	No
Country x Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Round Type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Flight Cancellation

	Log. Count Patents			Follow-on Funding (×100%)			Survival (×100%)			Exit (×100%)		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)	(3a)	(3b)	(3c)	(4a)	(4b)	(4c)
Post X Treat	-0.009	0.006	-0.003	-2.41 **	-8.85 ***	-8.09 ***	-0.00	-0.36 **	-0.38 *	-0.35 **	-0.24	-0.32 *
	(0.009)	(0.011)	(0.014)	(0.88)	(1.33)	(1.43)	(0.11)	(0.12)	(0.15)	(0.11)	(0.17)	(0.16)
X VC Reputation	0.016		0.016	-1.36		-1.26	0.03		0.03	0.13		0.13
	(0.011)		(0.011)	(1.13)		(1.10)	(0.13)		(0.13)	(0.14)		(0.14)
X VC Monitoring Tenure		-0.001	-0.001		1.39 ***	1.39 ***		0.09 ***	0.09 ***		-0.01	-0.01
		(0.002)	(0.003)		(0.21)	(0.21)		(0.02)	(0.02)		(0.03)	(0.03)
Num. Obs.	183,892	183,892	183,892	192,705	192,705	192,705	192,705	192,705	192,705	192,705	192,705	192,705
Mean DV	0.0534	0.0534	0.0534	11.4	11.4	11.4	99.7	99.7	99.7	0.336	0.336	0.336
Num. Dyads	8,305	8,305	8,305	8,305	8,305	8,305	8,305	8,305	8,305	8,305	8,305	8,305
Num. Treated Dyads	799	799	799	799	799	799	799	799	799	799	799	799
Dyad FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	No	No	No	No	No	No	No	No	No	No	No	No
Country x Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Round Type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Robust standard errors, clustered at the dyad level, are shown in parentheses. *VC Reputation* is based on the number of investments of a VC five years prior to the treatment. It is 1 for VCs which have a number larger than the median of investments and 0 otherwise. *VC Monitoring Tenure* is the number of years since the VC has first invested in the startup. Base outcome model for this analysis refers to (c) models in 3.3. Stars indicate significance level: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

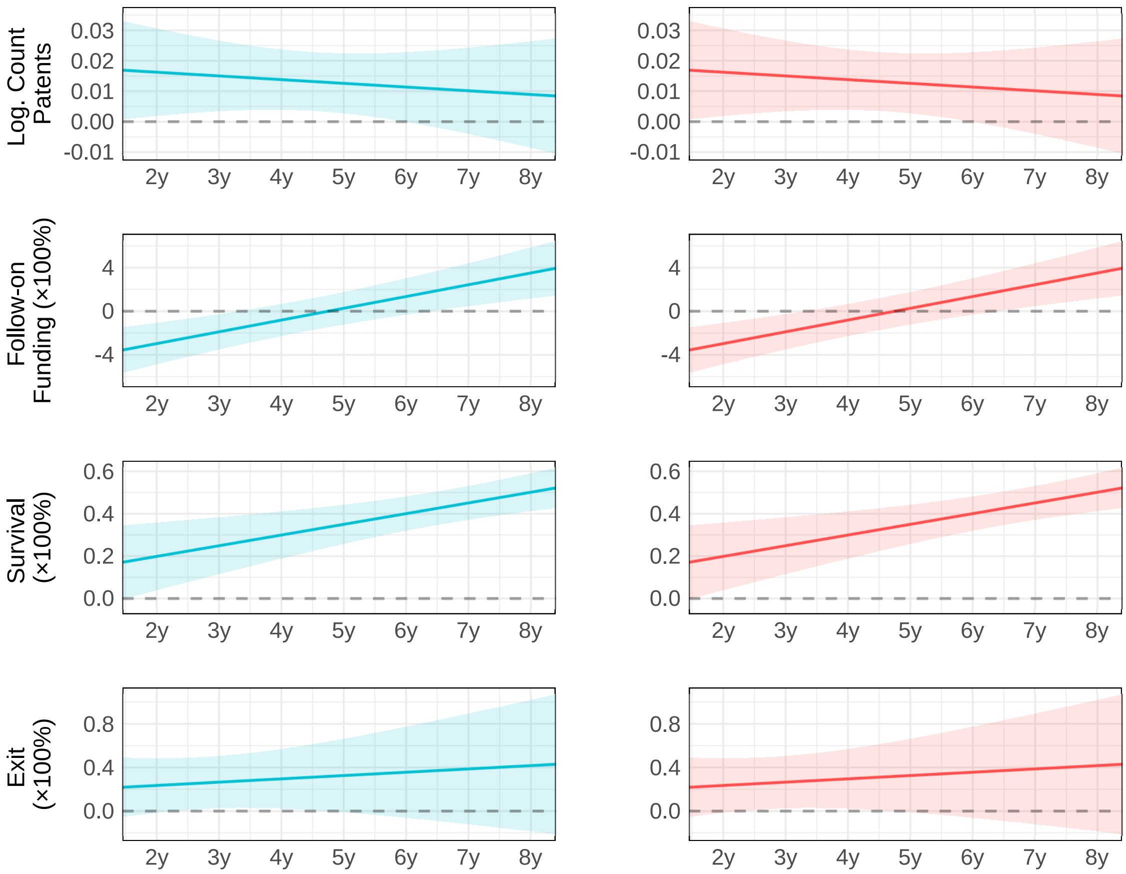
VC reputation. As hypothesized in H2a and H2b, we expect the treatment effects to be moderated by the reputation of the lead VC, increasing the magnitude of the shocks.

Our findings suggest that the effect of VC reputation on on-site monitoring is limited. The only observed positive coefficient is for the follow-on funding rate; however, the base effect was not statistically significant, and only when combined with long-involved dyads and a renowned VC does the estimated effect exceed zero.

Beyond this observation, we do not identify any additional moderation effects of VC reputation across either sample. Consequently, we find minimal or no evidence to support H2a and H2b, respectively. VC reputation does not appear to consistently moderate the effects of on-site monitoring.

VC monitoring tenure. In H3a and H3b, we hypothesize that the treatment effects are influenced by the length of VC involvement, refered to as VC monitoring tenure. The results are presented in Table 3.4 and Figure 3.6.

FIGURE 3.6: Effect Heterogeneity by VC Monitoring Tenure



Notes. Plots on the left side (blue) illustrate moderated treatment effects of the introduction of flights by time of involvement in years. Each subplot shows the effects for a different success measure. The subplots on the right side (red) show the same moderations for flight cancellations. Both refer to the colum (b) specification in 3.4. Values on the x-axis range from 10th percentile to 90th percentile. Confidence intervals are at the 95% level.

Interaction coefficients are significant and show in the same direction for follow-on funding and survival rates across both samples. They are also robust to the inclusion of the previously discussed interaction of treatment and VC reputation.

Thus, for the introduction sample, for follow-on funding, the effect is only positive for dyads with tenure of more than 3.5 years, as shown in Figure 3.6. In contrast, the moderating effects on survival rates are more pronounced, suggesting that dyads with longer relationships benefit more from improved travel connections in terms of increased survival chances.

A similar pattern emerges among dyads experiencing disrupted travel, where younger dyads appear to be more reactive to the treatment and suffer disproportionately from travel interruptions. Notably, the effect on survival begins to turn positive for dyads involved for more than 3.5 years, indicating a potential threshold for resilience to emerge. This finding highlights the importance of relationship duration in buffering the impacts of travel shocks.

While the results for H3a do not support our hypothesis, and, in the case of follow-on funding, the observed effects contradict expectations, we find limited support for H3b. Longer-involved dyads demonstrate reduced sensitivity to treatment effects, with follow-on funding and survival rates being less negatively affected. Potential explanations for these findings are explored in the discussion part.

3.6 Robustness Checks

3.6.1 Funding Amount

In addition to controlling for the startup's development stage, incorporating the amount of funding secured - an important proxy for externally perceived quality and value - can improve estimation accuracy. However, funding amounts are not consistently reported across the entire sample. To address this limitation in a robustness check, we apply two strategies: (1) we restrict the sample to dyads that consistently report complete funding amounts, excluding all observations with missing values, which substantially reduces the sample size; and (2) we impute missing funding values where reasonable and methodologically sound.

The imputation follows a three-step approach. First, when funding amounts are missing between two known rounds, we interpolate the missing value linearly, adjusting for the time interval between rounds. Second, based on startups with complete funding histories, we estimate a predictive model using the preceding round amount and total funding raised as explanatory variables. Where these predictors are available, we use the model to impute missing values. Third, to impute missing values for first-round funding - where prior round data does not exist - we compute the median ratio of first to second round funding (1.75 in our sample) and apply it to back-calculate the missing first-round amount. Startups for which a complete funding series cannot be imputed through these steps are excluded from the imputed sample.

The first approach, which relies on complete reporting, carries two limitations: the smaller sample size reduces statistical power, and the exclusion of non-reporting startups could introduce selection bias, as disclosure may be strategically motivated. The second approach depends on the validity of the imputation strategy, which assumes that funding dynamics can be accurately modeled based on available data.[18]

In our analysis, funding amounts are used in two ways: (i) as an outcome variable, and (ii) as a control variable, where we include the cumulative raised amount up to the observation period. To address skewness and the presence of outliers, all funding amounts are log-transformed.

The results presented in Table 3.5 largely corroborate the main analysis. Reductions in travel frictions are associated with increases in innovation activity, survival rates, and exit likelihoods, while follow-on funding appears less likely. Conversely, when travel frictions emerge, we observe a decline in both follow-on funding and exit rates. Additionally, one specification using imputed funding amounts suggests a potential decrease in innovation activity for the cancellation sample. The results with regard to funding amounts as outcome variable are consistent with what we show for funding rounds.

[18] In fact, since conditioning the introduction sample on always-reported funding amounts restricts the sample to only 47.3% of dyads, and even after imputation we retain just 78.4%, there may still be some selection bias - albeit a weaker one. The respective values for the cancellation sample are 43.1% and 79.8%.

TABLE 3.5: Main Hypothesis with Controls for Funding Amount

Panel A: Flight Introduction

	Log. Count Patents		Follow-on Funding (×100%)		Log. Follow-on Funding Amount		Survival (×100%)		Exit (×100%)	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)	(5a)	(5b)
Post X Treat	0.020 +	0.019 *	-0.55	-2.42 *	-0.09	-0.37 **	0.36 ***	0.25 **	0.42 +	0.29 +
	(0.012)	(0.008)	(1.13)	(0.98)	(0.16)	(0.14)	(0.11)	(0.09)	(0.22)	(0.15)
Num. Obs.	50,772	87,846	55,571	96,002	55,571	96,002	55,571	96,002	55,571	96,002
Mean DV	0.0371	0.039	10.3	12.7	99.7	99.8	0.263	0.253	1.46	1.8
Num. Dyads	3,239	5,400	3,344	5,546	3,344	5,546	3,344	5,546	3,344	5,546
Num. Treated Dyads	213	368	213	368	213	368	213	368	213	368
Dyad FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country x Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Round Type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Raised Amount	Complete	Imputed	Complete	Imputed	Complete	Imputed	Complete	Imputed	Complete	Imputed

Panel B: Flight Cancellation

	Log. Count Patents		Follow-on Funding (×100%)		Log. Follow-on Funding Amount		Survival (×100%)		Exit (×100%)	
	(6a)	(6b)	(7a)	(7b)	(8a)	(8b)	(9a)	(9b)	(10a)	(10b)
Post X Treat	0.009	-0.012 +	-2.67 **	-4.90 ***	-0.41 ***	-0.72 ***	0.11	0.12	-0.28 **	-0.33 ***
	(0.009)	(0.006)	(0.83)	(0.66)	(0.12)	(0.10)	(0.11)	(0.07)	(0.10)	(0.06)
Num. Obs.	76,275	146,035	79,641	152,830	79,641	152,830	79,641	152,830	79,641	152,830
Mean DV	0.0653	0.0633	10.1	12.4	99.7	99.7	0.392	0.359	1.51	1.84
Num. Dyads	3,578	6,627	3,578	6,627	3,578	6,627	3,578	6,627	3,578	6,627
Num. Treated Dyads	333	630	333	630	333	630	333	630	333	630
Dyad FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country x Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Round Type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Raised Amount	Complete	Imputed	Complete	Imputed	Complete	Imputed	Complete	Imputed	Complete	Imputed

Notes. Robust standard errors, clustered at the dyad level, are shown in parentheses. Specification (a) refers to sample with full funding information and specification (b) adds imputed funding information. Sample only contains dyad for which we consistenly observe monetary funding amount. Stars indicate significance level: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.6.2 Differentiated Exit Variable

To further examine the nature of the observed exit effects, we introduce alternative dependent variables that differentiate between IPOs and acquisitions as forms of exit. As shown in Table 3.6, the results are primarily driven by acquisitions. This pattern is likely attributable to the limited number of IPOs in the sample, with only 28 observed in the introduction sample and 52 in the cancellation sample. The low frequency of IPOs restricts statistical power and suggests that we should interpret exit as predominantly acquisitions.

TABLE 3.6: Analysis with Differentiated Exit Variable

Panel A: Flight Introduction

	Acquisition (×100%)			IPO (×100%)		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Post X Treat	0.28 *	0.45 *	0.30 *	-0.01	-0.03 +	-0.02
	(0.13)	(0.22)	(0.15)	(0.03)	(0.02)	(0.03)
Num. Obs.	122,099	55,571	96,002	122,099	55,571	96,002
Mean DV	0.211	0.225	0.225	0.0229	0.0378	0.0281
Num. Dyads	7,071	3,344	5,546	7,071	3,344	5,546
Num. Treated Dyads	452	213	368	452	213	368
Dyad FE	Yes	Yes	Yes	Yes	Yes	Yes
Country x Period FE	Yes	Yes	Yes	Yes	Yes	Yes
Round Type	Yes	Yes	Yes	Yes	Yes	Yes
Raised Amount		Complete	Imputed		Complete	Imputed

Panel B: Flight Cancellation

	Acquisition (×100%)			IPO (×100%)		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Post X Treat	-0.28 *	-0.25 **	-0.25 **	0.01	-0.03	-0.03
	(0.14)	(0.09)	(0.09)	(0.03)	(0.04)	(0.04)
Num. Obs.	192,705	79,641	79,649	192,705	79,641	79,649
Mean DV	0.309	0.357	0.357	0.0269	0.0351	0.0351
Num. Dyads	8,305	3,578	3,578	8,305	3,578	3,578
Num. Treated Dyads	799	333	333	799	333	333
Dyad FE	Yes	Yes	Yes	Yes	Yes	Yes
Country x Period FE	Yes	Yes	Yes	Yes	Yes	Yes
Round Type	Yes	Yes	Yes	Yes	Yes	Yes
Raised Amount		Complete	Imputed		Complete	Imputed

Notes. Robust standard errors, clustered at the dyad level, are shown in parentheses. In other analyses, acquisitions and IPOs are combined as exits. Stars indicate significance level: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.7 Discussion

Across various outcomes and the two shocks analyzed, we observe effects of moderate magnitude. Specifically, the introduction of direct travel options between lead VCs and portfolio startups leads to increases in innovation activity, survival rates, and exit rates while at the same time less new funding rounds are completed. In contrast, disruptions to travel options result in negative effects on follow-on funding and exit rates.

These findings can be partly explained by the differing nature of the two shocks affecting VC monitoring. The positive shock - introduced by new flight routes - reduced travel costs and permanently improved accessibility between VCs and portfolio startups. Enhanced opportunities for in-person collaboration and increased efficiency may have substituted for some of the support typically obtained through follow-on funding, potentially explaining the observed decline in funding frequency despite other positive outcomes. In contrast, the negative shock - flight cancellations during the COVID-19 pandemic - temporarily disrupted travel and increased monitoring costs. Although temporary, the uncertainty surrounding the duration of the pandemic at the time likely amplified its immediate impact on investor-startup relations and resource flows.

The following sections discuss the outcomes in detail and explore potential mechanisms underlying these results.

Innovation activity. The immediate effect of flight introductions on the number of patent applications is modest but becomes more pronounced after a few years, suggesting that the impact takes time to materialize. This aligns with Bernstein et al. (2016) and is consistent with the nature of patent applications, which often result from cumulative efforts unfolding over several years. In contrast, the analysis period following flight cancellations may have been too short to observe significant disruptions to these processes. Research projects already underway, with secured funding, may have continued uninterrupted despite the temporary travel disruptions.

Follow-on funding. For follow-on funding, we observe immediate effects following the travel disruptions caused by the pandemic. Over the observed time horizon, the likelihood of completing follow-on funding rounds decreases for dyads experiencing flight interruptions compared to those with uninterrupted travel options, with the effect slightly intensifying over time. Follow-on funding typically occurs at regular intervals, providing startups with the necessary resources to sustain growth. Disruptions to VC monitoring caused by flight cancellations likely impede the VC's ability to signal confidence and demonstrate active engagement to other investors. This uncertainty surrounding the startup's future increases difficulty in securing additional funding rounds.

In contrast, we find that flight introductions lead to a modest increase in follow-on funding during the first year, followed by a gradual decline and a negative trend in subsequent years. Because this pattern does not appear to come at the expense of other success outcomes, we cautiously interpret this as evidence that enhanced on-site monitoring may improve startup efficiency. While we are unable to directly test this mechanism due to data limitations, prior research offers supporting evidence that more intensive VC involvement can lead to more efficient operations and better strategic decision-making (Chemmanur et al., 2011; Hellmann and Puri, 2000).

Operational outcomes. Many startups operate with negative cash flow and rely heavily on follow-on funding to sustain their operations. However, while we observe a decline in follow-on funding rates following travel disruptions, we do not find a corresponding decrease in survival rates. One possible explanation is that survival outcomes take longer to materialize. Startups may still have access to prior funding resources that are not yet fully exhausted, aligning with the observed delayed positive effect of flight introductions on survival rates, which becomes significant only after a few years.

Another explanation could be the extensive business support schemes implemented during the pandemic in the countries included in our sample. These programs, targeting small and medium-sized enterprises (SMEs), provided financial relief through grants, reduced interest rates, guaranteed debt financing, wage subsidies, tax concessions, and subsidized business advice.[19] While these measures aimed to mitigate the adverse economic impacts of the pandemic, their immediacy may have temporarily delayed business closures, masking the true effect of VC monitoring on survival rates. This could explain the divergence between survival rates and follow-on funding findings with regard to travel disruptions.

With regard to exits, we find that the observed effects are predominantly driven by acquisitions. Due to the limited number of IPOs in the sample, no meaningful conclusions can be drawn about the effect on IPOs. The effects on acquisitions align with expectations. Acquisitions depend heavily on sustained trust, active signaling, and close collaboration among the startup, the lead VC, and potential acquirers. Temporary disruptions caused by flight cancellations hinder the VC's ability to provide hands-on support, facilitate negotiations, and signal confidence, leading to immediate negative effects on acquisition rates. Additionally, logistical challenges associated with flight interruptions can delay or derail time-sensitive acquisition processes, such as due diligence or deal closure. Reduced accessi-

[19] `https://www.worldbank.org/en/data/interactive/2020/04/14/map-of-sme-support-measures-in-response-to-covid-19` [accessed: 2024-12-20]

bility of affected startup locations to other investors and stakeholders further complicates the acquisition process.

In contrast, flight introductions lead to permanent improvements in travel accessibility, enhancing the VC's capacity to support startups and engage with acquirers over the long term. Improved travel logistics facilitate better relationship-building and more efficient acquisition processes, resulting in positive effects on acquisition rates.

VC reputation. An especially intriguing finding is the limited evidence of a moderation effect by VC reputation. At first glance, this appears surprising, as previous studies have consistently documented that reputable VCs drive positive success outcomes (Sørensen, 2007; Hochberg et al., 2007; Nahata, 2008; Clarysse et al., 2013; Chemmanur et al., 2011; Quas et al., 2021). A significant portion of the VC reputation effect is attributed to market signaling or certification power, whereby reputable VCs convey unobservable qualities of the startup to other investors and stakeholders, thereby validating the startup and its business model (Megginson and Weiss, 1991; Hochberg et al., 2007; Pollock et al., 2010; Quas et al., 2021).

Based on our findings, we argue that the signaling and certification roles of reputable VCs are not dependent on physical proximity. These functions likely operate effectively through other mechanisms, such as established networks or perceived credibility, without requiring face-to-face interactions. This may explain the lack of a moderation effect of VC reputation in the context of our study.

VC monitoring tenure. Interestingly, the duration of involvement between lead VCs and startups moderates the effects of flight shocks positively across both samples. Dyads with longer-established relationships demonstrate better follow-on funding outcomes compared to those with shorter relationships. Mature relationships are better positioned to leverage enhanced travel opportunities, enabling them to capitalize on improved travel connections more effectively. At the same time, they exhibit greater resilience to travel disruptions. Factors such as trust, well-established processes, and accumulated familiarity likely underpin this resilience, allowing these dyads to navigate the challenges posed by disruptions while maintaining stronger collaboration.

3.7.1 Empirical Contributions

Our study contributes to the empirical literature examining the impact of VCs beyond their provision of financial resources (see e.g., Bernstein et al., 2016; Bronzini et al., 2020; Kortum and Lerner, 2000; Quas et al., 2021; Bertoni et al., 2011; Sørensen, 2007; Chemmanur et al., 2011; Puri and Zarutskie, 2012). This body of work has analyzed a range of entrepreneurial success variables. We expand on these contributions by simultaneously examining multiple dimensions of startup activity and performance, including innovation activity, follow-on funding, survival rates, and exit rates. This comprehensive approach provides a more nuanced understanding of VCs' influence. Furthermore, our study sheds light on the differences between permanent shocks (flight introductions) and temporary shocks (flight cancellations), highlighting how both the direction and duration of these shocks affect outcomes. We also explore how these effects evolve over time and are moderated by factors such as VC reputation and the length of VC-startup collaboration.

On the methodological front, we build on an estimation strategy previously employed by Bernstein et al. (2016) in the same context and by others in related fields of scientific collaboration (Catalini et al., 2020), exchange between academic and industry (Baruffaldi and Poege, 2024), and plant investment (Giroud, 2013). By leveraging exogenous shocks, we isolate the effect of face-to-face or on-site VC monitoring, offering a significant advantage over studies that rely on matching on observables (Hellmann and Puri, 2002; Chemmanur et al., 2011; Puri and Zarutskie, 2012; Bertoni et al., 2011). Matching approaches assume that unobservable characteristics do not influence outcomes, a particularly restrictive assumption in the dynamic and opaque domain of startups.

By focusing specifically on the on-site monitoring effect, which excludes monitoring activities performed remotely, our study also contributes to the broader literature on remote work and the role of communication technologies in replacing in-person interactions, often associated with the "death of distance" hypothesis (Cairncross, 2001). While digital tools have been shown to complement physical proximity, empirical evidence consistently underscores the continued importance of face-to-face interactions for collaboration and success (Yang et al., 2021; Catalini et al., 2020; Forman et al., 2005; Blum and Goldfarb, 2006). Our findings reinforce this perspective, demonstrating that while technological advances enhance remote engagement, they do not fully substitute for the unique benefits of in-person interactions in VC monitoring and decision-making processes.

3.7.2 Practical Implications

Although the effects vary by the type of outcome and the nature of the shock, our findings show that the distance and ease of travel between lead VCs and portfolio startups significantly impact monitoring abilities. This has several practical implications for VCs, startups, policy-makers, and the broader entrepreneurial ecosystem.

When selecting and screening portfolio startups, VCs should consider the weight the put on incorporating travel accessibility into their decision-making processes, prioritizing investments in startups that are geographically closer or more accessible via reliable travel routes. In cases where physical proximity is lacking, VCs could adopt hybrid monitoring strategies, reserving face-to-face interactions for critical activities such as due diligence, trust-building, and facilitating acquisitions, while leveraging virtual tools for routine updates and communication.

Startups can also strategically position themselves in areas with better travel accessibility to attract and retain VC investment. Proximity to major travel hubs increases the likelihood of securing funding and enables startups to leverage their lead VC's involvement more effectively. This proximity can also enhance the startup's ability to signal quality to other investors, particularly during pivotal phases like follow-on funding and acquisitions.

For policy-makers, the Covid-19 pandemic underscored the fragility of global mobility and its impact on economic activities. Governments can use these insights to design better crisis management strategies that support startups and VCs during periods of travel disruption. Recognizing the limitations of remote work, policy-makers can also focus on fostering innovation hubs in well-connected cities or regions to enhance the natural clustering of startups and investors. Such efforts could improve access to funding, promote collaboration, and accelerate innovation.

3.7.3 Limitations and Future Research

Most effects are of moderate magnitude and sometimes, only when aggregated, turn out statistically significant. Two potential explanations may account for this observation. First, the relatively small number of treated dyads, particularly in the introduction sample, may lead to underpowered tests. Expanding the sample to include startups from a broader range of countries, not limited to the United States, Canada, the United Kingdom, and Germany, which have generally been well-connected in recent decades, could substantially increase the number of treated dyads. This expansion would not only enhance the statistical power of the tests but also provide valuable insights into the dynamics of VC monitoring and travel accessibility on a more global scale.

Second, our analysis focuses on the on-site monitoring effect, which represents only a portion of the total impact of VC monitoring. While advances in communication technologies over recent decades have enabled effective remote collaboration, our study specifically captures the aspects of monitoring that require physical proximity. As a result, the estimates provided in this analysis should be interpreted as a lower bound of the total effect of VC monitoring. The actual impact, which includes contributions from both on-site and virtual interactions, is likely larger in magnitude.

This limitation is not necessarily a drawback but instead highlights the unique aspects of VC monitoring that cannot be substituted by virtual interactions. These findings underscore the critical role of physical proximity for certain key activities, such as trust-building, due diligence, and strategic decision-making. Future research that explicitly disentangles the components of VC monitoring, those that can be replaced by virtual tools and those that cannot, could provide deeper and more nuanced insights into the mechanisms driving startup success.

Another fruitful avenue for future research would be to investigate the formation of investor-startup dyads contingent on travel costs and the implications of geographical accessibility for partnership decisions. It would also be valuable to examine whether travel costs influence the prevalence or structure of VC syndication, whether it encourages collaboration among multiple VCs or alters the dynamics of syndication agreements.

Chapter 4

When Startups Fail - The Role of Media in Shaping Entrepreneurial Trajectories

4.1 Introduction

Entrepreneurship is inherently uncertain, and failure is an inevitable part of the entrepreneurial journey. While some ventures achieve success, many do not survive, making failure a defining characteristic of entrepreneurial activity. However, failure is not merely an objective outcome; rather, it is a socially constructed phenomenon, shaped by the environment, institutional forces, and societal expectations. How failure is perceived and evaluated can have profound consequences for entrepreneurs, influencing their ability to recover, re-enter the market, and maintain their legitimacy in the broader entrepreneurial ecosystem.

The way entrepreneurial failure is interpreted might depend on various contextual factors, including the expectations placed on a startup before its failure. High levels of funding, industry hype, and deviations from established norms can all influence how failure is framed. Moreover, the audience evaluating the failure - whether investors, the media, or the public - is itself embedded in social and cultural contexts that shape its perception. Understanding these dynamics is essential for uncovering how entrepreneurial failure is framed and how it affects the future trajectories of failed entrepreneurs.

The perception of failure is often negative, despite its natural role in entrepreneurial processes (McGrath, 1999). The stark contrast between aspiration and outcome creates a stigma around failure, reinforcing the societal tendency to view it as a personal or strategic shortcoming rather than an expected consequence of risk-taking (Lubatkin and Chatterjee, 1994; Shepherd and Patzelt, 2017). Yet, this perception is neither universal nor static; rather, it is influenced by broader societal norms, cultural narratives, and institutional frameworks (Aldrich and Martinez, 2003).

A key institution shaping these evaluations is the media. Media coverage acts as a powerful mechanism through which entrepreneurial failure is framed and disseminated to the public (Pollock and Rindova, 2003). The media serves as an information intermediary, influencing stakeholder perceptions by legitimizing or delegitimizing failed entrepreneurs and their ventures (Überbacher, 2014). For startups, which already face the "liability of newness" (Singh et al., 1986; Stinchcombe, 2000), media narratives can be particularly consequential, shaping investor confidence, public opinion, and the entrepreneur's ability to recover. Despite its importance, the role of media in shaping the social evaluation of failure remains underexplored, particularly in how different contextual factors moderate its impact.

Empirically, scholars have widely used media reports as means to understand organizational legitimacy (Brown and Deegan, 1998; Deephouse, 1996; Pollock and Rindova, 2003; Deephouse and Carter, 2005; Elsbach, 1994; Lamertz and Baum, 1998), particularly the tone of coverage (Deephouse, 1996). It has also been shown that news media can categorize startups as legitimate or illegitimate (Simmons et al., 2014) and pressure firms into action (Brown and Deegan, 1998).

However, very few studies have empirically examined media coverage of business failure, with the exception of listed Chinese underperforming firms receiving more negative coverage (Sheng and Lan, 2019), firms involved in wrongdoing and recalling facing a reduction in social approval (Zavyalova et al., 2012), and a news sentiment comparison of entrepreneurial failure, finding that the decline is larger in the United States than in Germany (Jacobi, 2018).

This study examines how entrepreneurial failure is evaluated through media coverage, focusing on the contextual factors that shape its portrayal. Expectations play a key role in this process - factors such as funding amounts, industry hype, and the extent to which a startup aligns with or deviates from established categories and norms all influence how failure is framed. Additionally, the audience interpreting failure - embedded in specific social and cultural contexts - further shapes these evaluations. Understanding these dynamics is crucial for assessing how media narratives construct the societal perception of failure and influence the entrepreneurial re-entry process.

To address this theoretical and empirical challenge, we analyze a large dataset of startups matched with news articles published between 2005 and 2023. Leveraging state-of-the-art natural language processing models, we extract sentiments from 4,636,226 sentences covering 56,068 startups in the United States, the United Kingdom, and Germany. Using robust econometric methods, we find a consistently negative portrayal of entrepreneurial failure, with significant variation based on contextual factors. In particular, startups with higher funding levels experience harsher judgment, while atypical startups seem to be somewhat shielded from negative press. Additionally, media narratives differ across national contexts, reflecting broader cultural and institutional differences in how failure is perceived.

Beyond examining media sentiment, this study also explores its consequences for entrepreneurial re-entry. Our findings indicate that negative media coverage after failure reduces the likelihood of founders starting new ventures. Notably, we find no evidence that entrepreneurs can meaningfully influence the tone of media coverage through impression management strategies during times of crisis. These insights contribute to a broader understanding of the entrepreneurial life cycle, positioning failure not as a definitive endpoint but as a socially constructed phase with long-term implications for individuals and ecosystems alike.

Our work makes two key contributions. First, while many scholars have explored the role of media and legitimacy in entrepreneurship (Brown and Deegan, 1998; Deephouse, 1996; Deephouse and Carter, 2005; Elsbach, 1994; Lamertz and Baum, 1998; Pollock and Rindova, 2003), most of this research has concentrated on legitimacy building in the early stages of the entrepreneurial process. In contrast, limited attention has been given to how business failure is socially evaluated (Zavyalova et al., 2012; Sheng and Lan, 2019; Jacobi, 2018). We address in particular entrepreneurial failure, a critical yet underexplored phase of entrepreneurship (McGrath, 1999), by offering a nuanced analysis of the social evaluation of failure, recognizing it as a distinct and important aspect of the entrepreneurial journey that merits deeper investigation. We document a negative reaction to failure and additionally show that evaluation is negatively moderated by funding amount and positively moderated by atypicality, and the perception that atypical ventures demonstrate greater boldness in innovation.

Second, existing research on legitimacy formation has primarily focused on the diversity of stakeholders involved in the process (Überbacher, 2014; Fisher, 2020; Shepherd and Patzelt, 2015). We extend this work by demonstrating that evaluations of entrepreneurial failure occur within distinct social and cultural contexts (Lee et al., 2007), and are themselves socially constructed (Radu-Lefebvre et al., 2019). Our analysis highlights how these evaluations vary across audiences, using countries as proxies for institutional differences (Fisher et al., 2017). In doing so, we introduce the concept of audience diversity at a societal level, offering a broader perspective on the role of social context in legitimacy evaluations.

More broadly, we contribute to the understanding of the media's role in entrepreneurship and respond to calls for further research on the relationships between media, national culture, and organizations (Hindle and Klyver, 2007; Carroll and McCombs, 2003). Additionally, we expand the literature on the influence of national culture on entrepreneurship, a mainly survey-driven field that has predominantly focused on entrepreneurial behavior but has yielded inconclusive results (Stephan, 2022; Taylor and Wilson, 2012; Shane, 1993; Kleinhempel et al., 2022).

Our findings have several important implications. For founders, negative media coverage reduces their likelihood of restarting a new venture. Founders might feel stigmatized by the negative portrayal and, as a result, refrain from pursuing another entrepreneurial endeavor. Also, the results suggest that startups and founders have no ability to influence the sentiment of news coverage surrounding their failure. At the societal level, the harsh treatment of failed entrepreneurs in the media could discourage overall entrepreneurial activity, potentially stifling innovation and economic growth.

4.2 Theory

4.2.1 Entrepreneurial Failure

Entrepreneurial failure is fundamentally tied to the entrepreneurial process (Wiklund et al., 2010; Bruderl et al., 1992; Shane, 2009), largely due to the irreducible uncertainty faced by new ventures (Venkataraman, 2019). As Shepherd and Patzelt (2017) aptly describe, startup ventures resemble "experiments with unknowable outcomes," where the high-risk nature of returns (Lubatkin and Chatterjee, 1994) significantly increases the likelihood of failure. While much of the research on entrepreneurship focuses on success, failure is a more common outcome and should not be unexpected. Because failure contrasts sharply with the initial goals and aspirations of startups, society often associates failure with shame (McGrath, 1999) and views risk-taking favorably only when it leads to success, otherwise considering it akin to gambling (March and Shapira, 1987).

Generally defined, failure is the absence of success, but entrepreneurial or business failure has been approached with varying definitions in the literature (see Ucbasaran et al. (2013) for a review), ranging from broad definitions that cover any ownership discontinuity (including successful exits like IPOs or acquisitions) to narrow definitions limited to bankruptcy, which fail to capture cases where outcomes fall short of the entrepreneur's expectations. As researchers have stressed the importance to explicitly establish one's own definition (Shepherd and Patzelt, 2017), we draw on previous work (McGrath, 1999; Ucbasaran et al., 2013) to adopt a balanced definition.

Definition. *Failure is the termination of a venture or initiative that falls short of its goals, based on subjective assessments or unmet economic viability thresholds set by the entrepreneur(s).*

This definition encompasses bankruptcies, insolvencies, and instances where entrepreneurs opt to discontinue operations, potentially due to a more promising opportunity, whether it be another venture or traditional employment. However, it excludes successful exits such as acquisitions and IPOs.

4.2.2 Social Evaluation

Rather than being purely objective and existing in isolation, startups, as new entities, are subject to a social construction process (Aldrich and Martinez, 2003) and are shaped through collective human agreement. First, often emerging out of an entrepreneur's vision, the way founders present their startup relies heavily on society's values and expectations. But also the startup's acceptance is influenced by social perception and its perceived value to society. More specifically, due to the liability of newness (Singh et al., 1986; Stinchcombe, 2000), startups, unknown to most actors in their environment, seek legitimacy to increase their limited chances of survival (Singh et al., 1986; Stinchcombe, 2000). An organization can establish legitimacy and gain access to essential resources only once stakeholders such as clients, investors, employees, and business partners are sufficiently informed and confident, having reduced their uncertainty (Aldrich and Fiol, 1994; Bitektine, 2011; Zimmerman and Zeitz, 2002; Lounsbury and Glynn, 2001; Delmar and Shane, 2004; Dobrev and Gotsopoulos, 2010; Tornikoski, 2009).

Organizational legitimacy, described by Oliver (1991) as "social fitness" or by Deephouse (1996) as "the endorsement of an organization by social actors" or "acceptance of the organization by its environment", is a key concept within organizational research. Aldrich and Fiol (1994) and Bitektine (2011) differentiate between cognitive legitimacy judgments that answer whether the organization belongs to any familiar class or category and sociopolitical legitimacy judgments that refer to the organization's right to exist and benefit or hazard to an individual, a group, or the society. The liability of newness is a function of both types of legitimacy, but startups are generally more concerned with cognitive legitimacy (Aldrich and Fiol, 1994).

Entrepreneurial failure, like other components of the entrepreneurial process, is also subject to social construction (Aldrich and Martinez, 2003), with its meaning and consequences shaped by the societal, cultural, and institutional contexts in which it occurs. Social norms, cultural narratives, and evaluative frameworks vary across societies and audiences, influencing how failure is perceived (Cardon et al., 2011). In Silicon Valley, for example, the "fail fast, fail often" mentality celebrates resilience in the face of failure, while in other parts of the world, failure is viewed as a sign of incompetence (Mantere et al., 2013; Wiesenfeld et al., 2008), and recovery is often hindered by institutional and legal constraints (Lee et al., 2007). Importantly, both the legitimation process and the loss of legitimacy after failure are significantly influenced by the characteristics and attributes of the startup itself, the object being evaluated (Shepherd and Patzelt, 2015).

4.2.3 Media

A key actor in the social construction and evaluation process is the media (Gamson et al., 1992), which shapes the public's perception of organizations, including startups, by rendering social judgments. Without media, most stakeholders would remain uninformed about startups, precluding the formation of any substantive opinion. Consequently, media functions as an "infomediary" (Deephouse and Heugens, 2009), short for information intermediary organizations, facilitating the dissemination of relevant information to the general audience and creating a public forum for discussion.

Among the large body of literature in the field of mass communication and media studies, the agenda-setting theory (McCombs and Shaw, 1972) is one of the most prominent and widely used frameworks to explain how media affects public opinion and audience thinking. It suggests that media can influence the public agenda by choosing which issues and topics receive attention. Media cannot tell the audience “what to think”, but “what to think about”, thereby shaping its thinking and behavior (Hindle and Klyver, 2007) and prioritizing public concerns.

Although previously mainly applied in political contexts, Carroll and McCombs (2003) argue that the agenda-setting theory applies equally well to business news and corporate reputations. The authors suggest that the more positive news attention a company receives, whether in general or focused on specific aspects of the company, the more positive the public perception of the company and its corresponding qualities becomes. Drawing on agenda-setting theory, Deephouse and Heugens (2009) position professional news media as prototypical infomediaries, emphasizing their role as information processors and brokers who bring structure and rhythm to otherwise unstructured and unpredictable events (McQuail, 1985; Molotch and Lester, 1974). As gatekeepers between events and the public’s perception of those events, media controls the processes of gathering, filtering, framing, and disseminating information (Tuchman, 1978). Media institutions’ own biases, beliefs, and interests inevitably influence the selection and reporting of news. Through their influence on public knowledge and opinion about firms, as concluded by Deephouse (2000), and their active participation in the social construction of reality, as noted by Gamson et al. (1992), the media plays a pivotal role in shaping societal narratives.

4.2.4 Entrepreneurial Failure, Social Evaluation, and Media

For nascent startups, media coverage therefore plays a key role in their journey to gain organizational legitimacy. More generally for organizations, Zavyalova et al. (2012) point out that news media can actively determine a firm’s social approval through what we have just described as agenda-setting, by selecting how and what organizations to cover.

According to Simmons et al. (2014), startups can be categorized as legitimate or illegitimate through information intermediaries, and, following Bitektine (2011), media can render legitimacy by judging organizations based on their activities and assessing the overall value to society. Brown and Deegan (1998) find that, in line with agenda-setting theory, the intensity of news coverage can drive the community's concern about environmental issues and pressure firms into action.

Reviewing the relevant literature, Überbacher (2014) identifies five distinct perspectives on legitimation of new startups, of which the most widely used is an institutional perspective. Institutions shape how audiences form legitimacy judgments, and two types can be distinguished: cognitive and evaluative institutions, with the latter being most relevant to our study.[1] Evaluative institutions are powerful institutional entities such as media organizations that the audience relies on for its judgments and resource allocation decisions. In a literature review of media coverage of firms, Graf-Vlachy et al. (2020) identify the emphasis of news media as an institutional force in a firm's environment evaluating "appropriateness and desirability of actions" and affecting "impression formation and the legitimation of firms" (Pollock and Rindova, 2003; Fombrun and Shanley, 1990; Deeds et al., 2004).

Therefore, scholars pay particular attention to the tone of media coverage to quantify legitimacy of organizations (Brown and Deegan, 1998; Deephouse, 1996; Deephouse and Carter, 2005; Elsbach, 1994; Lamertz and Baum, 1998; Pollock and Rindova, 2003) and interpret it as a measure of an organization's legitimacy with the public (Deephouse, 1996). Graf-Vlachy et al. (2020) classify this research into news media "recording" social approval (Deephouse, 2000; Ferguson et al., 2000; Bundy and Pfarrer, 2015; Lamin and Zaheer, 2012; Vanacker and Forbes, 2016; Zavyalova et al., 2017) and actively "influencing" it (Deephouse, 2000; Durand and Vergne, 2015; McDonnell and King, 2013).

The effect of legitimation through media coverage has been mainly studied for public firms. Pollock and Rindova (2003) show that, in line with what is expected, for public offerings, the volume of media-provided information decreases underpricing and increases stock turnover. Fang and Peress (2008) and Gurun and Butler (2012) reveal particularly strong effects for firms that are small, are illiquid, have a low analyst following, or are mostly held by individual investors (Graf-Vlachy et al., 2020), which is consistent with our

[1] Cognitive institutions refer to beliefs held by the audience about what organizations have to adhere to and how to behave within a given field of activity (e.g. presenting a business plan or acting economically responsibly (Delmar and Shane, 2004; Karlsson and Honig, 2009; Meyer and Rowan, 1977; Wiklund et al., 2010))

view that media plays a particularly important role for nascent startups. Horverak (2009), Kaniel and Parham (2017), and Solomon et al. (2014) find the effect of media coverage to positively affect a diverse range of products. Inconsistent with other findings, Berger et al. (2010) show even negative coverage to be beneficial for sales by increasing awareness.

Failure incidents, in particular, are often considered newsworthy and covered in the press due to their potentially negative and undesirable nature. Even in the academic literature, McGrath (1999) observes a negative attitude towards entrepreneurial failure and in general, humans tend to show a negativity bias by giving greater weight to negative events (Rozin and Royzman, 2001). In media reporting, the presence of negativity bias has been observed as well. Humans are "often drawn towards more negative and emotionally triggering news" (Ahern and Sosyura, 2015). In the context of financial news, the addition of negative words to a headline increases consumption rates of online news (Robertson et al., 2023), which can also be shown by physiological reactions, such as changes in skin conductance levels and heart rate variability triggered by negative news (Soroka et al., 2019). An obvious business interest of media organizations to cover negative events follows.

Sheng and Lan (2019) find listed Chinese underperforming firms to receive more frequent and negative coverage, and Zavyalova et al. (2012) analyze news articles about wrongdoing and product recalls by US toy companies, documenting a drop in the level of social approval. Only Jacobi (2018) compares media evaluations of failure in a startup context, finding that the decline in news sentiment is larger in the United States than in Germany, likely due to a higher prior sentiment. Thus,

Hypothesis 1. Following failure, the startup's media coverage becomes more negative.

4.2.5 Contextual Factors

In the following paragraphs, we argue that the evaluation of startup failure is contingent on contextual factors. Media might scrutinize highly funded startups or startups assigned to high-profile, hyped industries differently, leading to varying consequences. Also, atypical startups, those diverging from conventional categories within an industry, might be partly shielded for breaking the norms. Moreover, the context of the audience, such as entrepreneurial attitudes in different countries, could influence how failure is reported.

Funding

Cardon et al. (2011) categorize the reasons for entrepreneurial failure into two main types: misfortunes and mistakes. Among the misfortunes, market forces (40%) and funding constraints (29%) are the most frequently cited causes. These categories are likely to prompt different evaluations, as entrepreneurs facing misfortunes are often viewed with more sympathy, while those responsible for professional mistakes are more likely to be held accountable. Consequently, startups with lower funding levels may rather be seen as victims of external circumstances, whereas better-funded ventures may be criticized as failure is perceived to be caused by inadequate execution.

Partly contrasting, Thornhill and Amit (2003) argue that failure in younger, and typically less-funded firms, stems from deficiencies in managerial knowledge and financial skills, while older firms tend to fail due to their inability to adapt to environmental changes.[2] Consequently, because younger, less-funded firms may not possess the same organizational skills and routines, this could lead to greater leniency toward their failures.

In addition to the attribution of failure, expectations play a key role in the reaction to a startup's demise. Media attention increases a startup's perceived potential and enhances its ability to secure venture capital funding (Petkova et al., 2013). It bolsters a startup's legitimacy (Pollock and Rindova, 2003; Deephouse, 2000), especially when it secures funding from high-status investors (Pollock and Gulati, 2007). However, this heightened legitimacy is accompanied by greater expectations, making failures of well-funded startups more visible and subject to harsher criticism. As Garud et al. (2014) point out, stories projecting future success are essential for gaining legitimacy but can become sources of disappointment and legitimacy loss in the event of failure. Moreover, since failure rates are particularly high in the early stages of startup development (Thornhill and Amit, 2003), startups that fail at later stages often face even greater disappointment, as they are expected to be more resilient and their failure is perceived as more unexpected compared to their younger counterparts. Thus,

Hypothesis 2. Following failure, the change in media coverage of startups with higher levels of funding is more negative than for startups with lower levels of funding.

[2]Here, age is merely a proxy for managerial competence.

Hype

Expectations arise not only from the amount of funding raised but also from the industry in which a startup operates, particularly when that industry is subject to significant hype. Logue and Grimes (2022) defines hype as "a collective vision and promise of a possible future, around which attention, excitement and expectations increase over time". However, these high expectations are frequently unmet, resulting in a loss of legitimacy. This dynamic explains why hype is often perceived as somewhat deceptive in public discourse (Van Lente et al., 2013; Ruef and Markard, 2010). A typical hype cycle involves an initial peak, where expectations and future potential are exaggerated, followed by a phase of disappointment as outcomes fall short, leading to societal and stakeholder disillusionment (Van Lente et al., 2013).

Hype can be measured at various levels, such as the startup or industry level. When researching entrepreneurial failure, industry hype is particularly relevant, as startups can gain legitimacy through their affiliation with a industry. For instance, Shepherd and Patzelt (2015) demonstrate that failed startups are judged more leniently when they focus on socially beneficial causes like environmental preservation and similarly, Aldrich and Fiol (1994) highlight that perceptions of a startup are often closely linked to the broader characteristics of the industry. On the example of the once praised market category of genetically modified foods, Deeds et al. (2004) argue that all participants in an industry are affected by its legitimacy. Consequently, emerging market categories provide spillover benefits, allowing startups with little or no track record to receive considerable recognition (Younger and Fisher, 2020). Thus, industry affiliation serves as a resource through which entrepreneurs can attract stakeholder support and access critical resources.

For the legitimation of industries, media plays a critical role (Navis and Glynn, 2010) by informing the general public, and creating and amplifying hype as expectations are "often spread through spoken and written words" (Van Lente et al., 2013). Increased media attention, driven by large expectations of future potential, helps establish legitimacy.

However, this initial appreciation and legitimacy can backfire if these startups fail to meet set expectations which causes a sharp decline in legitimacy (Garud et al., 2019). Since deceit is often associated with hype (Logue and Grimes, 2022), failure may result in greater scrutiny as media now identifies a startup which does not fit into the previous provided narrative and needs to explain why it deviates to the audience.

Interestingly, though disappointing at the startup level, it does not necessarily lead to disappointment at the industry level (Ruef and Markard, 2010). Consequently, because the gap opens between the hyped industry and the failed startup, we argue that, upon failure, a startup rapidly loses both its affiliation with the hyped industry and the legitimacy that came with it, leading to harsher evaluations by the media. When the hype surrounding the startup's industry fails to materialize due to the startup's failure and its failure to align to the hype, the heightened expectations and initial legitimacy can have particularly negative consequences on the assessment of the failure. Thus,

Hypothesis 3. Following failure, the change in media coverage of startups from industries with great hype is more negative than for startups from industries with little hype.

Atypicality

Evaluations likely depend on the positioning of a startup, i.e. how it traverses categorical boundaries present in its industry. Particularly the concept of atypicality might be crucial. According to Goldberg et al. (2016), atypicality refers to the defiance of conventional categorical boundaries. Startups that blend elements from different genres or markets create distinctive identities, diverging from industry norms and established categories. In the context of innovation, atypicality reflects the degree to which a startup explores novel combinations of ideas and practices, thereby distinguishing itself from the mainstream and established norms (Kovacs and Hannan, 2015; Wagner et al., 2019; Abbasiharofteh et al., 2023; Uzzi et al., 2013).

Given the higher levels of uncertainty but also potential associated with atypical ventures, these startups are likely to be differently evaluated than their more typical counterparts. On the one hand, atypical startups may suffer from liability of newness more intensely (e.g. Amason et al., 2006; Navis and Glynn, 2011). Atypicality may also signal creativity, yet experimental evidence by Mueller et al. (2012) demonstrates that it can elicit a negative bias, mediated by uncertainty. Although creativity is often seen as a desirable trait, paradoxically, creative ideas are frequently rejected. On the other hand, journalists, in their perpetual pursuit of exceptional stories, may value the boldness of innovation and attribute a startup's failure to external factors rather than to the venture itself, resulting in more lenient media coverage.

We expect that the evaluation of a failure event will differ based on the level of atypicality a startup exhibits and see reasons for either direction. This leads us to formulate a competing hypothesis.

Hypothesis 4. Following failure, the change in media coverage of atypical startups is (a) more negative or (b) less negative.

4.2.6 Audience

Scholars have stressed the importance of a more nuanced perspective on the complexities of the legitimation process (Fisher, 2020), the largely ignored societal perspective in the formation of social evaluations (Radu-Lefebvre et al., 2019), and in the context of entrepreneurial failure, that perceived severity cannot be aggregated across different contexts (Lee et al., 2007). Shepherd and Patzelt (2015) demonstrate that the evaluation of entrepreneurial failure is influenced not only by the characteristics of the failed startup but also by the attributes of the observer.

Therefore, the question arises how entrepreneurial failure is perceived across different audiences. So far, scholars have mainly stressed the importance of accounting for audience diversity in terms of stakeholder groups (Überbacher, 2014; Fisher, 2020), but leave audience diversity at the societal level unexplored. As societal institutions play a critical role in shaping the perception and evaluation of entrepreneurial failure (Radu-Lefebvre et al., 2019), and given that the media represents one of the most influential institutional forces (Adoni and Mane, 1984; Bitektine and Haack, 2015), variations in evaluations are likely to arise when these institutions are embedded in distinct social and cultural contexts, each governed by unique sets of norms and heuristics (Fisher et al., 2017). Thus, entrepreneurs and startups encounter diverse audiences, with the media functioning as a key institutional actor that interprets and disseminates perceptions of entrepreneurial activity.

Intra-National Evaluation

Entrepreneurship and its processes being socially and contextually embedded calls for the inclusion of cultural factors. Culture, encompassing shared value systems that reflect societal orientations, desirable goals, and aspired end-states (Kleinhempel et al., 2022), as well as "the interactive aggregate of common characteristics that influence a human group's response to its environment" (Hofstede, 1980), serves as a key factor distinguishing nations. It is frequently used to explain variations in entrepreneurial activity that cannot be accounted for by institutions or levels of economic development.

A large stream of literature has dealt with the relationship between culture and entrepreneurship (Stephan, 2022; Hayton et al., 2002) and finds, for example, that entrepreneurship thrives in a friendly, cooperative, and supportive societal climate by encouraging experimentation and increasing legitimacy (Stephan and Uhlaner, 2010). Several studies have focused on measures of culture (Hofstede, 2001; Schwartz, 2006; House and Global Leadership and Organizational Behavior Effectiveness Research Program, 2004) to detect associations and find that particularly individualism, but also certain types of collectivism, such as patriotism and nationalism, positively affect entrepreneurship in a country (Taylor and Wilson, 2012). Similar results with regard to individualism were reported by Shane (1993) earlier and by McGrath et al. (1992), who find that entrepreneurs score higher on power distance, individualism, and masculinity, and lower on uncertainty avoidance.

Drawing on studies that media is an influential factor in culture and social behavior (McQuail, 1985; McDonald, 2004) and that values and culture matter for entrepreneurship (Shane, 1992; Tiessen, 1997; Thomas and Mueller, 2000; Stephan and Uhlaner, 2010), Hindle and Klyver (2007) reveal a significant positive association between the volume of entrepreneurship media stories and a nation's volume of people running a young business[3], but do not find association with either opportunity searching or actual startup activity. However, they document a positive association with participation in opportunity- but not necessity-based entrepreneurship. In a two-country study similar to ours, Jacobi (2018) demonstrates startup coverage to be more positive in the United States than in Germany.

The majority of studies rely on survey data to explore national culture, which raises concerns as entrepreneurship, economic development, formal institutions, and culture are deeply intertwined, and therefore, isolating the effect of culture is challenging. Kleinhempel et al. (2022) overcome this challenge by studying self-employment choices of second-generation immigrants in the US and European countries and find evidence for intergenerational transmission of cultural imprints influencing the likelihood of becoming an entrepreneur.

Concerning entrepreneurial failure, instead of being treated as personal failures, mistakes are treated as an element of experimentation in socially supportive cultures (Stephan

[3] 3-42 months old by definition of the Global Entrepreneurship Monitor (Bosma, 2018)

and Uhlaner, 2010). McGrath (1999) and Cardon et al. (2011) suggest that in individualist societies, business failures may be forgiven, while in collectivist societies, the adverse effects are more severe.

Hypothesis 5a. Following failure, the change in media coverage of domestic startups varies by country.

Cross-National Evaluation

Although for startups, the main market focus is typically at home, at least at an early stage and for well-developed countries that offer attractive markets, the question, of whether and how the previous theorizing extends to cross-national coverage holds relevance as well. At some point, and for some startups even at the beginning, plans of scaling their operations and entering markets abroad arise.

An important concept for cross-group relationships established in social sciences is in-group bias, which, when applied to nations, is also called ethnocentrism, and was originally formulated by Sumner (2002) in 1906. By a strict definition, it states the belief in one's own group superiority and negative evaluation of other groups, but we will rely on a broader definition by Raden (2003) who defines a general in-group bias as "giving ratings to one's own group's that are higher than those given to another" and use the term in-group favoritism interchangeably when referring to a positive in-group bias, which would translate into a more positive sentiment towards domestic startups.

Fischer and Derham (2016) find proof for cultural differences in in-group favoritism and for the uncertainty avoidance theory proposed by Hogg (2000) which states that societies with high levels of uncertainty avoidance display larger in-group favoritism in order to reduce uncertainty. However, across ten nations, Cuddy et al. (2009) show that more collectivist cultures exhibit weaker levels of in-group bias and Ma-Kellams et al. (2011) find in-group derogation, a phenomenon counterintuitive to the rather established in-group favoritism, to be more present among East-Asian than North American countries, indicating differences between collectivist (East-Asian) and individualist (North American) cultural groups. This is in line with Begley (2001), who generalize that the level of shame from business failure is more pronounced in collectivist societies and in-group derogation is more common among minority groups than in dominant groups.

Specifically in the context of media reporting on companies, Gurun and Butler (2012) demonstrate that local media tend to use fewer negative terms when covering local firms compared to those based elsewhere, although this positive bias is partly attributed to the advertising expenditures that local companies invest in local media outlets. Thus,

Hypothesis 5b. Following failure, the change in media coverage of startups is less negative for domestic than for foreign startups.

4.2.7 Stigmatization

Failure is generally viewed negatively due to its immediate and visible impacts on stakeholders (Shepherd and Patzelt, 2017), leading to blame (Semadeni et al., 2008), shame (Wiesenfeld et al., 2008), and "embarrassment, anger, and loss of self-esteem" (Sutton and Callahan, 1987). The literature refers to these social costs as "stigma" (e.g. Sutton and Callahan, 1987; Cardon et al., 2011; Semadeni et al., 2008). For example, Cardon et al. (2011) find that 38% of failure accounts in major U.S. newspapers created a stigma around failed entrepreneurs and Sutton and Callahan (1987) note that many stakeholders involved with bankrupt firms themselves use the word stigma proactively in interviews. Stigma reduces a person from being "whole" to being "tainted" (Goffman, 2022), with "conduct stigma" specifically disapproving behavior that deviates from societal norms. Failing entrepreneurs risk losing status, legitimacy, and face negative impacts on relationships (Cope, 2011; D'Aveni, 1990; Sutton and Callahan, 1987).

Prevention

A strategy to counteract this stigma created by news coverage is referred to as impression management. It describes the deliberate actions taken by individuals and organizations to maintain or restore legitimacy in the face of negative attributions, such as entrepreneurial or organizational failure (Shepherd and Haynie, 2011; Sutton and Callahan, 1987). For entrepreneurs, this process is particularly significant because the legitimacy of the organization and the founder are closely intertwined. Failure can result in both reputational damage and personal costs, including stigmatization, denigration, and diminished social standing (Wiesenfeld et al., 2008; Brown, 1997). Entrepreneurs are driven by a self-esteem motive, which explains why both organizational and individual behaviors are geared towards preserving a positive image (Brown, 1997).

According to impression management theory, individuals take various actions to maintain favorable impressions of themselves, often in response to perceived external threats (Shepherd and Haynie, 2011; Elsbach, 1994). These actions are intended to mitigate the negative effects of failure, which can include blame, loss of legitimacy, and social stigmatization. Sutton and Callahan (1987) propose a hierarchy of stigma-management strategies to address this, including concealing unfavorable information, presenting events in a positive light, denying or accepting responsibility, and withdrawing from the public eye. These strategies aim to repair or protect the damaged reputations of both top management and the firm itself.

On an individual level, impression management techniques can include providing verbal explanations to defend, excuse, justify, or enhance behaviors, for example using press releases, following events that threaten legitimacy (Elsbach, 1994; Tedeschi, 1981). Additionally, organizations act strategically to defend their public image, often diluting negative media attention and cultivating positive perceptions among key stakeholders (McDonnell and King, 2013; Oliver, 1991). Ultimately, impression management plays a critical role in entrepreneurial failure by influencing how both the entrepreneur and the venture are perceived, thus shaping the aftermath of the failure. Thus,

Hypothesis 6. By releasing positive information, failed entrepreneurs positively affect media coverage of their startups.

Entrepreneurial Re-Entry

Both individual (internal) and environmental (external) factors contribute to venture failure (Cardon et al., 2011), yet media attribution often focuses on blaming the founders, partly due to the inability to accurately distinguish between internal and external causes. Whether this responsibility is real or only perceived hardly matters for the resulting social costs. Founders judged as responsible face social sanctions (McGrath, 1999), as they are assumed to have inside information and should have anticipated the failure (Semadeni et al., 2008). In fact, a founder can be stigmatized purely by association with the business failure (Kulik et al., 2008).

Stigma can deter entrepreneurs from future ventures (Shepherd and Patzelt, 2017), and the social costs of failure may reduce overall entrepreneurial activity and national economic value (McGrath, 1999). In addition to emotional consequences, failed entrepreneurs could face reduced access to financial and human resources (Cope, 2011; Shepherd et al., 2011; Sutton and Callahan, 1987), further lowering the likelihood of restarting a venture. However, some research suggests that venture capitalists do not automatically blame failed entrepreneurs, and access to VC funding remains open to them (Cope et al., 2004). Using a natural experiment, Cahn et al. (2021) show that, in fact, when entrepreneurs are not legally required to provide information on their past failure, the probability of starting a new business increases by at least 19%. Therefore, media gains further importance as information intermediary and as its evaluation of a failure amplifies its social costs and significantly influences future entrepreneurial decisions (Shepherd and Patzelt, 2017).

Bases on these findings, we hypothesize

Hypothesis 7. The more negative the media coverage is in case of failure, the less likely the affected founders are to found again.

4.3 Data and Variables

4.3.1 Data

Startups

For our initial startup database, we merged startups from PitchBook[4] and Crunchbase[5] with a valid funding trajectory, resulting in a total of 151,387 startups that were active in the United States (121,038), the United Kingdom (21,390) and in Germany (8,959) in the years 2006 to 2023.[6] Out of those startups, 34,095 startups underwent closure in our sampling period (24,277, 2,940, and 1,242, respectively). Because for all countries, the initial proportion of closed startups contradicts the observation that the majority of startups fail[7], we add failure data by querying the German Handelsregister[8] and OpenCorporates[9] to improve the accuracy of our estimates. Besides funding information, description, and categories, our database contains founder and employee information for some startups. For a better coverage of founder and employee data, we enriched the data with a dataset provided by CoreSignal based on professional profiles of people associated with these startups.

[4] `https://pitchbook.com/`
[5] `https://www.crunchbase.com/`
[6] See Appendix for more details on the merging.
[7] `https://hbr.org/2021/05/why-start-ups-fail`
[8] `https://www.handelsregister.de/` (German public company register)
[9] `https://opencorporates.com/` (large database of company and corporate entity information from various countries)

News articles

During our data collection period, ranging from 1st June 2022 to 31st December 2023, we collected news articles published since 2005 for all startups in our database using LexisNexis as our data source.[10] LexisNexis is a service provider with an extensive volume of news outlets, relying on 4.5m daily added documents and 80K sources[11], among those newspapers from around the world and global online news sources.

After fetching the articles, we performed further local preprocessing to identify the language using the fastText library (Joulin et al., 2016). To avoid false matches and noise in the search results, we filter the articles on filter words specifically extracted from the description of each startup. Based on the number of occurrences of the startup name and the specific filter words, we decide to keep or remove search results. While for English articles, there is more noise due to generic startup names[12], this issue is less common for German text, as generic startup names are typically in English for German startups as well. Following careful manual testing, we limit the English sample to two startup name occurrences and four filter word instances, requiring only one of each in the German sample. Our resulting dataset consists of 1,328,587 unique combinations of startups and articles[13], which cover 58,737 startups, with 45,291 startups being from the United States, 9,204 from the United Kingdom, and 4,242 from Germany, meaning that for more than half of the startups, we did not find any articles.

We have information about the publishing outlet, however, unfortunately, only a fraction is assigned to a country. Therefore, we relied on external lists and manual research to classify the outlets to a country.[14]

4.3.2 Variables

Dependent Variable

Following prior empirical work (Brown and Deegan, 1998; Deephouse, 1996; Pollock and Rindova, 2003; Deephouse and Carter, 2005; Elsbach, 1994; Lamertz and Baum, 1998), our dependent variable measures organizational legitimacy conferred by media through

[10] One year earlier than the database startup activity to include articles leading up to activity.
[11] `https://internationalsales.lexisnexis.com/nexis-daas`
[12] For instance, an UK-based food startup is called "THIS".
[13] An article can cover more than one startup. Therefore, we mention the unique combination of startups and articles.
[14] Sources and approach to identify the country of publication can be found in the appendix.

the sentiment of news articles.

There are several problems related to extracting sentiments from news articles. First, articles might not be entirely dedicated to a single startup but might cover related issues or several startups. Secondly, startups are covered on an infrequent base and the amount of coverage is far from being similar between startups. To increase the accuracy of our results and balance the weight of extensively and less extensively covered startups, we restrict our analysis to sentences that mention a startup by name and aggregate observations at the startup-quarter level. Our resulting sample consists of 413,399 startup-quarters, which we aggregated from 4,636,226 sentences. It should be mentioned that repeated cross-sections are not a flaw of our data collection, but instead a mere reflection of the actual data-generating process and still allow modeling heterogeneity at the startup level.[15]

Because of the nature of our cross-national analysis, we first translate the German articles into English and subsequently extract sentiments to ensure complete comparability of the scores. For both steps, we rely on state-of-the-art transformer models, which have proven to be highly effective and performative across a variety of tasks and benchmarks related to natural language processing. The core of transformer models is the self-attention mechanism, introduced by Vaswani et al. (2017), which improves translation tasks by its improved understanding of contexts and word relationships.

For the translation from German to English, we rely on the model "wmt19-de-en" (Ng et al., 2019), developed by the Facebook AI research group, which, from 2008 to 2021, ranks first 12 times in the news translation task "DEU-ENG" (German $\rightarrow$ English)[16]. This news translation task is carried out every year at the Conference on Machine Translation (WMT), the main event for machine translation research. In the year of contribution, the model even outperforms human translation (Barrault et al., 2019).

After translating German to English text, we extract a sentiment for each of the sentences using the RoBERTa model published on Huggingface[17] by the Research Group in Natural Langauge Processing at Cardiff University[18], which is accessible through the popular transformers library (Wolf et al., 2020). It is based on a RoBERTa model described by Rozado et al. (2022) that is trained on 124m tweets from January 2018 to December 2021

[15] Startups are covered on irregular basis and therefore, compared to a perfect panel, 90% are "missing".
[16] https://opus.nlpl.eu/leaderboard/index.php?model=top&model1=unknown&model2=unknown&test=all&scoreslang=all&src=deu&trg=eng
[17] https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest
[18] https://www.cardiff.ac.uk/research/explore/research-units/natural-language-processing

and it is fine-tuned for sentiment analysis with the TweetEval benchmark (Barbieri et al., 2020). We acknowledge that tweets are different from sentences in news articles. Due to their brevity, tweets are more concise and opinion-driven, with creators often employing sensationalism to increase their audience size. On the opposite, news articles typically provide more in-depth coverage, context, and analysis and adhere to editorial standards. However, despite differing writing styles, tweets and startup news sentences share similar sizes and both convey diverse information. While tweets generally have a more extreme sentiment range, our model effectively distinguishes negative, neutral, and positive sentiments in news sentences.

We applied the model to each sentence, resulting in a vector of probabilities $\mathbf{p} = [p_{\text{pos}}, ; p_{\text{neu}}, ; p_{\text{neg}}]$, which indicates the likelihood of the sentence expressing positive, neutral, or negative sentiment and sums to 1. By multiplying these probabilities with commonly used values for respective sentiments (pos $= 1$, neu $= 0$, neg $= -1$), we can compute a sentiment score for each sentence.

$$sentiment = p_{\text{pos}} * 1 + p_{\text{neu}} * 0 + p_{\text{neg}} * (-)1$$

Consequently, sentiment scores fall into the range $[-1, 1]$, consistent with the notion that "legitimation is a continuous, not a dichotomous variable (...) Indeed, it runs the full gamut from highly supportive to highly oppositional" (Etzioni, 1987).

To gain deeper insights into the drivers of sentiment change, we also compute the number of positive, neutral, and negative sentences. Each sentence is classified based on the highest probability in its probability vector. Given the highly skewed distribution of sentence counts, we follow standard practice in management research by applying a logarithmic transformation. To avoid infinite values, we add 1 prior to the transformation.

Treatment Variable

The treatment variable that represents failure is 0 for operating startups and switches on ($= 1$) in the quarter a startup undergoes closure and remains so for the subsequent periods. We follow the definition of failure provided in the theory part and consider failure as the termination of a venture that falls short of its goals, based on subjective assessments or unmet economic viability thresholds set by the entrepreneur. This definition is rather broad and we acknowledge that a variety of reasons can be the cause of closing a business. Although closing a business does not necessarily mean failure (Headd, 2003), shutting down a business implies the absence of promised value and cessation of goals pursuit and we observe how this event is covered. In the following analyses, we also address the issue that failure is not a sudden event that can be traced back to a particular date and that the legal date of failure might be different from the de facto date of failure.

Moderators

It is important to note that, for estimating moderation effects, we split the startups into equally sized groups based on their rank within the overall distribution. Thus, it is not the absolute value that matters, but the relative position in the distribution.

For H2, we measure funding as the cumulative amount a startup has raised across its funding rounds. For rounds with no available funding information, we assign a value of zero only if a subsequent round discloses the raised amount. Startups with no available funding data at any stage are excluded from the analysis.

To investigate H3, whether industry hype influences the judgment of a startup's failure, for a focal startup, we define hype as the average sentiment in its industry one quarter prior to the failure. We also test alternative measures, such as news volume, hype within one year prior to failure, or number of funding rounds.

The measure of atypicality for analyzing H4 relies on the description of startups. For each startup we extract terms (noun phrases) that contain information about the business model, purpose and focus. Then, we create the co-occurence matrix, which contains information on how often terms occur together, i.e. *cryptocurrency* appears often with *blockchain*, but very rarely with *footwear*.

To get a refined measure of the semantic similarity, for each pair of terms included in the co-occurrence matrix, we retrieve the Jaccard similarity, which indicates the relatedness of the two terms based on their number of co-occurences in startup descriptions, weighted by the count of total occurrences together or separate of each other.

$$J(i, j) = \frac{|i \cap j|}{|i \cup j|} \tag{4.1}$$

Perfect similarity (=1) occurs when terms always co-occur, while perfect dissimilarity (=0) is when they never appear together. For computational reasons, and because some terms are very rarely used, we include only terms that have appeared at least 50 times.

Atypicality is defined as the combination of terms that do not typically co-occur. Therefore, instead of similarity, we are more interested in the distance between two terms, which can be calculated as an inverse term of their similarity, following Kovacs and Hannan (2015) and Shepard (1987).[19]

$$dist(i,j) = -\frac{ln(J(i,j))}{\gamma} \tag{4.2}$$

Having obtained the pairwise distances, we compute a startup's atypicality as a function of the average pairwise distance between the terms used to describe it. The more terms assigned to a startup, and the greater the distance between them, the more atypical the startup is considered to be. This measure is proportional to both the number of labels and the conceptual distance between them, providing a numeric value for comparing startups' atypicality levels.

Because atypicality is inherently dynamic and categorical boundaries shift over time, we compute atypicality at the treatment dates within the industry of the focal startup. Measuring general atypicality across industries can lead to biased results as a startup can potentially benefit from the industry it operates in when being evaluated (Aldrich and Martinez, 2003; Shepherd and Patzelt, 2015).

To measure diversity regarding the audience and the cultural context in H5-H7, we use the country of publication and startup country. We take a country perspective, as the instiutional and cultural landscape within a country is rather similar. However, other researchers have argued that differences at a finer grade like the regional could also be of relevance (Cardon et al., 2011).

Control variables

We include control variables expected to be correlated with media sentiment to control for variation in legitimacy not caused by failure and we control for variables potentially

[19]Note that the choice of γ does not matter for our study as we only focus on where a particular startup is ranked in the distribution.

correlated with failure to isolate the hypothesized effects as precisely as possible (Cinelli et al., 2024).

The characteristics of the founder team often play a central role in media reports and build cognitive legitimacy (Shepherd, 2003). Particularly for young and nascent startups, founders are inextricably linked to the startup and in case of failure, the founders are the ones facing stigma. We include the number of female founders, the number of founders holding various academic degrees, and the number of founders with prior entrepreneurial experience. We base these variables on prior literature showing the importance of the founder team's human capital (Tzabbar and Margolis, 2017; Hashai and Zahra, 2022).

The funding a startup is able to secure is often seen as one dimension of startup success and a very plausible factor for explaining differences in media sentiment. Funding rounds themselves constitute relevant news. Therefore, we control for both the number of successfully completed funding rounds and the logarithm of the total funding raised in US dollars. The funding amount is measured cumulatively up to the quarter in which the news article appears. Similarly, funding rounds are captured through a set of dummy variables, where the most recently completed round takes the value 1 and all others are set to 0.

Additionally, we account for whether a startup has secured funding from a venture capital firm as this backing provides an important signal about the quality of the startup (Davila et al., 2003; Megginson and Weiss, 1991; Pollock et al., 2010) and could be seen as a validation of the startup's potential and add to the legitimacy. It could also be related to the professionalization of a startup with regard to issuing press releases.

We also control for the volume of media exposure and startup age, which, according to Schultz et al. (2001), are often related to reputation.

Fixed Effects

We include both main effects for the publication quarter and publication country and the interaction of both components. This way, we make sure to account for any unobserved shocks that are specific to a publication quarter, a publication country, or a publication country for a specific quarter.

Lastly, we include fixed effects for the industry a startup operates in due to the discussed industry spillover effects with regard to legitimacy. In the absence of formal industry classifications in our dataset, we employ a K-Means clustering approach based on the same terms we compute atypicality on. This method ensures semantic similarity among startups in the same cluster. Using the elbow method, we determined that 15 clusters, comparable to the SIC or NAICS 1-digit industry classifications, offer a reasonable partitioning of the data.

First, we create a document-term matrix, in this case, a startup-term matrix, where all startups are rows and all terms used to describe them are columns. The values are either 0 if a startup is not described by a particular term or 1 if it is. Using the FAISS algorithm (Douze et al., 2024), a library with highly efficient similarity search and clustering algorithms, we then perform K-Means clustering. Iteratively, each startup is assigned to its nearest centroid, which is a mean of all startups it contains and is updated in each iteration. Mathematically, the vector x_i (startup) is assigned to its nearest centroid C_k based on the $||.||$ L2 Euclidean distance. Both are of same dimension.

$$k = argmin_k ||C_k - x_i|| \tag{4.3}$$

We set the number of iterations to 20 and repeat the process five times to avoid sensitivity to the initial guess of parameters.[20] Based on the objective function, which indicates clustering quality via the total sum of squares, we select the best clustering outcome.

4.4 Estimation Strategy

4.4.1 Identification

The ideal experiment to isolate the media's reaction to startup failure would involve observing the same startup over time in two counterfactual states: one in which it continues operating and one in which it fails, with all other factors held constant. Any differences in media coverage between these two states could then be directly attributed to the effect of failure. However, such an experiment is impossible to implement.

The next best alternative, often regarded as the gold standard of causal inference, would be a randomized experiment in which startups are randomly assigned to a treatment group (failure) or a control group (continuing operations). While theoretically feasible, such an

[20] As suggested by the rule of thumb here: `https://github.com/facebookresearch/faiss/wiki/FAQ` [accessed: 2024-12-20]

experiment is practically unworkable and far too expensive considering the personal, social, and financial investments at stake. Not only would it require startups to fail artificially, but it would also necessitate keeping the media uninformed about the experimental setup, an implausible requirement given the transparency inherent in media reporting. Even in scenarios where exogenous shocks might cause some startups to fail while others do not, it is highly likely that the media would incorporate the shock itself into their coverage, further complicating the isolation of the treatment effect.

Therefore, the most plausible and feasible approach is to adopt a quasi-experimental design, a research strategy that addresses the absence of random assignment into treatment and control groups. While quasi-experiments are a powerful tool, their most significant limitation compared to randomized experiments is the inability to completely rule out the influence of confounding factors. Measures such as incorporating variables related to both the treatment and the outcome can mitigate this limitation, but only when these variables are observed.

In our context, we expect failing startups and operating startups to differ in terms of their underlying quality. To account for these quality differences, we include control variables in the regression that are widely recognized as predictors of quality. These include the qualifications of the founder team, the stage of the startup, and the amount of funding it has secured (see 4.3.2 for more details). These controls aim to reduce bias and improve the robustness of the estimated treatment effects.

4.4.2 Estimation

We observe media coverage before and after failure for failed startups and coverage throughout the sample period for operating startups. Traditionally, this setting would have been analyzed with a DiD design using a two-way fixed effects estimator (TWFE), comparing changes in outcomes of units that are affected by the treatment to changes in units that are unaffected. However, recent econometric literature has shown that when treatment timing is staggered, i.e. units adopt the treatment at different points in time, TWFE can yield severely biased estimates when there is treatment effect heterogeneity across treatment times and across units (Goodman-Bacon, 2021; Borusyak et al., 2024; de Chaisemartin and D'Haultfœuille, 2023; Sun and Abraham, 2020; Wooldridge, 2021). More precisely, the bias is caused by "forbidden comparisons" (Borusyak et al., 2024) between late-treated units and early-treated units in settings with staggered adoption.

Because excluding treatment heterogeneity is a very strict assumption and, in our application, we could imagine that, for example, media sentiment after failure recovers after some time, we rely on the imputation-based estimator proposed by Borusyak et al. (2024), which is robust to the described issue. The estimator is capable of handling repeated cross-sectional data like ours, and generally, DiD designs work with repeated cross-sections as long as there is no change in the compositional settings of the cross-sections (Sant'Anna and Zhao, 2020), which we see no reason to suspect since, for each quarter, we have observations from a very similar pool of companies. Moreover, we only include startups which we observe at least once before and after treatment.

There has been a variety of proposed estimators robust to treatment effect heterogeneity for staggered treatment adoption (see e.g. Callaway and Sant'Anna, 2021; Sun and Abraham, 2020; Borusyak et al., 2024) which can be broadly split into two categories: imputation-based estimators and 2x2 aggregation. Because, for our application, imputation-based estimators most adequately exploit the richness of the large number of untreated observations and extensive numbers of covariates with potential explanatory power, we choose the estimator proposed by Borusyak et al. (2024). Moreover, the estimator can be applied to our triple-difference design for the analysis of moderation effects.

In the following paragraphs, we briefly describe the imputation based approach and the assumptions it relies on. In short, the first stage regresses outcomes for untreated observations on fixed effects and covariates, and the second stage subtracts estimated counterfactual outcomes from observed outcomes to obtain a residualized outcome, which is then regressed on treatment weights. For our computations, we build on the R package *didimputation*.[21]

First stage. Following Borusyak et al. (2024), using the subsample of untreated observations $D_{itp} = 0$, we estimate a first stage that, similar to the ordinary least square regression for the base level of sentiment, regresses the untreated observed outcome on the described startup-specific covariates X_{itp} and fixed effects for publication country α_p, publication quarter λ_t, publication country-quarter η_{itp}, and industry ν_i, where p is publication country, t is publication quarter and i startup. Note that the startup country is included in

[21] `https://github.com/kylebutts/didimputation`

the startup-specific covariates and ϵ_{itp} is the error term.

$$Y_{itp}(0) = \alpha_p + \lambda_t + \eta_{pt} + \nu_i + \beta X'_{itp} + \epsilon_{itp} \tag{4.4}$$

Particular attention must be given to the inclusion of time-varying covariates in DiD research designs (Huntington-Klein, 2023). Although differences in the levels of covariates or outcome are not necessarily an issue in DiD settings and for the assumption of parallel trends, covariates are generally included to increase the plausibility of the inherently untestable (conditional) parallel trends assumption. Unfortunately, there are many ways to include "bad controls". One obvious way of including bad controls is through the post-treatment bias (Acharya et al., 2016; Montgomery et al., 2018), which arises when controlling for a variable that itself is affected by treatment (e.g. funding variables could block the treatment effect). Therefore, we follow recent literature (Sant'Anna and Zhao, 2020; Callaway and Sant'Anna, 2021; Abadie, 2005) and the recommendation by Gardner (2022) to include time-varying or time-fixed covariates in the first stage and adjust outcomes for the second stage by using the latest pre-treatment realizations.[22]

Second stage. Using the estimated coefficients and fixed effects from the first stage, we impute $\widehat{Y}_{itp}(0)$, the counterfactual outcome for the treated units $D_{itp} = 1$, to subtract it from the observed outcome and obtain the adjusted outcome $\tau_{itp} = Y_{itp} - \widehat{Y}_{itp}(0)$ which contains what cannot be explained by covariates and fixed effects. Now, we can specify weights ω for any estimation target, such as the overall or any conditional ATT.

$$\tau_w = \sum_{itp \in D_{itp}=1} \omega_{it} \tau_{itp} \tag{4.5}$$

Assumptions The imputation-based approach relies on the following assumptions to be satisfied:

(1) (Conditional) parallel trends assumption: in absence of treatment, treated startups would follow a similar trend as untreated startups (conditional on covariates).

[22] We run robustness test without fixing the variables and come to the same result.

(2) No anticipation: the failure of a startup is not anticipated prior to the failure.

(3) Correct specification of the first stage: the parametric model for $Y(0)$ needs to be correctly specified to impute the counterfactual outcomes.

One advantage of the estimator is that we can estimate an event study, which on the one hand, serves the purpose of checking the plausibility of the parallel trends assumption (1) by performing placebo tests prior to the treatment and on the other hand, estimates the evolution of treatment effects over time. Moreover, the event study helps identifying anticipatory effects (2) that can be accounted for by setting the treatment period to a quarter prior to the actual treatment. Regarding (3), with our richness in explaining variables we are confident to specify and estimate the correct model.

4.5 Results

Table 4.1 presents descriptive statistics for our aggregated startup-quarters across the treatment (closed) and control (operating) groups. The descriptive summaries are informative on their own and highlight differences across both the treatment and time dimensions. For most variables, both groups appear broadly similar, though quality-related observations in the control group tend to be slightly higher, possibly reflecting more developed and promising startups, on average. Moreover, the outcome variables appear to change following the treatment.

Importantly, the DiD design does not require the treatment and control groups to be comparable or similar before the treatment but instead assumes that they follow the same trend over time. Although this assumption is inherently untestable, the plausibility of parallel trends is often enhanced by incorporating covariates to create a balance between the groups. For that method to be successful, there has to be distributional overlap between the groups, which the statistics confirm.

Table 4.1 shows that the majority of startup-quarters are from the United States and that, in all countries, most reporting occurs while the startup is still operational. The table also indicates that attention to failed companies is slightly higher in the United States and the United Kingdom compared to Germany.

Examining the outcome plots in Figure 4.1, we see that the levels of treatment and control groups are similar. Additionally, at first glance, as expected, the reported sentiment

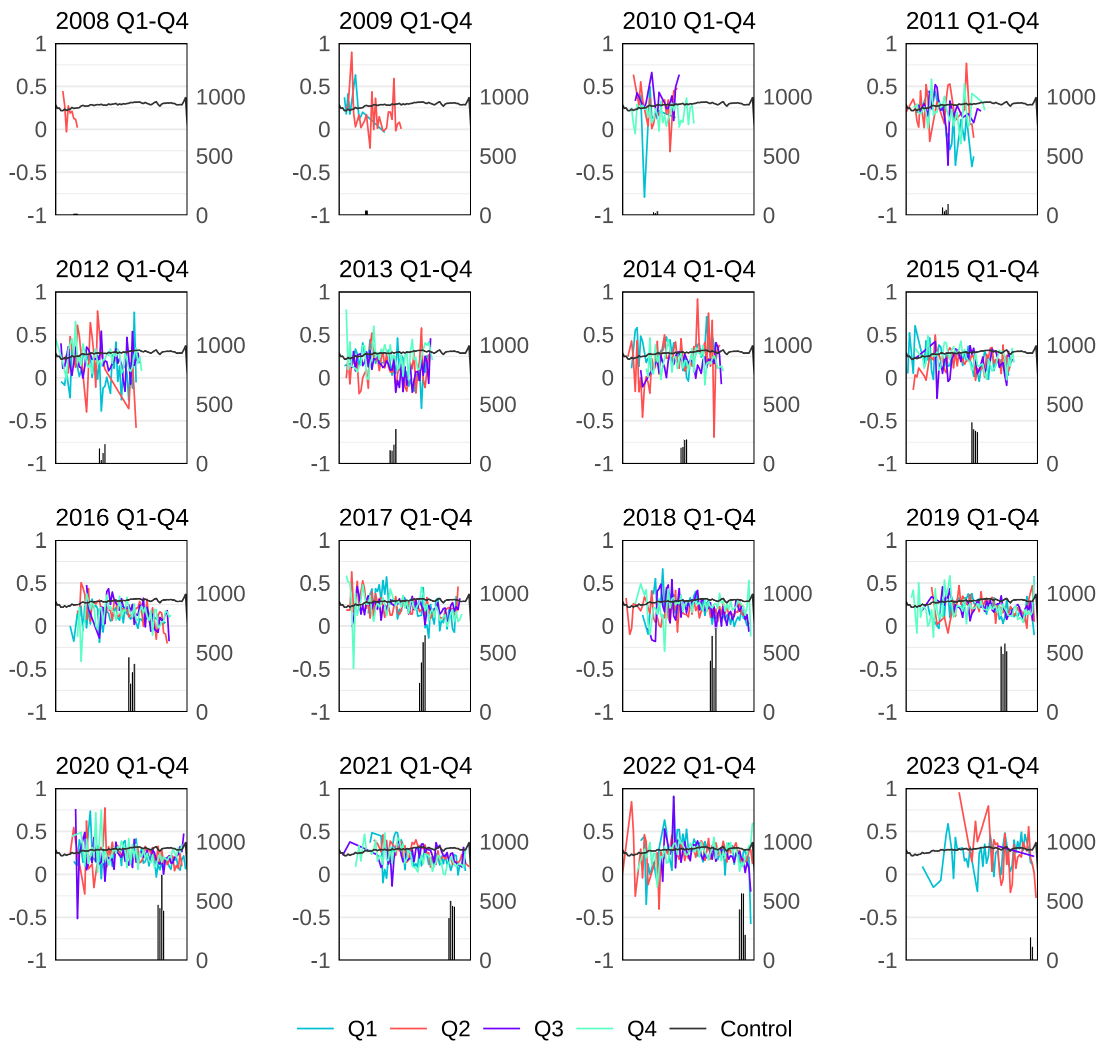
TABLE 4.1: Descriptive Statistics of Startup-Quarters

	Operating				Closed			
	Pre (N=269,595)		Post (N=0)		Pre (N=16,799)		Post (N=5,998)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
DV								
Average Sentiment	0.31	0.25	-	-	0.27	0.26	0.16	0.31
Num. of Neutral Sentences	7.52	28.71	-	-	6.91	22.23	10.72	38.48
Num. of Negative Sentences	0.59	6.72	-	-	0.76	8.19	1.37	5.42
Num. of Positive Sentences	3.11	10.80	-	-	2.44	5.72	3.26	14.61
IV: General								
Number of Articles	3.21	14.59	-	-	2.76	7.63	4.23	14.17
Age in Quarters	5.77	3.22	-	-	4.14	2.57	7.07	2.75
Country (Media : Startup)	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
USA : USA	136,984	50.8	-	-	9,232	55.0	2,809	46.8
USA : GBR	11,289	4.2	-	-	739	4.4	302	5.0
USA : DEU	3,194	1.2	-	-	139	0.8	97	1.6
GBR : USA	67,721	25.1	-	-	3,474	20.7	1,728	28.8
GBR : GBR	22,692	8.4	-	-	1,516	9.0	330	5.5
GBR : DEU	2,953	1.1	-	-	106	0.6	49	0.8
DEU : USA	12,453	4.6	-	-	922	5.5	511	8.5
DEU : GBR	2,683	1.0	-	-	133	0.8	59	1.0
DEU : DEU	9,626	3.6	-	-	538	3.2	113	1.9
IV: People								
Num. of Founder	4.33	12.15	-	-	3.48	2.04	3.18	1.71
Num. of Female	0.63	9.82	-	-	0.39	0.79	0.38	0.75
Num. of Experienced	1.07	2.79	-	-	0.73	2.09	0.61	1.70
Num. of PhD	0.84	1.98	-	-	0.61	1.34	0.31	0.82
Num. of Master	2.64	6.94	-	-	1.99	2.30	1.49	1.95
Num. of Bachelor	3.24	7.14	-	-	2.67	2.24	2.35	2.03
IV: Funding								
Raised Amount	61.7M	306.4M	-	-	25.7M	119.4M	31.4M	184.1M
VC Entry	0.91	0.29	-	-	0.84	0.36	0.80	0.40
Round Type	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Seed	49,471	18.4	-	-	5,036	30.0	1,497	25.0
Series A	46,660	17.3	-	-	3,715	22.1	1,559	26.0
Series B	39,739	14.7	-	-	2,549	15.2	1,023	17.1
Series C	31,583	11.7	-	-	1,730	10.3	660	11.0
Series D	23,999	8.9	-	-	1,054	6.3	401	6.7
Series E	17,808	6.6	-	-	768	4.6	230	3.8
Series F	12,740	4.7	-	-	469	2.8	221	3.7
Series G	8,671	3.2	-	-	354	2.1	93	1.6
Series H	5,575	2.1	-	-	227	1.4	77	1.3
Series I	10,474	3.9	-	-	196	1.2	95	1.6
Other	22,875	8.5	-	-	701	4.2	142	2.4

Notes. This table presents summary statistics for repeated cross-sectional startup-quarters. Startups are included only if they have coverage during the specified period, as sentiment cannot be computed otherwise. Industries are omitted for simplicity, and the raised amount is presented in levels instead of logarithmic values. For the countries, USA refers to United States, GBR to United Kingdom, and DEU to Germany as defined by ISO 3166-1 alpha-3.

drops after failure. Most of the failed startups in our sample fail between 2016 and 2022, while there are relatively few startups in the cohorts prior to 2012.

FIGURE 4.1: Observed News Sentiment across Treatment Cohorts



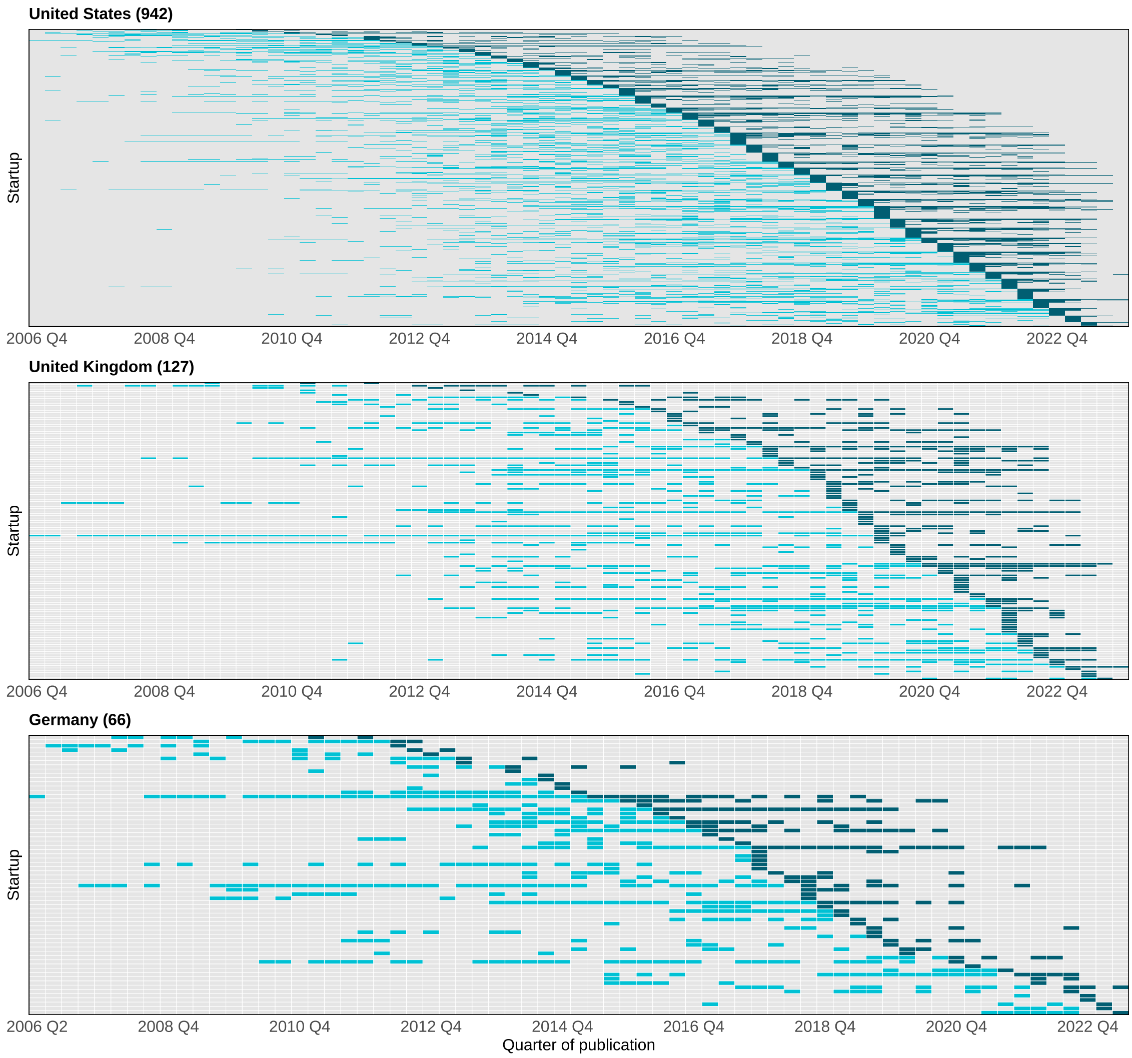
Notes. This figure shows the mean sentiment values observed for each cohort, where cohorts are defined as startups that failed in the same quarter. For clarity, cohorts are displayed across multiple panels, with plot titles and legends indicating the respective cohort groups. In all panels, the control group is depicted by a grey line. The relative number of startups in each cohort is shown as black bar plots accompanying the sentiment lines and references on the right x-axis.

In more detail, Figure 4.2 illustrates the treatment adoption and the composition of the sample. Unlike large public firms, startups are rarely covered both before and after failure, with media coverage gradually fading out in the periods following failure.

4.5.1 Response to Failure

First, we run an event study to check the validity of our conditional parallel trends assumption. It is inherently untestable due to the fundamental problem of causal inference,

FIGURE 4.2: Staggered Treatment Adoption and Coverage

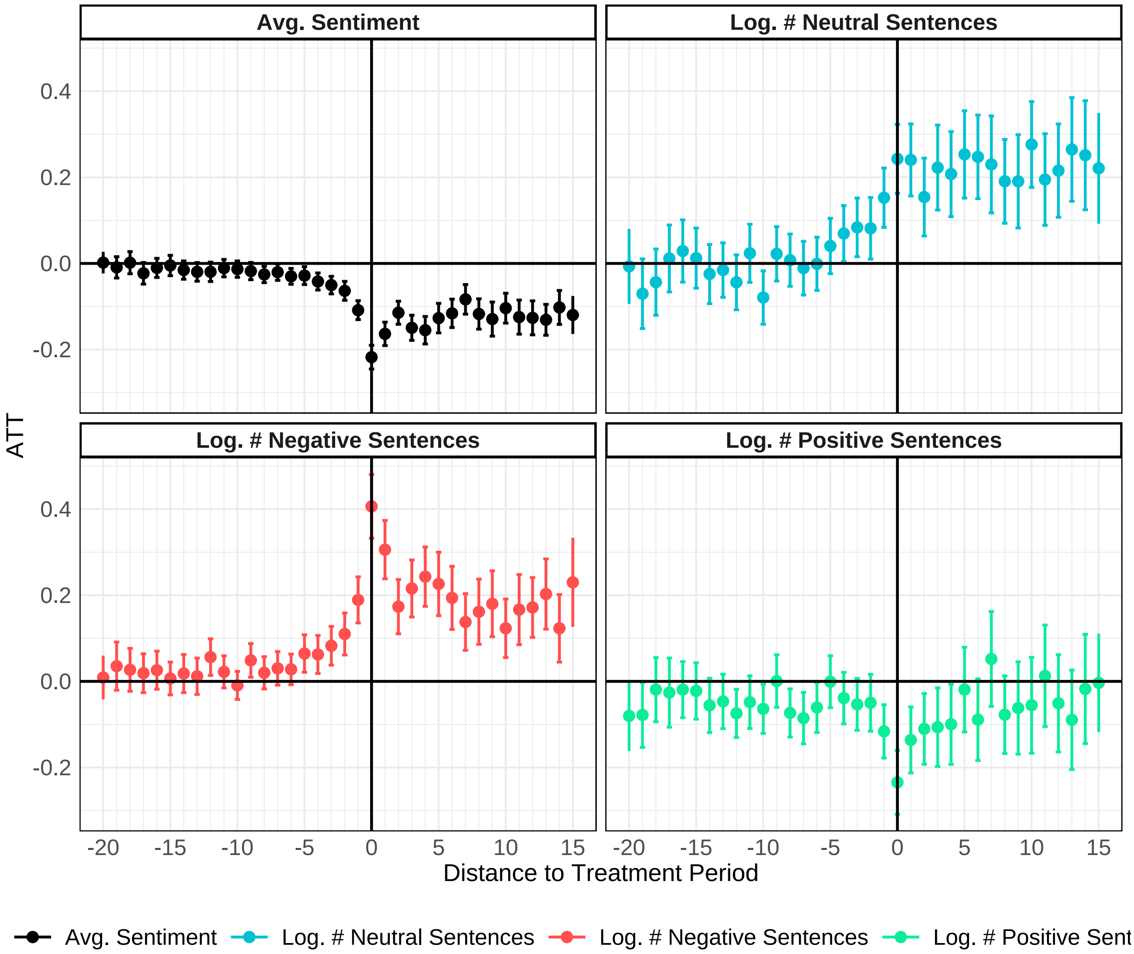


Notes. This figure illustrates the treatment adoption of failed startups and the presence of domestic news coverage, categorized by country. A light blue bar indicates news coverage before failure, a dark blue bar indicates coverage after failure, and a grey bar represents quarters with no observed coverage. The total number of unique startups in each country is shown in parentheses. When we assume anticipation, the number of treated startups increases to 1145, 164, and 83 for the United States, the United Kingdom and Germany, respectively.

as we only observe one state for each startup at a time, operating or failed. However, a common way to check whether the assumption of parallel trends is likely to hold is to perform placebo tests and check for treatment effects before the treatment period. Treatment effects before the treatment period should be indistinguishable from zero and consequently, treatment and control group should follow a similar trend prior to treatment.

We graph 20 lead and 15 lags in a dynamic specification of the imputation-based DiD approach in Figure 4.3. Until a few periods prior to the treatment, the placebo treatment effects are indistinguishable from zero but take values different from zero already prior to the treatment. As these effects have the same sign as the hypothesized and observed treatment effects but are smaller, we argue that these are anticipatory effects and startups receive negative coverage already before their definite closure. In fact, it is quite common that prior to failure, a startup struggles, which is picked up by the media and reported. From an analysis point of view, anticipatory effects can be accounted for by shifting the treatment period to an earlier period, which we incorporate in subsequent analyses by setting anticipation to one quarter before failure.

FIGURE 4.3: Event Study of Media Coverage and Entrepreneurial Failure



Notes. This figure presents estimated treatment effects (placebo and actual) with 95% confidence intervals over a period of 20 quarters before and 15 quarters after the treatment. The actual closure date, denoted as $t = 0$, is marked by a black vertical line.

While the changes around the treatment date are already visible in the event study, we test H1 more formally using a static specification of the imputation-based DiD approach in Table 4.2.[23] For our main outcome, average sentiment, we observe a significant negative

[23]In Table 4.2, we show only the second stage of the imputation-based DiD approach, and the control variables, such as funding and founders, are included in the first stage. Due to the residualized outcome, we can leave out the intercept.

TABLE 4.2: Difference-in-Differences Estimation of Entrepreneurial Failure on Media Coverage

	Avg. Sentiment	Log. # Neutral Sentences	Log. # Negative Sentences	Log. # Positive Sentences
Failure (*Anticipation* = 0,	−0.13 ***	0.23 ***	0.21 ***	−0.08 **
Num. Obs = 282,259)	(0.01)	(0.03)	(0.02)	(0.03)
Failure (*Anticipation* = 1,	−0.13 ***	0.22 ***	0.21 ***	−0.08 **
Num. Obs = 282,178)	(0.01)	(0.03)	(0.02)	(0.03)
Failure (*Anticipation* = 2,	−0.12 ***	0.20 ***	0.19 ***	−0.08 **
Num. Obs = 282,061)	(0.01)	(0.03)	(0.02)	(0.03)
Mean DV	0.294	1.488	0.191	0.88
Media Country X Quarter FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Series FE	Yes	Yes	Yes	Yes
Controls in First Stage	Yes (Funding + Founders)			

Notes. This table presents results from three separate models, accounting for 0, 1, and 2 quarters of anticipation, respectively. The models include main effects for publication quarter and publication country, their interaction, and industry fixed effects. Standard errors, clustered at the company level, are reported in parentheses. Control variables are included in the first stage. Significance levels are indicated as follows: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

shift of $[-.13, -.12]$[24] sentiment units, which relates to about half a standard deviation. From the coefficients of the other outcomes, we find that, on average, the proportion of neutral $[.2, .23]$ and negative sentences $[.19, .21]$ increases following failure, while the proportion of positive sentences decreases $[-.08]$. Accounting for the identified anticipatory effects has minimal impact on the magnitude of these effects.

In summary, across all outcomes and specifications regarding anticipatory effects, we find robust differences in the tonality of media coverage in response to failure and confirm H1.

4.5.2 Contextual Factors

For hypotheses H2-H4, we use the same underlying approach, but instead of estimating the average treatment effect, we compute heterogeneous treatment effects based on the hypothesized moderators. Unlike the classical TWFE regression, our method does not allow for the estimation of coefficients or slopes for continuous variables under the assumption of a linear effect. As an alternative, we divide the treated sample of startups into four equally sized groups based on their moderator values. We split the distribution at its quartiles ($q_{25\%}$, $q_{50\%}$, and $q_{75\%}$) to form four groups.[25] For example, the second group includes startups with moderator values greater than the bottom 25% but less than or equal to the top 50%. One advantage of this approach is that it reveals non-linear effects, which would

[24] Values in square brackets represent the minimum and maximum values across specifications.
[25] Note that we split the groups to contain the same number of startups, which does not result in equal sizes by startup-quarter.

be obscured in a ordinary TWFE model.

Funding

H2 predicts that the greater the amount a startup has raised, the more pronounced the shift in media sentiment. As shown in Table 4.3, startups in the lowest funding quartile experience only a $-.07$ reduction in sentiment units, approximately one-third the magnitude of the effect observed for startups in the top quartile ($-.20$). This heterogeneity in media coverage appears to be primarily driven by a reduction in neutral and positive sentences for startups that have raised more than $9.6M, with the most notable trend being the significant decrease in positive sentences. In line with H2, compared to the average treatment effect, there is substantial variation in the effects when conditioned on the funding sum, and we see a clear tendency of startups that "burned" more money to be treated more harshly than those with less funding.

Hype

H3 predicts amplified negative reactions to failure of startups in hyped industries. However, we do not observe heterogeneity for any outcome. Alternative measures of hype, which are not shown here, yield similar null results. Because we do not find support for the claim that startups in hyped industries receive a larger drop in sentiment, we do not confirm H3. However, it follows that even after failure, startups in hyped industries are more positively evaluated than startups from less hyped industries, i.e. the difference persists.

Atypicality

H4 proposes a competing hypothesis on whether atypicality helps or hurts a failed startup in terms of media coverage. Conditioning on how atypical a startup is within its industry, we find a weak tendency for startups that are atypical to be more shielded from negative press. Startups in the top quartile of the atypicality distribution experience a relatively weak drop in average sentiment with $-.09$ units compared to the most typical quartile that is treated more harshly with $-.15$. Thus, we find limited support for H4.[26]

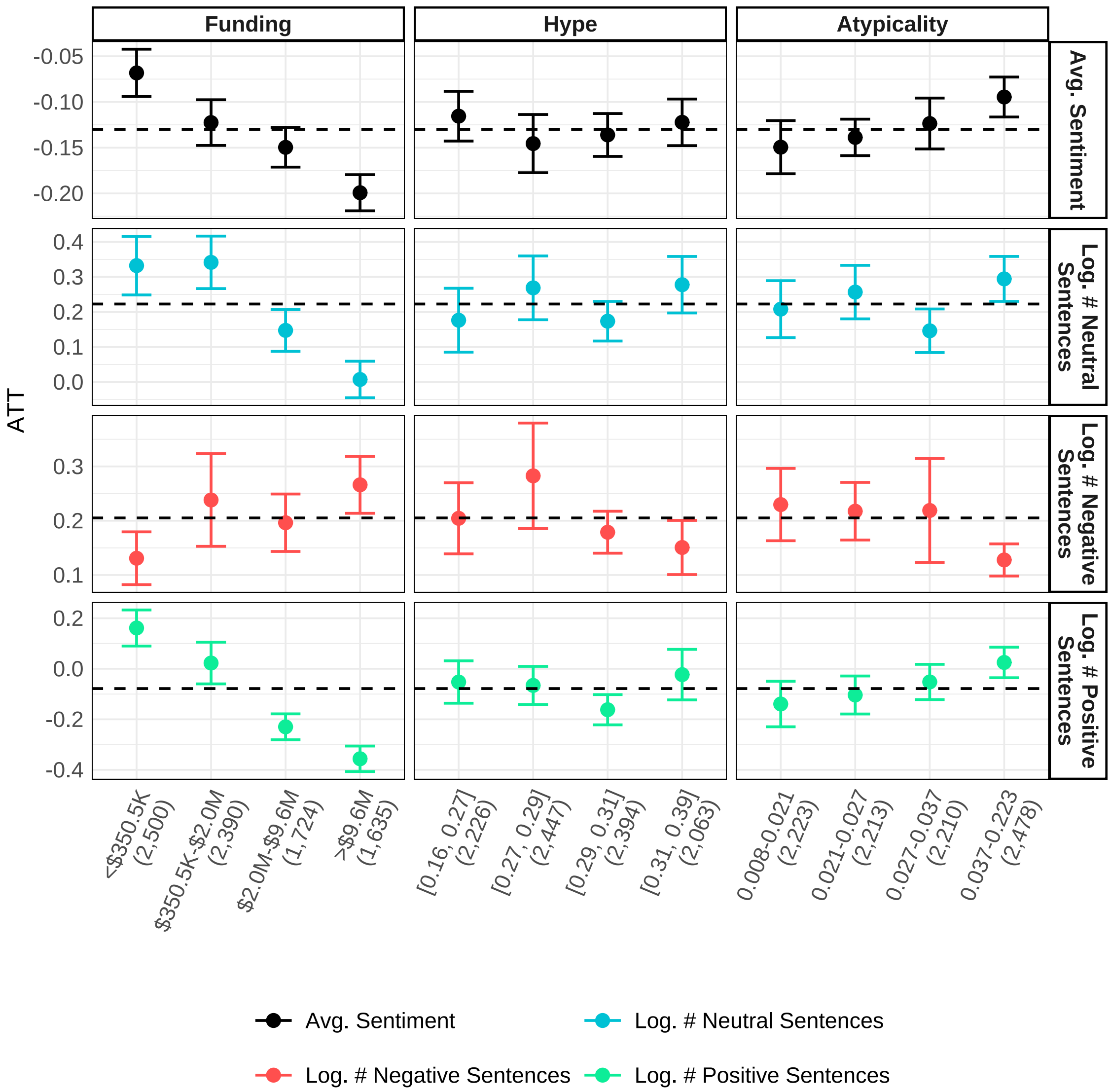
[26] We also run the analysis without restricting atypicality within industry but instead compute a global atypicality measure. The results are very similar and can be found in the appendix.

TABLE 4.3: Heterogeneous Treatment Effects of Entrepreneurial Failure on Media Coverage

	Avg. Sentiment	Log. # Neutral Sentences	Log. # Negative Sentences	Log. # Positive Sentences
Funding $q_{[0\%-25\%]}$	-0.07 ***	0.33 ***	0.13 ***	0.16 ***
	(0.01)	(0.04)	(0.02)	(0.04)
Funding $q_{[25\%-50\%]}$	-0.12 ***	0.34 ***	0.24 ***	0.02
	(0.01)	(0.04)	(0.04)	(0.04)
Funding $q_{[50\%-75\%]}$	-0.15 ***	0.15 ***	0.20 ***	-0.23 ***
	(0.01)	(0.03)	(0.03)	(0.03)
Funding $q_{[75\%-100\%]}$	-0.20 ***	0.01	0.27 ***	-0.36 ***
	(0.01)	(0.03)	(0.03)	(0.03)
Hype $q_{[0\%-25\%]}$	-0.12 ***	0.18 ***	0.20 ***	-0.05
	(0.01)	(0.05)	(0.03)	(0.04)
Hype $q_{[25\%-50\%]}$	-0.15 ***	0.27 ***	0.28 ***	-0.07 .
	(0.02)	(0.05)	(0.05)	(0.04)
Hype $q_{[50\%-75\%]}$	-0.14 ***	0.17 ***	0.18 ***	-0.16 ***
	(0.01)	(0.03)	(0.02)	(0.03)
Hype $q_{[75\%-100\%]}$	-0.12 ***	0.28 ***	0.15 ***	-0.02
	(0.01)	(0.04)	(0.03)	(0.05)
Atypicality $q_{[0\%-25\%]}$	-0.15 ***	0.21 ***	0.23 ***	-0.14 **
	(0.01)	(0.04)	(0.03)	(0.05)
Atypicality $q_{[25\%-50\%]}$	-0.14 ***	0.26 ***	0.22 ***	-0.10 **
	(0.01)	(0.04)	(0.03)	(0.04)
Atypicality $q_{[50\%-75\%]}$	-0.12 ***	0.15 ***	0.22 ***	-0.05
	(0.01)	(0.03)	(0.05)	(0.04)
Atypicality $q_{[75\%-100\%]}$	-0.09 ***	0.29 ***	0.13 ***	0.02
	(0.01)	(0.03)	(0.02)	(0.03)
Media (USA) : Startup (USA)	-0.14 ***	0.21 ***	0.22 ***	-0.12 ***
	(0.01)	(0.03)	(0.02)	(0.03)
Media (USA) : Startup (GBR)	-0.12 ***	0.19 **	0.22 ***	-0.10 *
	(0.01)	(0.06)	(0.03)	(0.04)
Media (USA) : Startup (DEU)	-0.05 ***	-0.14 ***	0.00	-0.20 ***
	(0.01)	(0.03)	(0.02)	(0.02)
Media (GBR) : Startup (USA)	-0.13 ***	0.36 ***	0.21 ***	0.04
	(0.01)	(0.04)	(0.02)	(0.04)
Media (GBR) : Startup (GBR)	-0.14 ***	0.25 ***	0.25 ***	-0.06
	(0.02)	(0.04)	(0.04)	(0.04)
Media (GBR) : Startup (DEU)	-0.14 ***	0.06 **	0.05 ***	-0.32 ***
	(0.01)	(0.02)	(0.02)	(0.02)
Media (DEU) : Startup (USA)	-0.09 ***	0.09	0.12	-0.10 ***
	(0.02)	(0.06)	(0.09)	(0.02)
Media (DEU) : Startup (GBR)	-0.03 ***	-0.12 **	-0.02	-0.10 ***
	(0.01)	(0.04)	(0.02)	(0.03)
Media (DEU) : Startup (DEU)	-0.25 ***	-0.24 ***	0.28 ***	-0.35 ***
	(0.01)	(0.03)	(0.02)	(0.02)
Num.Obs.	282,178	282,178	282,178	282,178
Mean DV	0.294	1.488	0.191	0.88
Media Country X Quarter FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Series FE	Yes	Yes	Yes	Yes
Controls in First Stage	Yes (Funding + Founders)			

Notes. This table presents results from a model with treatment effect interactions 1 quarter of anticipation. Both main effects of publication quarter and publication country and the interaction and industry fixed effects are included. Standard errors are clustered at the company level. Control variables are included in the first stage. Anticipation is set to 1. Stars indicate significance level: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

FIGURE 4.4: Treatment Effect Heterogeneity



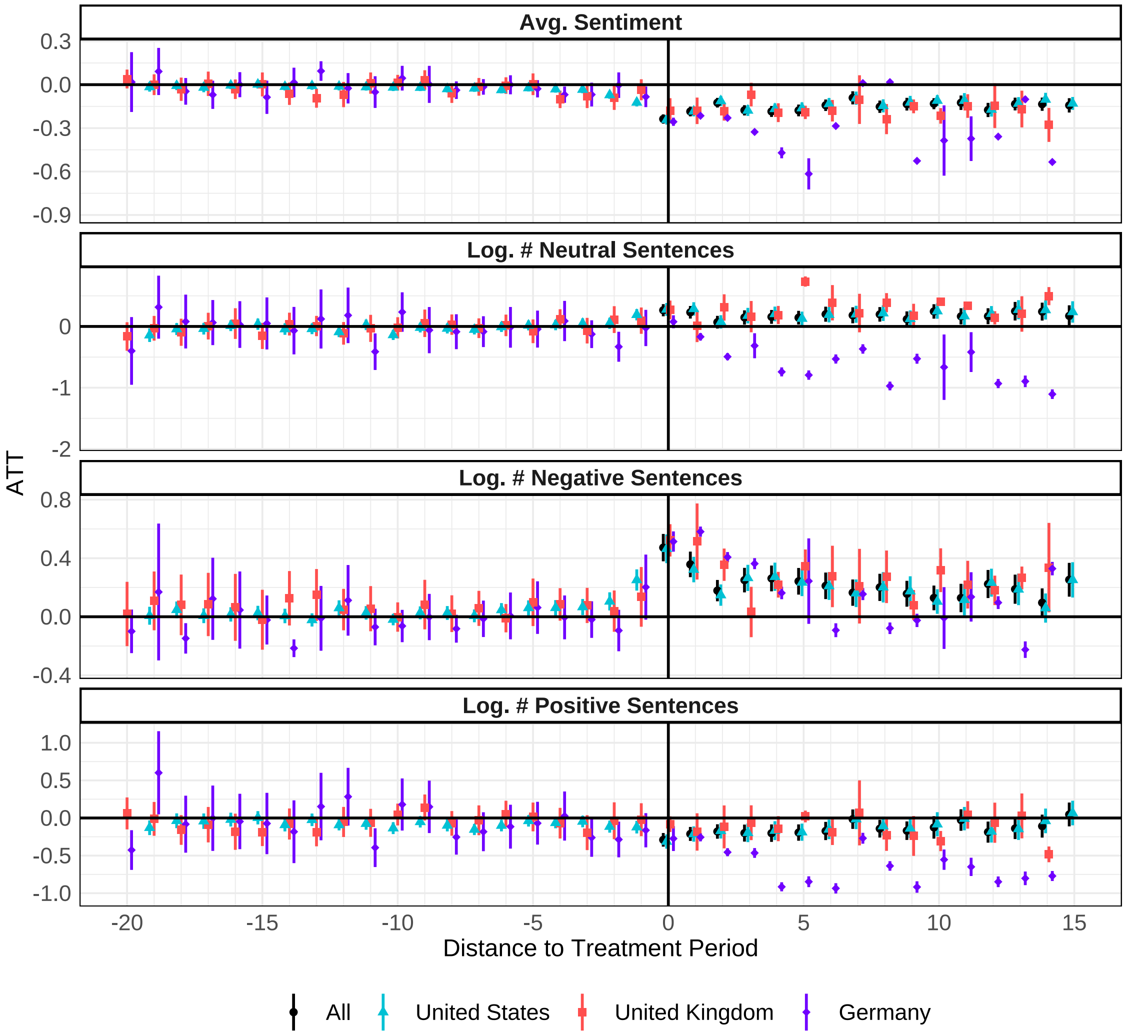
Notes. This figure depicts treatment effect heterogeneity by funding, hype, and atypicality, with 95% confidence intervals. For each moderator, startups are divided into quartiles. Moderators are represented along the horizontal axis, while outcomes are shown on the vertical axis. The baseline effect is indicated by a dashed horizontal line.

4.5.3 Audience

Intra-national

H5a predicts that it matters where the failure takes place, and who the audience is. To test this, we measure heterogeneity across the countries in our sample. In Figure 4.5, the event study of the "domestic" subset, where the startup's country of origin matches the country of the publishing outlet, reveals substantial variation. The response to failure is similar in the United States and the United Kingdom, but German media appears to take a harsher stance. The average sentiment change following failure is significantly more negative in Germany, driven by a reduction in neutral and positive sentences. Interestingly, the change in negative sentences shows less variation. The treatment effects, shown in Table 4.3, are −.14, −.14, and −.25 for the United States, the United Kingdom, and Germany, respectively, with Germany's reaction being notably stronger than in the Anglo-American countries.

FIGURE 4.5: Event Study of Domestic Treatment Effects by Country



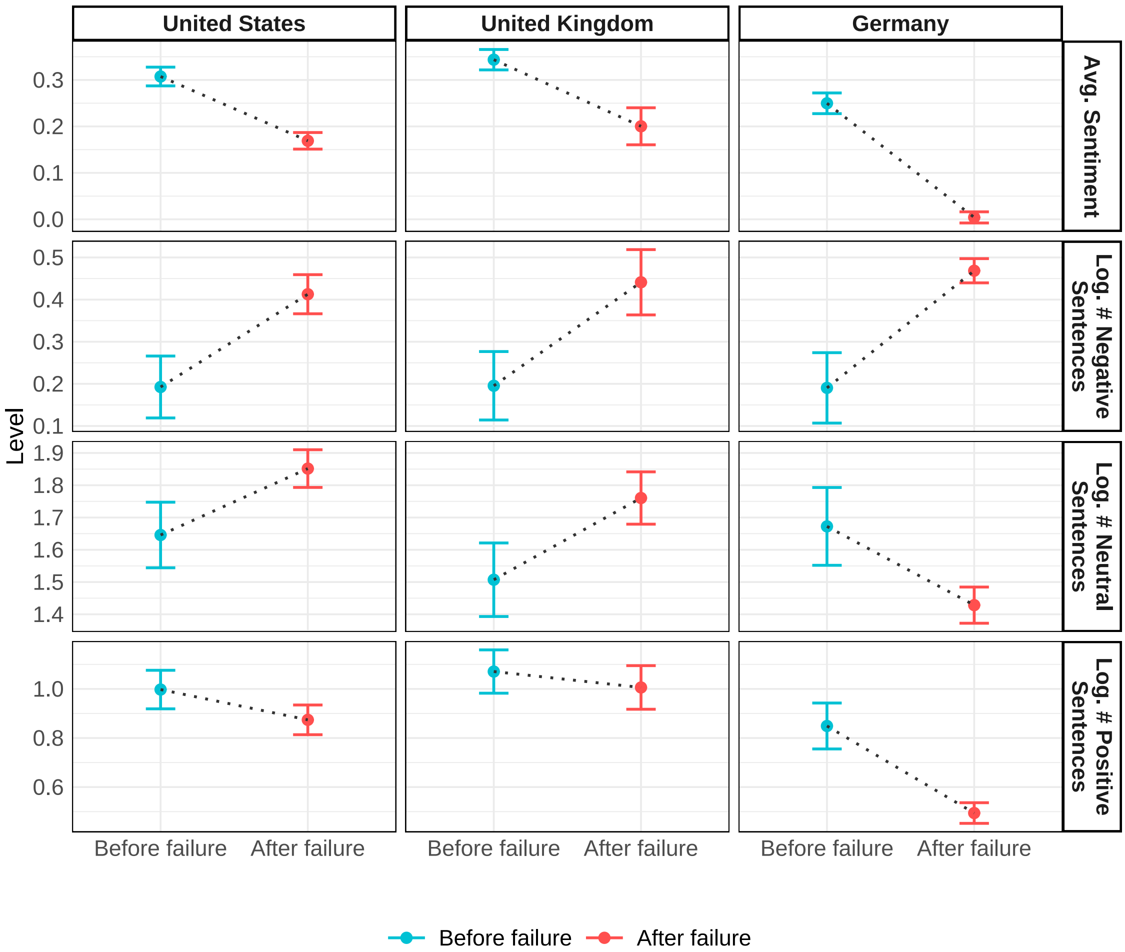
Notes. This figure shows estimated treatment effects (placebo and actual) with 95% confidence intervals over 20 quarters before and 15 quarters after failure, highlighting effect heterogeneity by country compared to the overall effect. The y-axis scales vary by outcome. Consistent with the base event study, anticipation effects are not accounted for. The actual closure date, denoted as $t = 0$, is marked by a black vertical line.

To investigate whether this difference stems from higher baseline values in Germany, and whether post-failure values remain at the same level, we conducted an OLS regression on the not-yet-treated sample[27], in order to be able to compare country differences before

[27]Detailed results of the regression can be found in the appendix. Not-yet-treated does not include startups that never failed.

and after failure. In Figure 4.6, we show that this is not the case. In fact, sentiment levels for operating startups in Germany are already lower. Moreover, the number of neutral sentences increases in both the United States and the United Kingdom, while it drops in Germany, indicating a lower level of factual attention. Overall, we find strong support for H5a, with the caveat that the differences in the effects between the United States and the United Kingdom are negligible.

FIGURE 4.6: Sentiment Level and Change by Country

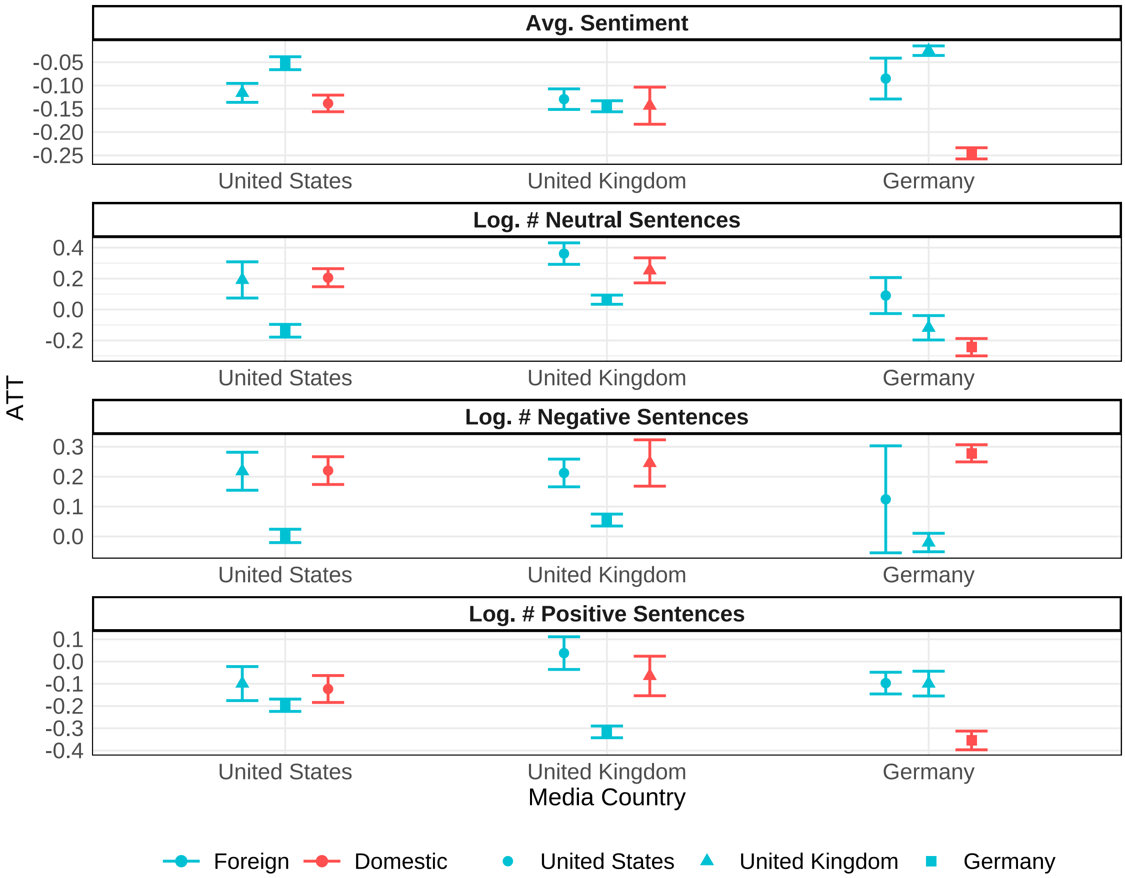


Notes. This figure compares estimated sentiment before and after failure across countries. Blue points represent coefficients from an OLS regression on the not-yet-treated sample, while red points indicate coefficients from the difference-in-differences estimator adjusted for the baseline level. Confidence intervals are set at 95%.

Cross-National

To test H5b, we examine the cross-national treatment interactions and present them graphically in Figure 4.7. H5b predicts that domestic startups would receive more sympathetic coverage compared to foreign startups. However, we find no evidence of a positive in-group bias or favoritism. Instead, our results show that, in general, there is no bias, though in Germany, domestic startups are treated more harshly than foreign ones. Specifically in Germany, American failed startups face a sentiment reduction of −.09, and British startups −.03, while German startups experience a significantly larger reduction of −.25. All other cross-national treatment effects are also negative, though the differences are quite moderate. In summary, we reject H5b but discover an unexpected negative in-group bias for Germany, which complements our prior results of a less entrepreneurial-friendly attitude in Germany.

FIGURE 4.7: Analysis of In-Group Bias



Notes. This graph compares media coverage of domestic versus foreign failed startups. The x-axis represents the country of the media outlet, and the startup country is indicated by the shape of the points. Domestic media-startup pairs are shown in red. Confidence intervals are set at 95%.

TABLE 4.4: Analysis of Press Release and News Sentiment

	Contemporaneous		Lagged		Both
	(1a)	(1b)	(2a)	(2b)	(3)
Press Release ($t = 0$)	0.41 ***	0.41 ***			0.38 ***
	(0.00)	(0.00)			(0.00)
Press Release ($t = -1$)			0.26 ***	0.25 ***	0.14 ***
			(0.00)	(0.00)	(0.00)
Failure		−0.09 ***		−0.11 ***	−0.05 ***
		(0.01)		(0.01)	(0.01)
Failure x Press Release ($t = 0$)		0.01			−0.02
		(0.02)			(0.03)
Failure x Press Release ($t = -1$)				0.05 *	−0.02
				(0.02)	(0.03)
Num.Obs.	168,810	168,810	129,392	129,392	96,467
Mean DV	0.32	0.32	0.3	0.3	0.31
R^2	0.139	0.141	0.056	0.058	0.161

Notes. This table presents results from ordinary least squares regressions of news sentiment on contemporaneous and/or lagged press release sentiment. Standard errors are reported in parentheses, and significance levels are denoted as follows: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

4.5.4 Stigmatization

Prevention

In H6, we predict that startups or entrepreneurs can influence their own media coverage in case of failure. We therefore investigate the relationship between what a startup communicates in its press releases and what is reported about the startup, with a particular focus on the failure event, when the need for justification is likely at its peak.

An initial indication comes from the test of H2: while well-funded organizations are typically better positioned to engage in impression management, we find that they are judged more harshly. To test H6 more formally, in Table 4.4, we regress media sentiment on press release sentiment to assess the magnitude of the correlation between these two variables, though we note that this does not imply causality. When a startup and a journalist report positively about a particular event, it might just be the case that the event happens to be a positive event. As we cannot control for this confounding factors, we are not able to draw causal conclusions.

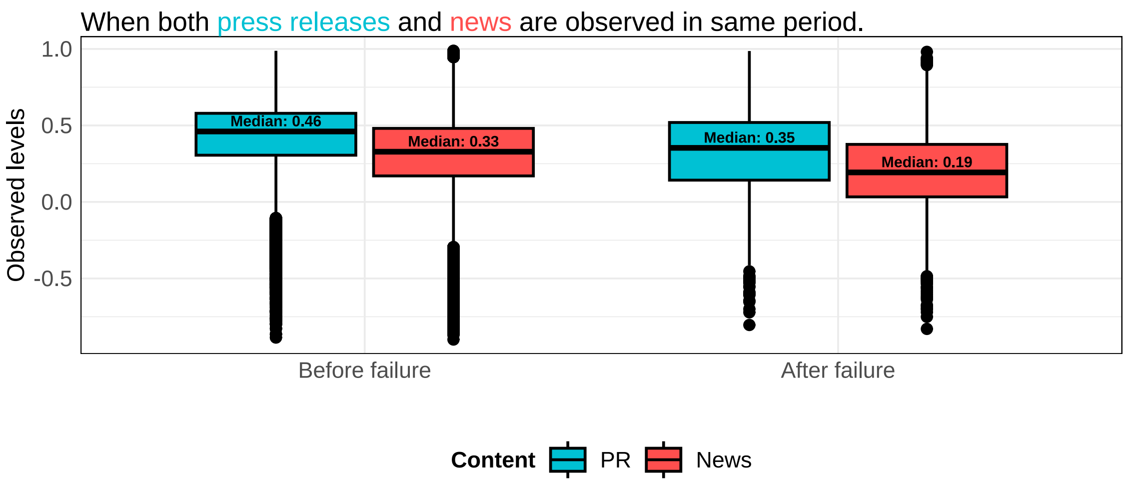
However, what we can infer is that the correlation between press releases and media sentiment in the same month is high [.38, .41], and for press releases from one month prior it is somewhat lower, but still significant [.14, .26]. However, the correlation does not change after the failure event.[28] This suggests that startups, or more specifically founders, cannot (or do not) engage in more effective impression management during times of crisis.

What stands out when comparing press releases to news reports in Figure 4.8 is the

[28] We suspect the interaction effect in the lagged specification is due to random chance.

difference in tone. Press releases tend to be significantly more positive than news coverage (0.46 and 0.33), and this gap persists even after failure (as indicated by the non-zero interaction effect in the regression), although both sentiments decline in magnitude (0.35 and 0.16).

FIGURE 4.8: Relationship Between News and Press Releases



Notes. This graph illustrates descriptive differences in actual sentiment levels between press releases (blue) and news articles (red). It includes only startup-quarters where both a press release and a news article were observed for the same startup.

In conclusion, while journalists may rely on press releases to some extent, a significant portion of the reporting appears to reflect independent journalistic work. In line with these findings, we do not find support for H6, which predicts that founders are able to exert some control over news coverage in the event of failure.

Entrepreneurial Re-Entry

In H7, we hypothesize that the more negatively founders are evaluated following entrepreneurial failure, the less likely they are to pursue entrepreneurial re-entry. The results are shown in Table 4.5.

TABLE 4.5: Logistic Regression of Entrepreneurial Re-Entry

	Sentiment		Sentiment & Name		Sentiment & Volume		Sentiment & Funding	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Sentiment Drop	−0.56 +	−0.61 +	−0.56 +	−0.57	−0.87 *	−0.88 *	−1.53	0.65
	(0.30)	(0.37)	(0.30)	(0.37)	(0.35)	(0.43)	(2.07)	(2.46)
Log. Volume					0.08	0.15		
					(0.15)	(0.16)		
Sentiment Drop X Founder Name			0.00	−0.02				
			(0.03)	(0.05)				
Sentiment Drop X Log. Volume					1.07 +	0.87		
					(0.62)	(0.70)		
Sentiment Drop X Log. Raised Amount							0.06	−0.08
							(0.14)	(0.16)
Gender		0.01		0.00		0.01		0.01
		(0.33)		(0.33)		(0.33)		(0.33)
Entrepreneurial Experience		0.06		0.06		0.06		0.06
		(0.05)		(0.05)		(0.05)		(0.05)
Years Spent at Failed Startup		−0.07		−0.07		−0.06		−0.07
		(0.05)		(0.05)		(0.05)		(0.05)
Degree: Bachelor		−0.36		−0.35		−0.36		−0.37
		(0.33)		(0.33)		(0.33)		(0.33)
Degree: Master		−0.85 *		−0.84 *		−0.87 *		−0.86 *
		(0.34)		(0.34)		(0.34)		(0.34)
Degree: PhD		−1.18 **		−1.19 **		−1.14 *		−1.20 **
		(0.44)		(0.44)		(0.44)		(0.44)
Log. Raised Amount		0.03		0.03		0.01	0.00	0.03
		(0.05)		(0.05)		(0.05)	(0.04)	(0.05)
Num.Obs.	616	521	616	521	616	521	585	521
R^2	0.039	0.076	0.039	0.076	0.045	0.081	0.041	0.076
Mean DV	0.422	0.413	0.422	0.413	0.422	0.413	0.422	0.413
Other Controls	-	Series	-	Series	-	Series	-	Series

Notes. This table presents results from a logit model analyzing the main independent variable, sentiment drop, along with interactions involving founder name, media volume, and funding. Each specification is shown both with and without control variables. The analysis is restricted to failed founders observed after their failure, capturing either the founding of a new venture or a transition to regular employment. Standard errors are reported in parentheses, and significance levels are indicated as follows:+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

We regress entrepreneurial restart on the drop in media sentiment following the failure and include only founders, where we observe any record of employment or a new venture after failure. We report the effect with and without controlling for characteristics of the individual and failed startup and test whether the effect is moderated by the occurrence of the founder's name in the article, the volume of news, or the funding sum of the failed startup. With controls, the main effect is $-.61$, which translates into a reduction of 4.3%pt (percentage points) and 11.9% across the interquartile range (from $q_{25\%}$ to $q_{75\%}$). Except for one significant coefficient when interacting volume and the reduction in sentiment (however, without controls), we do not find evidence of any effect heterogeneity.[29]

In conclusion, we show that media coverage of failure, among other factors, indeed influences the decision for entrepreneurial re-entry and find support for H7.

4.6 Discussion

Exploiting a unique dataset, our study examines the social evaluation of entrepreneurial failure through a large-scale analysis. Our findings reveal that failure is consistently portrayed in a negative light by the news media, though significant heterogeneity exists depending on contextual factors, including both startup characteristics and the institutional audience. In a follow-up exploration, we find no evidence that entrepreneurs can influence the tone of media coverage during times of crisis. Instead, they are subject to the media's evaluation, which in turn impacts their prospects for starting future entrepreneurial ventures.

4.6.1 Empirical Contributions

Few studies have examined the portrayal of business failure in the media (Sheng and Lan, 2019; Zavyalova et al., 2012), while a substantial body of literature has demonstrated the power of media institutions and reports in shaping legitimacy (e.g., Brown and Deegan (1998); Deephouse (1996); Deephouse and Carter (2005); Elsbach (1994); Lamertz and Baum (1998); Pollock and Rindova (2003); Überbacher (2014)). Legitimacy is crucial for startups as they strive to overcome the liability of newness (Singh et al., 1986; Stinchcombe, 2000). With the exception of Jacobi (2018), our study is the first to investigate how the

[29] On a different note, it appears that higher-skilled individuals are less likely to found a new venture; for example, an individual with a PhD is 6.8%pt or 19% less likely to restart than an individual whose highest degree is a Master's.

media evaluates entrepreneurial failure, which is so inherently tied to the entrepreneurial process (Wiklund et al., 2010; Bruderl et al., 1992; Shane, 2009). Media evaluation is especially important for startups, which differ significantly from established firms. Founders, who are deeply intertwined with their ventures, rely heavily on favorable media portrayals to send positive signals to various stakeholders. Our findings suggest that media evaluations of entrepreneurial failure are consistently negative, with notable heterogeneity based on startup characteristics. Startups with more funding are judged more harshly than their less-funded counterparts, while those that are atypical within their market category appear to be somewhat shielded from negative reporting. Interestingly, we find no evidence that industry hype plays a moderating role, suggesting that the gap between startups in hyped and non-hyped industries remains unchanged even after failure.

We not only examine who is reported on, but also who reports and to whom. Entrepreneurial failure is a social construct (Radu-Lefebvre et al., 2019) that unfolds within a social context, shaped by institutional actors such as the media. Ventures and their ideas are received within specific contexts and climates that heavily influence the entrepreneurial experience (Cardon et al., 2011). As such, the startup's country of origin and the media's country of reporting play a crucial role. Our findings demonstrate that the countries in our sample, despite being well-developed and culturally similar, produce markedly distinct results. This aligns with Fisher et al. (2017), who argue that institutions are deeply embedded in their unique social and cultural contexts, each governed by different sets of rules and heuristics. Similarly, Lee et al. (2007) show that societal institutions impact entrepreneurial friendliness. While previous research has emphasized the importance of audience diversity in the legitimation process across stakeholder groups (Überbacher, 2014; Fisher, 2020), we add to the literature by considering and demonstrating the importance of audience diversity at the societal level.

By extending the analysis to cross-country media coverage, we also complement the literature by gaining valuable insights into the presence of in-group bias, which, contrary to expectations, does not appear to exist. Surprisingly, we find that in Germany, the least positive country in our study, there is even a negative in-group bias. This unexpected finding complicates legitimation efforts and reinforces negative responses to failure, consistent with Jacobi (2018).

Exploring the role of media in entrepreneurship, we respond to calls by Hindle and

Klyver (2007) and Carroll and McCombs (2003) and contribute to the literature on national culture and entrepreneurship (e.g., Stephan and Uhlaner (2010); Kleinhempel et al. (2022); Taylor and Wilson (2012)). We highlight the importance of comparative entrepreneurship research, addressing the call by McGrath (1999) to investigate entrepreneurial payoff structures across diverse cultures. Our analysis of media coverage illustrates the heterogeneous nature of entrepreneurial phenomena across nations. In doing so, we emphasize the well-recognized challenge of generalizing findings in management science (Terjesen et al., 2016; Rosenzweig, 1994; Aharoni and Burton, 1994; Cardon et al., 2011).

4.6.2 Practical Implications

Our analysis of the relationship between press releases and news articles suggests that startups and founders cannot influence the sentiment of news coverage during failure events. Although we observe a strong correlation between press release sentiment and news sentiment, the interaction with failure does not yield a significant effect. This is surprising, as one might expect startups to be more proactive in times of crisis and engage in impression management. While we cannot definitively determine whether startups engage in such efforts, our findings suggest, that if they do, these efforts are ineffective.

Additionally, we explore the implications of negative media coverage for founders and their likelihood of restarting a new venture, finding a negative effect. It appears that some founders feel stigmatized by the failure and are discouraged from pursuing new entrepreneurial endeavors. While these implications primarily affect individuals, they may also have broader societal implications. Media coverage shapes public opinion about entrepreneurial actors (Deephouse, 2000), and entrepreneurial legitimacy, a form of aggregated legitimacy (Radu-Lefebvre et al., 2019), influences career choices and the diffusion of entrepreneurial culture at a societal level. Therefore, the harsh treatment of failed entrepreneurs is likely to reduce overall entrepreneurial activity.

4.6.3 Limitations and Future Research

We recognize certain limitations in our study, providing avenues for future research. First, we acknowledge potential problems with our failure indicator. The observed failure rate in our sample is lower than what would have been expected from known statistics (Bosma, 2018). It is plausible that some startups in our sample are in fact closed without our knowledge. Assuming unobserved failed startups are treated similarly to observed ones, this introduces bias that reduces the estimated effects, suggesting that the true effect may be even more pronounced. However, variations in failure data availability across countries or other moderators could potentially distort effects and relative rankings, though we lack indications of such discrepancies.

Moreover, business closures can stem from highly specific and diverse reasons, such as founders transitioning to salaried positions, entering retirement, experiencing insufficient returns from startups, or encountering legal issues. Our study, however, does not differentiate between these various reasons. While the cause of business closure may moderate treatment effects and exhibit variations in composition across countries, potentially influencing our results and conclusions, this nuance is not explicitly considered in our analysis. In future studies, addressing the challenge of identifying failure dates and associated causes could benefit from the application of advanced natural language processing techniques which would allow a more nuanced and accurate analysis.

Secondly, we did not account for potential behavioral changes of founders induced by sentiment or fear of expected sentiment. When founders react by prematurely or delaying ceasing operations in response to expected or actual coverage, and sentiment levels vary across countries, we might introduce bias.

Thirdly, our sentiment analysis is limited to the identification of negative and positive emotions. This approach, while capturing basic sentiments, results in significant information loss, as text inherently conveys more nuanced meanings. What our study labels as negative could encompass criticism, regret, hostility, or other emotions with negative connotations. Future research could enhance the analysis by employing a more fine-grained sentiment extraction approach to achieve more precise sentiment predictions. More nuanced approaches may also be able to detect impression management strategies and their effectiveness in greater detail, potentially leading to different conclusions.

Fourthly, studies could further explore entrepreneurial re-entry in more detail by analyzing the characteristics of founders who restart ventures, as well as the performance and success of these new ventures. Specifically, research could examine whether a learning effect is observable and how it depends on contextual factors. Preliminary findings from our analysis suggest that highly skilled founders tend to refrain from restarting, indicating the potential existence of a negative selection effect.

Chapter 5

Overall Discussion and Conclusion

This dissertation provides a comprehensive exploration of the entrepreneurial life cycle by examining three critical stages: the precursors of entrepreneurial action, the operational phase of startups, and the aftermath of failure, including entrepreneurial re-entry. Each study contributes uniquely to our understanding of the interconnected nature of these phases and their implications for entrepreneurship theory and practice.

5.1 Discussion

Our findings highlight the dynamic and cyclical nature of entrepreneurship, emphasizing how individual, organizational, and societal factors interact across different stages. In study 1, we demonstrate the foundational role of entrepreneurial attitudes, showing that team autonomy positively influences entrepreneurial intention but not self-efficacy, with the effect mediated by team diversity. Effects from peer interactions, however, only show up for self-efficacy, and only when some dimension of autonomy is guaranteed. This study establishes the importance of structured interventions in fostering entrepreneurial attitudes, key precursors to venture creation.

Building on this, in study 2, we investigate the operational phase and find that VC monitoring, facilitated by physical proximity, significantly impacts startup outcomes such as innovation, survival, and acquisitions. By leveraging natural experiments, we emphasize the critical role of external enablers in shaping entrepreneurial success and highlight the differential effects of permanent versus temporary disruptions to VC-startup interactions.

Finally, in study 3, we address the aftermath of entrepreneurial failure, showing that media portrayals significantly influence the likelihood of founders restarting ventures. We find that negative coverage exacerbates stigma and discourages re-entry, whereas balanced narratives can mitigate these effects. By linking failure to re-entry, this study underscores the potentially cyclical nature of the entrepreneurial process and highlights the societal role in fostering entrepreneurial resilience.

5.2 Contributions

Our research makes several contributions to the field of entrepreneurship. Thematically, we advance understanding across the entire entrepreneurial life cycle, offering insights into the interplay between individual agency, external enablers, societal perceptions, and entepreneurial phenomena. Collectively, our studies demonstrate that entrepreneurship is not a linear journey but a cyclical process, in which recovery and re-entry after failure are just as academically and practically significant as the initial stages and operational success.

From a methodological perspective, this work advances entrepreneurship research by applying rigorous identification strategies designed to address the empirical challenges inherent in studying the entrepreneurial process. By utilizing experimental designs in study 1, natural experiments in study 2, and quasi-experimental approaches in study 3, we demonstrate effective methods for inferring causal effects across settings with varying levels of empirical complexity. These rigorous approaches strengthen the validity of our findings and lay a solid groundwork for future research.

5.3 Implications

Our findings have important implications across individual, organizational, and societal levels. For educators and policymakers, study 1 highlights the need to balance autonomy and diversity in entrepreneurial training programs to effectively nurture entrepreneurial attitudes. For venture capitalists and startups, study 2 underscores the importance of physical proximity and strategic location in fostering effective monitoring relationships. Finally, study 3 points to the critical role of media narratives in shaping societal perceptions of failure, informing policy-makers to reduce stigma and encourage entrepreneurial re-entry.

At a broader level, our findings highlight the interconnected nature of the entrepreneurial ecosystem. By supporting entrepreneurs from intention to re-entry, we can help ensure a dynamic entrepreneurial landscape capable of driving innovation, employment, and economic growth.

5.4 Outlook and Future Research

This dissertation opens avenues for further research in various ways. For instance, future studies could explore the effects of team autonomy on entrepreneurial outcomes and attitudes in more heterogenous settings, disentangle the components of remote and in-person venture capital monitoring, and investigate the mechanisms underlying media narratives and their influence on entrepreneurial resilience.

Bibliography

Abadie, A. 2005. Semiparametric Difference-in-Differences Estimators. ***The Review of Economic Studies***, 72(1): 1–19.

Abbasiharofteh, M., Kogler, D. F., and Lengyel, B. 2023. Atypical combinations of technologies in regional co-inventor networks. ***Research Policy***, 52(10): 1–17.

Acharya, A., Blackwell, M., and Sen, M. 2016. Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects. ***American Political Science Review***, 110(3): 512–529.

Adoni, H. and Mane, S. 1984. Media and the social construction of reality: Toward an integration of theory and research. ***Communication Research***, 11(3): 323–340.

Aghion, P., Van Reenen, J., and Zingales, L. 2013. Innovation and Institutional Ownership. ***American Economic Review***, 103(1): 277–304.

Agrawal, A., Catalini, C., and Goldfarb, A. 2015. Crowdfunding: Geography, Social Networks, and the Timing of Investment Decisions. ***Journal of Economics & Management Strategy***, 24(2): 253–274.

Agrawal, A. and Goldfarb, A. 2008. Restructuring Research: Communication Costs and the Democratization of University Innovation. ***American Economic Review***, 98(4): 1578–1590.

Aharoni, Y. and Burton, R. M. 1994. Is Management Science International: In Search of Universal Rules. ***Management Science***, 40(1): 1–3.

Ahern, K. R., Duchin, R., and Shumway, T. 2014. Peer Effects in Risk Aversion and Trust. ***Review of Financial Studies***, 27(11): 3213–3240.

Ahern, K. R. and Sosyura, D. 2015. Rumor Has It: Sensationalism in Financial Media. ***The Review of Financial Studies***, 28(7): 2050–2093.

Ajzen, I. 1991. The theory of planned behavior. ***Organizational Behavior and Human Decision Processes***, 50(2): 179–211.

Aldrich, H. E. and Fiol, C. M. 1994. Fools Rush in? The Institutional Context of Industry Creation. ***The Academy of Management Review***, 19(4): 105–127.

Aldrich, H. E. and Martinez, M. 2003. Entrepreneurship as Social Construction: A Multilevel Evolutionary Approach. In Acs, Z. J. and Audretsch, D. B.(Eds.), ***Handbook of Entrepreneurship Research***, 359–399. Boston, MA: Springer US.

Amason, A. C., Shrader, R. C., and Tompson, G. H. 2006. Newness and novelty: Relating top management team composition to new venture performance. ***Journal of Business Venturing***, 21(1): 125–148.

Amornsiripanitch, N., Gompers, P. A., and Xuan, Y. 2019. More than Money: Venture Capitalists on Boards. ***The Journal of Law, Economics, and Organization***, 35(3): 513–543.

Ando, Y. 2024. Dynamics of High-Growth Young Firms and the Role of Venture Capitalists. ***PIER Working Paper Archive***, 24–012.

Angrist, J. D. 2014. The perils of peer effects. ***Labour Economics***, 30: 98–108.

Audretsch, D. B. 2007. ***The Entrepreneurial Society***. Oxford, [England] New York: Oxford University Press.

Bagozzi, R. P., Bergami, M., and Leone, L. 2003. Hierarchical representation of motives in goal setting. ***Journal of Applied Psychology***, 88(5): 915–943.

Bailyn, L. 1985. Autonomy in the industrial R&D lab. ***Human Resource Management***, 24(2): 129–146.

Baker, M. P. and Gompers, P. A. 2000. The Determinants of Board Structure at the Initial Public Offering. ***SSRN Electronic Journal***.

Bandura, A. 1982. Self-efficacy mechanism in human agency. ***American Psychologist***, 37(2): 122–147.

Bandura, A. 2012. Self-Efficacy. In Ramachandran, V. S.(Ed.), ***Encyclopedia of Human Behavior***. San Diego, CA, USA: Elsevier Science & Technology Books.

Barbieri, F., Camacho-Collados, J., Neves, L., and Espinosa-Anke, L. 2020. TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification.

Barrault, L., Bojar, O., Costa-jussà, M. R., Federmann, C., Fishel, M., Graham, Y., Haddow, B., Huck, M., Koehn, P., Malmasi, S., Monz, C., Müller, M., Pal, S., Post, M., and Zampieri, M. 2019. Findings of the 2019 Conference on Machine Translation (WMT19). ***Proceedings of the Fourth Conference on Machine Translation (Volume 2: Shared Task Papers, Day 1)***.

Bartling, B., Fehr, E., and Herz, H. 2014. The Intrinsic Value of Decision Rights: Intrinsic Value of Decision Rights. ***Econometrica***, 82(6): 2005–2039.

Bartling, B., Fehr, E., and Schmidt, K. M. 2012. Screening, Competition, and Job Design: Economic Origins of Good Jobs. ***American Economic Review***, 102(2): 834–864.

Baruffaldi, S. and Poege, F. 2024. Like Stars: How Firms Learn at Scientific Conferences. ***Management Science***, 71(3): 1865–1888.

Begley, T. 2001. East Asian and Anglo-SaxonCountries. ***Journal of International Business Studies***, 32(3): 537–553.

Benz, M. and Frey, B. S. 2008a. Being Independent is a Great Thing: Subjective Evaluations of Self-Employment and Hierarchy. ***Economica***, 75(298): 362–383.

Benz, M. and Frey, B. S. 2008b. The value of doing what you like: Evidence from the self-employed in 23 countries. ***Journal of Economic Behavior & Organization***, 68(3-4): 445–455.

Berger, J., Sorensen, A. T., and Rasmussen, S. J. 2010. Positive Effects of Negative Publicity: When Negative Reviews Increase Sales. ***Marketing Science***, 29(5): 815–827.

Bernstein, S., Giroud, X., and Townsend, R. R. 2016. The Impact of Venture Capital Monitoring. ***The Journal of Finance***, 71(4): 1591–1622.

Bertoni, F., Colombo, M. G., and Grilli, L. 2011. Venture capital financing and the growth of high-tech start-ups: Disentangling treatment from selection effects. ***Research Policy***, 40(7): 1028–1043.

Bitektine, A. 2011. Toward a Theory of Social Judgments of Organizations: The Case of Legitimacy, Reputation, and Status. ***Academy of Management Review***, 36(1): 151–179.

Bitektine, A. and Haack, P. 2015. The "Macro" and the "Micro" of Legitimacy: Toward a Multilevel Theory of the Legitimacy Process. ***Academy of Management Review***, 40(1): 49–75.

Block, J. and Koellinger, P. 2009. I Can't Get No Satisfaction—Necessity Entrepreneurship and Procedural Utility. ***Kyklos***, 62(2): 191–209.

Blum, B. S. and Goldfarb, A. 2006. Does the internet defy the law of gravity? ***Journal of International Economics***, 70(2): 384–405.

Borusyak, K., Jaravel, X., and Spiess, J. 2024. Revisiting Event-Study Designs: Robust and Efficient Estimation. ***Review of Economic Studies***, 91(6): 3253–3285.

Bosma, N. 2018. ***Global entrepreneurship monitor: 2018/2019 Global report***. Chile: Universidad del Desarrollo.

Bosma, N., Hessels, J., Schutjens, V., Praag, M. V., and Verheul, I. 2012. Entrepreneurship and role models. ***Journal of Economic Psychology***, 33(2): 410–424.

Boss, V., Dahlander, L., Ihl, C., and Jayaraman, R. 2023. Organizing Entrepreneurial Teams: A Field Experiment on Autonomy over Choosing Teams and Ideas. ***Organization Science***, 34(6): 2097–2118.

Bottazzi, L., Rin, M. D., and Hellmann, T. 2008. Who are the active investors? Evidence from venture capital. ***Journal of Financial Economics***, 89(3): 488–512.

Bronzini, R., Caramellino, G., and Magri, S. 2020. Venture capitalists at work: A diff-in-diff approach at late-stages of the screening process. ***Journal of Business Venturing***, 35(3): 1–20.

Brown, A. D. 1997. Narcissism, Identity, and Legitimacy. ***Academy of Management Review***, 22(3): 643–686.

Brown, N. and Deegan, C. 1998. The public disclosure of environmental performance information—a dual test of media agenda setting theory and legitimacy theory. ***Accounting and Business Research***, 29(1): 21–41.

Bruderl, J., Preisendorfer, P., and Ziegler, R. 1992. Survival Chances of Newly Founded Business Organizations. ***American Sociological Review***, 57(2): 227–247.

Bundy, J. and Pfarrer, M. D. 2015. A Burden of Responsibility: The Role of Social Approval at the Onset of a Crisis. ***Academy of Management Review***, 40(3): 345–369.

Butts, K. and Gardner, J. 2022. {did2s}: Two-Stage Difference-in-Differences. ***The R Journal***, 14(3): 162–173.

Büyükboyaci, M. and Robbett, A. 2019. Team formation with complementary skills. ***Journal of Economics & Management Strategy***, 28(4): 713–733.

Caetano, C. and Callaway, B. 2024. Difference-in-Differences when Parallel Trends Holds Conditional on Covariates.

Caetano, C., Callaway, B., Payne, S., and Rodrigues, H. S. 2024. Difference in Differences with Time-Varying Covariates.

Cahn, C., Girotti, M., and Landier, A. 2021. Entrepreneurship and information on past failures: A natural experiment. ***Journal of Financial Economics***, 141(1): 102–121.

Cairncross, F. 2001. ***The Death of Distance: How the Communications Revolution Is Changing Our Lives***. Boston: Harvard Business School Press, completely new ed edition.

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

Cardon, M. S., Stevens, C. E., and Potter, D. R. 2011. Misfortunes or mistakes? ***Journal of Business Venturing***, 26(1): 79–92.

Carroll, C. E. and McCombs, M. 2003. Agenda-setting Effects of Business News on the Public's Images and Opinions about Major Corporations. ***Corporate Reputation Review***, 6(1): 36–46.

Cassar, L. and Meier, S. 2018. Nonmonetary Incentives and the Implications of Work as a Source of Meaning. ***Journal of Economic Perspectives***, 32(3): 215–238.

Catalini, C., Fons-Rosen, C., and Gaulé, P. 2020. How Do Travel Costs Shape Collaboration? ***Management Science***, 66(8): 3340–3360.

Cattani, G., Ferriani, S., Mariani, M. M., and Mengoli, S. 2013. Tackling the "Galacticos" effect: Team familiarity and the performance of star-studded projects. ***Industrial and Corporate Change***, 22(6): 1629–1662.

Chemmanur, T. J., Krishnan, K., and Nandy, D. K. 2011. How Does Venture Capital Financing Improve Efficiency in Private Firms? A Look Beneath the Surface. ***Review of Financial Studies***, 24(12): 4037–4090.

Chen, H., Gompers, P., Kovner, A., and Lerner, J. 2010. Buy local? The geography of venture capital. ***Journal of Urban Economics***, 67(1): 90–102.

Chen, R. 2017. Coordination with endogenous groups. ***Journal of Economic Behavior & Organization***, 141: 177–187.

Chen, R. and Gong, J. 2018. Can self selection create high-performing teams? ***Journal of Economic Behavior & Organization***, 148: 20–33.

Cinelli, C., Forney, A., and Pearl, J. 2024. A Crash Course in Good and Bad Controls. ***Sociological Methods & Research***, 53(3): 1071–1104.

Clarysse, B., Bobelyn, A., and Del Palacio Aguirre, I. 2013. Learning from own and others' previous experience: The contribution of the venture capital firm to the likelihood of a portfolio company's trade sale. ***Small Business Economics***, 40(3): 575–590.

Cope, J. 2011. Entrepreneurial learning from failure: An interpretative phenomenological analysis. ***Journal of Business Venturing***, 26(6): 604–623.

Cope, J., Cave, F., and Eccles, S. 2004. Attitudes of venture capital investors towards entrepreneurs with previous business failure. ***Venture Capital***, 6(2-3): 147–172.

Cuddy, A. J. C., Fiske, S. T., Kwan, V. S. Y., Glick, P., Demoulin, S., Leyens, J.-P., Bond, M. H., Croizet, J.-C., Ellemers, N., Sleebos, E., Htun, T. T., Kim, H.-J., Maio, G., Perry, J., Petkova, K., Todorov, V., Rodríguez-Bailón, R., Morales, E., Moya, M., Palacios, M., Smith, V., Perez, R., Vala, J., and Ziegler, R. 2009. Stereotype content model across cultures: Towards universal similarities and some differences. ***British Journal of Social Psychology***, 48(1): 1–33.

Daft, R. L. and Lengel, R. H. 1986. Organizational Information Requirements, Media Richness and Structural Design. ***Management Science***, 32(5): 554–571.

D'Aveni, R. A. 1990. Top Managerial Prestige and Organizational Bankruptcy. ***Organization Science***, 1(2): 121–142.

Davidsson, P. and Honig, B. 2003. The role of social and human capital among nascent entrepreneurs. ***Journal of Business Venturing***, 18(3): 301–331.

Davila, A., Foster, G., and Gupta, M. 2003. Venture capital financing and the growth of startup firms. ***Journal of Business Venturing***, 18(6): 689–708.

de Chaisemartin, C. and D'Haultfœuille, X. 2023. Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. ***The Econometrics Journal***, 26(3): 1–30.

Deci, E. L. and Ryan, R. M. 1985. ***Intrinsic Motivation and Self-Determination in Human Behavior***. Boston, MA: Springer US.

Deci, E. L. and Ryan, R. M. 2000. The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. ***Psychological Inquiry***, 11(4): 227–268.

Deeds, D. L., Mang, P. Y., and Frandsen, M. L. 2004. The Influence of Firms' and Industries' Legitimacy on the Flow of Capital into High-Technology Ventures. ***Strategic Organization***, 2(1): 9–34.

Deephouse, D. L. 1996. Does Isomorphism Legitimate? ***Academy of Management Journal***, 39(4): 1024–1039.

Deephouse, D. L. 2000. Media Reputation as a Strategic Resource: An Integration of Mass Communication and Resource-Based Theories. ***Journal of Management***, 26(6): 1091–1112.

Deephouse, D. L. and Carter, S. M. 2005. An Examination of Differences Between Organizational Legitimacy and Organizational Reputation*. ***Journal of Management Studies***, 42(2): 329–360.

Deephouse, D. L. and Heugens, P. P. M. A. R. 2009. Linking Social Issues to Organizational Impact: The Role of Infomediaries and the Infomediary Process. ***Journal of Business Ethics***, 86(4): 541–553.

Delmar, F. and Shane, S. 2004. Legitimating first: Organizing activities and the survival of new ventures. ***Journal of Business Venturing***, 19(3): 385–410.

Dobrev, S. D. and Gotsopoulos, A. 2010. Legitimacy vacuum, structural imprinting, and the first mover disadvantage. ***Academy of Management Journal***, 53(5): 1153–1174.

Douze, M., Guzhva, A., Deng, C., Johnson, J., Szilvasy, G., Mazaré, P.-E., Lomeli, M., Hosseini, L., and Jégou, H. 2024. The Faiss library.

Durand, R. and Vergne, J.-P. 2015. Asset divestment as a response to media attacks in stigmatized industries. ***Strategic Management Journal***, 36(8): 1205–1223.

Elsbach, K. D. 1994. Managing Organizational Legitimacy in the California Cattle Industry: The Construction and Effectiveness of Verbal Accounts. ***Administrative Science Quarterly***, 39(1): 57–88.

Engle, R. L., Schlaegel, C., and Delanoe, S. 2011. The Role of Social Influence, Culture, and Gender on Entrepreneurial Intent. ***Journal of Small Business & Entrepreneurship***, 24(4): 471–492.

Etzioni, A. 1987. Entrepreneurship, adaptation and legitimation. ***Journal of Economic Behavior & Organization***, 8(2): 175–189.

Ewens, M. and Marx, M. 2018. Founder Replacement and Startup Performance. ***The Review of Financial Studies***, 31(4): 1532–1565.

Falk, A. and Kosfeld, M. 2006. The Hidden Costs of Control. ***American Economic Review***, 96(5): 1611–1630.

Fang, L. H. and Peress, J. 2008. Media Coverage and the Cross-Section of Stock Returns. ***SSRN Electronic Journal***.

Fehr, E., Herz, H., and Wilkening, T. 2013. The Lure of Authority: Motivation and Incentive Effects of Power. ***American Economic Review***, 103(4): 1325–1359.

Feld, J. and Zölitz, U. 2017. Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects. ***Journal of Labor Economics***, 35(2): 387–428.

Ferguson, T. D., Deephouse, D. L., and Ferguson, W. L. 2000. Do strategic groups differ in reputation? ***Strategic Management Journal***, 21(12): 1195–1214.

Fischer, R. and Derham, C. 2016. Is in-group bias culture-dependent? A meta-analysis across 18 societies. ***SpringerPlus***, 5(1): 1–9.

Fisher, G. 2020. The Complexities of New Venture Legitimacy. ***Organization Theory***, 1(2): 1–25.

Fisher, G., Kuratko, D. F., Bloodgood, J. M., and Hornsby, J. S. 2017. Legitimate to whom? The challenge of audience diversity and new venture legitimacy. ***Journal of Business Venturing***, 32(1): 52–71.

Fisher, R. A. 1992. Statistical Methods for Research Workers. In Kotz, S. and Johnson, N. L.(Eds.), ***Breakthroughs in Statistics***, 66–70. New York, NY: Springer New York.

Fombrun, C. and Shanley, M. 1990. What's in a Name? Reputation Building and Corporate Strategy. ***Academy of Management Journal***, 33(2): 233–258.

Forman, C., Goldfarb, A., and Greenstein, S. 2005. How did location affect adoption of the commercial Internet? Global village vs. urban leadership. ***Journal of Urban Economics***, 58(3): 389–420.

Frey, B. S., Benz, M., and Stutzer, A. 2004. Introducing Procedural Utility: Not Only What, but Also How Matters. ***Journal of Institutional and Theoretical Economics***, 160(3): 377–401.

Gamson, W. A., Croteau, D., Hoynes, W., and Sasson, T. 1992. Media Images and the Social Construction of Reality. ***Annual Review of Sociology***, 18(1): 373–393.

Gardner, J. 2022. Two-stage differences in differences. arXiv:2207.05943.

Gartner, W., Shaver, K., Carter, N., and Reynolds, P. 2004. ***Handbook of Entrepreneurial Dynamics: The Process of Business Creation***. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc.

Garud, R., Lant, T. K., and Schildt, H. A. 2019. Generative imitation, strategic distancing and optimal distinctiveness during the growth, decline and stabilization of Silicon Alley. ***Innovation***, 21(1): 187–213.

Garud, R., Schildt, H. A., and Lant, T. K. 2014. Entrepreneurial Storytelling, Future Expectations, and the Paradox of Legitimacy. ***Organization Science***, 25(5): 1479–1492.

Gaspar, J. and Glaeser, E. L. 1998. Information Technology and the Future of Cities. ***Journal of Urban Economics***, 43(1): 136–156.

Giroud, X. 2013. Proximity and Investment: Evidence from Plant-Level Data*. ***The Quarterly Journal of Economics***, 128(2): 861–915.

Goffman, E. 2022. ***Stigma: Notes on the Management of Spoiled Identity***. Penguin Modern Classics. Dublin: Penguin Books.

Goldberg, A., Hannan, M. T., and Kovács, B. 2016. What Does It Mean to Span Cultural Boundaries? Variety and Atypicality in Cultural Consumption. ***American Sociological Review***, 81(2): 215–241.

Gompers, P. A., Gornall, W., Kaplan, S. N., and Strebulaev, I. A. 2020. How do venture capitalists make decisions? ***Journal of Financial Economics***, 135(1): 169–190.

Gompers, P. A., Huang, K., and Wang, S. Q. 2017. Homophily in Entrepreneurial Team Formation. ***SSRN Electronic Journal***.

Goodman-Bacon, A. 2021. Difference-in-differences with variation in treatment timing. ***Journal of Econometrics***, 225(2): 254–277.

Gorman, M. and Sahlman, W. A. 1989. What do venture capitalists do? ***Journal of Business Venturing***, 4(4): 231–248.

Graf-Vlachy, L., Oliver, A. G., Banfield, R., König, A., and Bundy, J. 2020. Media Coverage of Firms: Background, Integration, and Directions for Future Research. ***Journal of Management***, 46(1): 36–69.

Graham, B. S. and Hahn, J. 2005. Identification and estimation of the linear-in-means model of social interactions. ***Economics Letters***, 88(1): 1–6.

Gulati, R. and Gargiulo, M. 1999. Where Do Interorganizational Networks Come From? ***American Journal of Sociology***, 104(5): 1439–1493.

Gurun, U. G. and Butler, A. W. 2012. Don't Believe the Hype: Local Media Slant, Local Advertising, and Firm Value. ***The Journal of Finance***, 67(2): 561–598.

Hacamo, I. and Kleiner, K. 2018. Believe in Yourself: Do Peers Promote or Discourage Entrepreneurship? ***SSRN Electronic Journal***.

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., and Tatlow, H. 2021. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). ***Nature Human Behaviour***, 5(4): 529–538.

Hamilton, B. H., Nickerson, J. A., and Owan, H. 2003. Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. ***Journal of Political Economy***, 111(3): 465–497.

Hansen, Z., Owan, H., and Pan, J. 2015. The impact of group diversity on class performance: Evidence from college classrooms. ***Education Economics***, 23(2): 238–258.

Hasan, S. and Koning, R. 2019. Prior ties and the limits of peer effects on startup team performance. ***Strategic Management Journal***, 40(9): 1394–1416.

Hashai, N. and Zahra, S. 2022. Founder team prior work experience: An asset or a liability for startup growth? ***Strategic Entrepreneurship Journal***, 16(1): 155–184.

Hayton, J. C., George, G., and Zahra, S. A. 2002. National Culture and Entrepreneurship: A Review of Behavioral Research. ***Entrepreneurship Theory and Practice***, 26(4): 33–52.

Headd, B. 2003. Redefining Business Success: Distinguishing between Closure and Failure. ***Small Business Economics***, 21(1): 51–61.

Heath, D., Seegert, N., and Yang, J. 2023. Team Production and the Homophily Trap: Evidence from Open Source Software. ***SSRN Electronic Journal***.

Hellmann, T. and Puri, M. 2000. The Interaction Between Product Market and Financing Strategy: The Role of Venture Capital. ***The Review of Financial Studies***, 13(4).

Hellmann, T. and Puri, M. 2002. Venture Capital and the Professionalization of Start-Up Firms: Empirical Evidence. ***The Journal of Finance***, 57(1): 169–197.

Hindle, K. and Klyver, K. 2007. Exploring the relationship between media coverage and participation in entrepreneurship: Initial global evidence and research implications. ***International Entrepreneurship and Management Journal***, 3(2): 217–242.

Hochberg, Y. V., Ljungqvist, A., and Lu, Y. 2007. Whom You Know Matters: Venture Capital Networks and Investment Performance. ***The Journal of Finance***, 62(1): 251–301.

Hochberg, Y. V., Ljungqvist, A., and Vissing-Jørgensen, A. 2014. Informational Holdup and Performance Persistence in Venture Capital. ***Review of Financial Studies***, 27(1): 102–152.

Hofstede, G. 1980. Culture and Organizations. ***International Studies of Management & Organization***, 10(4): 15–41.

Hofstede, G. H. 2001. ***Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations across Nations***. Thousand Oaks, Calif: Sage Publications, 2nd ed edition.

Hogg, M. A. 2000. Subjective Uncertainty Reduction through Self-categorization: A Motivational Theory of Social Identity Processes. ***European Review of Social Psychology***, 11(1): 223–255.

Hoogendoorn, S., Parker, S. C., and Van Praag, M. 2017. Smart or Diverse Start-up Teams? Evidence from a Field Experiment. ***Organization Science***, 28(6): 1010–1028.

Horverak, Ø. 2009. Research Note - Wine Journalism - Marketing or Consumers' Guide? ***Marketing Science***, 28(3): 573–579.

House, R. J. and Global Leadership and Organizational Behavior Effectiveness Research Program(Eds.). 2004. ***Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies***. Thousand Oaks, Calif: Sage Publications.

Hsu, D. H. 2004. What Do Entrepreneurs Pay for Venture Capital Affiliation? ***The Journal of Finance***, 59(4): 1805–1844.

Huckman, R. S., Staats, B. R., and Upton, D. M. 2009. Team Familiarity, Role Experience, and Performance: Evidence from Indian Software Services. ***Management Science***, 55(1): 85–100.

Huntington-Klein, N. 2023. Controls in Difference-in-Differences Don't Just Work. https://nickchk.substack.com/p/controls-in-difference-in-differences (Accessed: 2023-10-31).

Ingram, P. and Morris, M. W. 2007. Do People Mix at Mixers? Structure, Homophily, and the "Life of the Party". ***Administrative Science Quarterly***, 52(4): 558–585.

Jacobi, M. 2018. Media Judgment of Entrepreneurial Failure - Implications for Founders [Doctoral dissertation, Hamburg University of Technology]. ***TUHH Open Research***.

Joulin, A., Grave, E., Bojanowski, P., and Mikolov, T. 2016. Bag of Tricks for Efficient Text Classification.

Kaiser, U. and Mueller, B. 2013. Team Heterogeneity in Startups and its Development Over Time. ***SSRN Electronic Journal***.

Kamm, J. B., Shuman, J. C., Seeger, J. A., and Nurick, A. J. 1990. Entrepreneurial Teams in New Venture Creation: A Research Agenda. ***Entrepreneurship Theory and Practice***, 14(4): 7–17.

Kaniel, R. and Parham, R. 2017. WSJ Category Kings – The impact of media attention on consumer and mutual fund investment decisions. ***Journal of Financial Economics***, 123(2): 337–356.

Kaplan, S. N. and Schoar, A. 2005. Private Equity Performance: Returns, Persistence, and Capital Flows. ***The Journal of Finance***, 60(4): 1791–1823.

Karlsson, T. and Honig, B. 2009. Judging a business by its cover: An institutional perspective on new ventures and the business plan. ***Journal of Business Venturing***, 24(1): 27–45.

Kautonen, T., Van Gelderen, M., and Fink, M. 2015. Robustness of the Theory of Planned Behavior in Predicting Entrepreneurial Intentions and Actions. ***Entrepreneurship Theory and Practice***, 39(3): 655–674.

Kerr, W. R., Lerner, J., and Schoar, A. 2014. The Consequences of Entrepreneurial Finance: Evidence from Angel Financings. ***Review of Financial Studies***, 27(1): 20–55.

Kiessling, L., Radbruch, J., and Schaube, S. 2022. Self-Selection of Peers and Performance. ***Management Science***, 68(11): 8184–8201.

Kleinhempel, J., Klasing, M. J., and Beugelsdijk, S. 2022. Cultural Roots of Entrepreneurship: Evidence from Second-Generation Immigrants. ***Organization Science***, 34(5): 1800–1819.

Kortum, S. and Lerner, J. 2000. Assessing the Contribution of Venture Capital to Innovation. ***The RAND Journal of Economics***, 31(4): 674–692.

Kovacs, B. and Hannan, M. 2015. Conceptual Spaces and the Consequences of Category Spanning. ***Sociological Science***, 2: 252–286.

Krueger, N. F. and Carsrud, A. L. 1993. Entrepreneurial intentions: Applying the theory of planned behaviour. ***Entrepreneurship & Regional Development***, 5(4): 315–330.

Kulik, C. T., Bainbridge, H. T. J., and Cregan, C. 2008. Known by the Company We Keep: Stigma-By-Association Effects in the Workplace. ***Academy of Management Review***, 33(1): 216–230.

Lamertz, K. and Baum, J. A. C. 1998. The Legitimacy of Organizational Downsizing in Canada: An Analysis of Explanatory Media Accounts. ***Canadian Journal of Administrative Sciences / Revue Canadienne des Sciences de l'Administration***, 15(1): 93–107.

Lamin, A. and Zaheer, S. 2012. Wall Street vs. Main Street: Firm Strategies for Defending Legitimacy and Their Impact on Different Stakeholders. ***Organization Science***, 23(1): 47–66.

Lazear, E. P. 1998. ***Personnel Economics for Managers***. New York Weinheim: Wiley.

Lazear, E. P. 2005. Entrepreneurship. ***Journal of Labor Economics***, 23(4): 649–680.

Lee, S.-H., Peng, M. W., and Barney, J. B. 2007. Bankruptcy Law and Entrepreneurship Development: A Real Options Perspective. ***Academy of Management Review***, 32(1): 257–272.

Lee, S. H. and Wong, P. K. 2004. An exploratory study of technopreneurial intentions: A career anchor perspective. ***Journal of Business Venturing***, 19(1): 7–28.

Lerner, J. 1995. Venture Capitalists and the Oversight of Private Firms. ***The Journal of Finance***, 50(1): 301–318.

Lerner, J. and Malmendier, U. 2013. With a Little Help from My (Random) Friends: Success and Failure in Post-Business School Entrepreneurship. ***Review of Financial Studies***, 26(10): 2411–2452.

Lerner, J., Sorensen, M., and Strömberg, P. 2011. Private Equity and Long-Run Investment: The Case of Innovation. ***The Journal of Finance***, 66(2): 445–477.

Logue, D. and Grimes, M. 2022. Living Up to the Hype: How New Ventures Manage the Resource and Liability of Future-Oriented Visions within the Nascent Market of Impact Investing. ***Academy of Management Journal***, 65(3): 1055–1082.

Lopez, M. J. and Gutman, R. 2017. Estimation of Causal Effects with Multiple Treatments: A Review and New Ideas. ***Statistical Science***, 32(3): 432–454.

Lounsbury, M. and Glynn, M. A. 2001. Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources. ***Strategic Management Journal***, 22(6-7): 545–564.

Lovas, B. and Ghoshal, S. 2000. Strategy as guided evolution. ***Strategic Management Journal***, 21(9): 875–896.

Lubatkin, M. and Chatterjee, S. 1994. Extending Modern Portfolio Theory into the Domain of Corporate Diversification: Does It Apply? ***Academy of Management Journal***, 37(1): 109–136.

Ma-Kellams, C., Spencer-Rodgers, J., and Peng, K. 2011. I Am Against Us? Unpacking Cultural Differences in Ingroup Favoritism via Dialecticism. ***Personality and Social Psychology Bulletin***, 37(1): 15–27.

Madden, T. J., Ellen, P. S., and Ajzen, I. 1992. A Comparison of the Theory of Planned Behavior and the Theory of Reasoned Action. ***Personality and Social Psychology Bulletin***, 18(1): 3–9.

Manski, C. F. 1993. Identification of Endogenous Social Effects: The Reflection Problem. ***The Review of Economic Studies***, 60(3): 531–542.

Mantere, S., Aula, P., Schildt, H., and Vaara, E. 2013. Narrative attributions of entrepreneurial failure. ***Journal of Business Venturing***, 28(4): 459–473.

March, J. G. and Shapira, Z. 1987. Managerial Perspectives on Risk and Risk Taking. ***Management Science***, 33(11): 1404–1418.

McCombs, M. E. and Shaw, D. L. 1972. The Agenda-Setting Function of Mass Media. ***Public Opinion Quarterly***, 36(2): 176–187.

McDonald, D. 2004. Twentieth-Century Media Effects Research, 183–200. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc.

McDonald, R. M. and Eisenhardt, K. M. 2020. Parallel Play: Startups, Nascent Markets, and Effective Business-model Design. ***Administrative Science Quarterly***, 65(2): 483–523.

McDonnell, M.-H. and King, B. 2013. Keeping up Appearances: Reputational Threat and Impression Management after Social Movement Boycotts. ***Administrative Science Quarterly***, 58(3): 387–419.

McGee, J. E., Peterson, M., Mueller, S. L., and Sequeira, J. M. 2009. Entrepreneurial Self–Efficacy: Refining the Measure. ***Entrepreneurship Theory and Practice***, 33(4): 965–988.

McGrath, R. G. 1999. Falling Forward: Real Options Reasoning and Entrepreneurial Failure. ***The Academy of Management Review***, 24(1): 13–30.

McGrath, R. G., MacMillan, I. C., and Scheinberg, S. 1992. Elitists, risk-takers, and rugged individualists? An exploratory analysis of cultural differences between entrepreneurs and non-entrepreneurs. ***Journal of Business Venturing***, 7(2): 115–135.

McMullen, J. S. and Dimov, D. 2013. Time and the Entrepreneurial Journey: The Problems and Promise of Studying Entrepreneurship as a Process. ***Journal of Management Studies***, 50(8): 1481–1512.

McPherson, M., Smith-Lovin, L., and Cook, J. M. 2001. Birds of a Feather: Homophily in Social Networks. ***Annual Review of Sociology***, 27(1): 415–444.

McQuail, D. 1985. Sociology of Mass Communication. ***Annual Review of Sociology***, 11(1): 93–111.

Megginson, W. L. and Weiss, K. A. 1991. Venture Capitalist Certification in Initial Public Offerings. ***The Journal of Finance***, 46(3): 879–903.

Meyer, J. W. and Rowan, B. 1977. Institutionalized Organizations: Formal Structure as Myth and Ceremony. ***American Journal of Sociology***, 83(2): 340–363.

Molotch, H. and Lester, M. 1974. News as Purposive Behavior: On the Strategic Use of Routine Events, Accidents, and Scandals. ***American Sociological Review***, 39(1): 101–112.

Montgomery, J. M., Nyhan, B., and Torres, M. 2018. How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It. ***American Journal of Political Science***, 62(3): 760–775.

Mueller, J. S., Melwani, S., and Goncalo, J. A. 2012. The Bias Against Creativity: Why People Desire but Reject Creative Ideas. ***Psychological Science***, 23(1): 13–17.

Mutz, D. C., Pemantle, R., and Pham, P. 2019. The Perils of Balance Testing in Experimental Design: Messy Analyses of Clean Data. ***The American Statistician***, 73(1): 32–42.

Nahata, R. 2008. Venture capital reputation and investment performance. ***Journal of Financial Economics***, 90(2): 127–151.

Nanda, R. and Sørensen, J. B. 2010. Workplace Peers and Entrepreneurship. ***Management Science***, 56(7): 1116–1126.

Navis, C. and Glynn, M. A. 2010. How New Market Categories Emerge: Temporal Dynamics of Legitimacy, Identity, and Entrepreneurship in Satellite Radio, 1990–2005. ***Administrative Science Quarterly***, 55(3): 439–471.

Navis, C. and Glynn, M. A. 2011. Legitimate Distinctiveness and the Entrepreneurial Identity: Influence on Investor Judgments of New Venture Plausibility. ***Academy of Management Review***, 36(3): 479–499.

Ng, N., Yee, K., Baevski, A., Ott, M., Auli, M., and Edunov, S. 2019. Facebook FAIR's WMT19 News Translation Task Submission.

Oliver, C. 1991. Strategic Responses to Institutional Processes. ***The Academy of Management Review***, 16(1): 145–179.

Owens, D., Grossman, Z., and Fackler, R. 2014. The Control Premium: A Preference for Payoff Autonomy. ***American Economic Journal: Microeconomics***, 6(4): 138–161.

Packard, M. D., Clark, B. B., and Klein, P. G. 2017. Uncertainty Types and Transitions in the Entrepreneurial Process. ***Organization Science***, 28(5): 840–856.

Parker, S.(Ed.). 2006. ***The Life Cycle of Entrepreneurial Ventures***. Number 3 in International Handbook Series on Entrepreneurship. Boston, MA: Springer Science+Business Media, Inc.

Petkova, A. P., Rindova, V. P., and Gupta, A. K. 2013. No News Is Bad News: Sensegiving Activities, Media Attention, and Venture Capital Funding of New Technology Organizations. ***Organization Science***, 24(3): 865–888.

Pollock, T. G., Chen, G., Jackson, E. M., and Hambrick, D. C. 2010. How much prestige is enough? Assessing the value of multiple types of high-status affiliates for young firms. ***Journal of Business Venturing***, 25(1): 6–23.

Pollock, T. G. and Gulati, R. 2007. Standing out from the crowd: The visibility-enhancing effects of IPO-related signals on alliance formation by entrepreneurial firms. ***Strategic Organization***, 5(4): 339–372.

Pollock, T. G. and Rindova, V. P. 2003. Media Legitimation Effects in the Market for Initial Public Offerings. ***Academy of Management Journal***, 46(5): 631–642.

Portyanko, S., Reinmoeller, P., Hussels, S., and Turner, N. 2023. Peer effects and intentional entrepreneurial behaviour: A systematic literature review and research agenda. ***International Journal of Management Reviews***, 25(3): 515–545.

Puri, M. and Zarutskie, R. 2012. On the Life Cycle Dynamics of Venture-Capital- and Non-Venture-Capital-Financed Firms. ***The Journal of Finance***, 67(6): 2247–2293.

Quas, A., Martí, J., and Reverte, C. 2021. What money cannot buy: A new approach to measure venture capital ability to add non-financial resources. ***Small Business Economics***, 57(3): 1361–1382.

Raden, D. 2003. Ingroup Bias, Classic Ethnocentrism, and Non-Ethnocentrism Among American Whites. ***Political Psychology***, 24(4): 803–828.

Radu-Lefebvre, M., Loué, C., and Redien-Collot, R. 2019. Contextualizing Entrepreneurial Legitimacy: The Interplay of Internal and External Social Representations. ***Journal of Enterprising Culture***, 27(01): 1–33.

Reagans, R., Zuckerman, E., and McEvily, B. 2004. How to Make the Team: Social Networks vs. Demography as Criteria for Designing Effective Teams. ***Administrative Science Quarterly***, 49(1): 101–133.

Robertson, C. E., Pröllochs, N., Schwarzenegger, K., Pärnamets, P., Van Bavel, J. J., and Feuerriegel, S. 2023. Negativity drives online news consumption. ***Nature Human Behaviour***, 7(5): 812–822.

Rosenzweig, P. M. 1994. When Can Management Science Research Be Generalized Internationally? ***Management Science***, 40(1): 28–39.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. 2023. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. ***Journal of Econometrics***, 235(2): 2218–2244.

Rozado, D., Hughes, R., and Halberstadt, J. 2022. Longitudinal analysis of sentiment and emotion in news media headlines using automated labelling with Transformer language models. ***PLOS ONE***, 17(10): 1–14.

Rozin, P. and Royzman, E. B. 2001. Negativity Bias, Negativity Dominance, and Contagion. ***Personality and Social Psychology Review***, 5(4): 296–320.

Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. ***Journal of Educational Psychology***, 66(5): 688–701.

Ruef, A. and Markard, J. 2010. What happens after a hype? How changing expectations affected innovation activities in the case of stationary fuel cells. ***Technology Analysis & Strategic Management***, 22(3): 317–338.

Ruef, M., Aldrich, H. E., and Carter, N. M. 2003. The Structure of Founding Teams: Homophily, Strong Ties, and Isolation among U.S. Entrepreneurs. ***American Sociological Review***, 68(2): 195–222.

Sacerdote, B. 2011. Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? In ***Handbook of the Economics of Education***, volume 3, 249–277. Elsevier.

Sant'Anna, P. H. and Zhao, J. 2020. Doubly robust difference-in-differences estimators. ***Journal of Econometrics***, 219(1): 101–122.

Schultz, M., Mouritsen, J., and Gabrielsen, G. 2001. Sticky Reputation: Analyzing a Ranking System. ***Corporate Reputation Review***, 4(1): 24–41.

Schwartz, S. 2006. A Theory of Cultural Value Orientations: Explication and Applications. ***Comparative Sociology***, 5(2-3): 137–182.

Seabright, M. A., Levinthal, D. A., and Fichman, M. 1992. Role of Individual Attachments in the Dissolution of Interorganizational Relationships. ***Academy of Management Journal***, 35(1): 122–160.

Semadeni, M., Cannella Jr., A. A., Fraser, D. R., and Lee, D. S. 2008. Fight or flight: Managing stigma in executive careers. ***Strategic Management Journal***, 29(5): 557–567.

Seru, A. 2014. Firm boundaries matter: Evidence from conglomerates and R&D activity. ***Journal of Financial Economics***, 111(2): 381–405.

Shane, S. 1993. Cultural influences on national rates of innovation. ***Journal of Business Venturing***, 8(1): 59–73.

Shane, S. 2009. Why encouraging more people to become entrepreneurs is bad public policy. ***Small Business Economics***, 33(2): 141–149.

Shane, S. A. 1992. Why do some societies invent more than others? ***Journal of Business Venturing***, 7(1): 29–46.

Shapero, A. and Sokol, L. 1982. The social dimensions of entrepreneurship. In ***Encyclopedia of Entrepreneurship***. Englewood Cliffs, NJ: Prentice-Hall.

Sheng, J. and Lan, H. 2019. Business failure and mass media: An analysis of media exposure in the context of delisting event. ***Journal of Business Research***, 97: 316–323.

Shepard, R. N. 1987. Toward a Universal Law of Generalization for Psychological Science. ***Science***, 237(4820): 1317–1323.

Shepherd, D. A. 2003. Learning from Business Failure: Propositions of Grief Recovery for the Self-Employed. ***The Academy of Management Review***, 28(2): 318–328.

Shepherd, D. A. and Haynie, J. M. 2011. Venture failure, stigma, and impression management: A self-verification, self-determination view. ***Strategic Entrepreneurship Journal***, 5(2): 178–197.

Shepherd, D. A. and Patzelt, H. 2015. Harsh Evaluations of Entrepreneurs Who Fail: The Role of Sexual Orientation, Use of Environmentally Friendly Technologies, and Observers' Perspective Taking: Harsh Evaluations of Entrepreneurs Who Fail. ***Journal of Management Studies***, 52(2): 253–284.

Shepherd, D. A. and Patzelt, H. 2017. ***Trailblazing in Entrepreneurship***. Cham: Springer International Publishing.

Shepherd, D. A., Patzelt, H., and Wolfe, M. 2011. Moving Forward from Project Failure: Negative Emotions, Affective Commitment, and Learning from the Experience. ***Academy of Management Journal***, 54(6): 1229–1259.

Simmons, S. A., Wiklund, J., and Levie, J. 2014. Stigma and business failure: Implications for entrepreneurs' career choices. ***Small Business Economics***, 42(3): 485–505.

Singh, J. V., Tucker, D. J., and House, R. J. 1986. Organizational Legitimacy and the Liability of Newness. ***Administrative Science Quarterly***, 31(2): 171–193.

Solomon, D. H., Soltes, E., and Sosyura, D. 2014. Winners in the spotlight: Media coverage of fund holdings as a driver of flows. ***Journal of Financial Economics***, 113(1): 53–72.

Sørensen, M. 2007. How Smart Is Smart Money? A Two-Sided Matching Model of Venture Capital. ***The Journal of Finance***, 62(6): 2725–2762.

Soroka, S., Fournier, P., and Nir, L. 2019. Cross-national evidence of a negativity bias in psychophysiological reactions to news. ***Proceedings of the National Academy of Sciences***, 116(38): 18888–18892.

Splawa-Neyman, J., Dabrowska, D. M., and Speed, T. P. 1990. On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. ***Statistical Science***, 5(4): 465–472.

Steffens, P., Terjesen, S., and Davidsson, P. 2012. Birds of a feather get lost together: New venture team composition and performance. ***Small Business Economics***, 39(3): 727–743.

Stephan, U. 2022. Cross-Cultural Innovation and Entrepreneurship. ***Annual Review of Organizational Psychology and Organizational Behavior***, 9(1): 277–308.

Stephan, U. and Uhlaner, L. M. 2010. Performance-based vs socially supportive culture: A cross-national study of descriptive norms and entrepreneurship. ***Journal of International Business Studies***, 41(8): 1347–1364.

Stinchcombe, A. L. 2000. Social structure and organizations. In ***Advances in Strategic Management***, volume 17, 229–259. Bingley: Emerald (MCB UP).

Storper, M. and Venables, A. J. 2004. Buzz: Face-to-face contact and the urban economy. ***Journal of Economic Geography***, 4(4): 351–370.

Sumner, W. G. 2002. ***Folkways: A Study of Mores, Manners, Customs and Morals***. Mineola, N.Y: Dover Publications.

Sun, L. and Abraham, S. 2020. Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects.

Sutton, R. I. and Callahan, A. L. 1987. The Stigma of Bankruptcy: Spoiled Organizational Image and Its Management. ***Academy of Management Journal***, 30(3): 405–436.

Tartari, V., Perkmann, M., and Salter, A. 2014. In good company: The influence of peers on industry engagement by academic scientists. ***Research Policy***, 43(7): 1189–1203.

Taylor, M. Z. and Wilson, S. 2012. Does culture still matter?: The effects of individualism on national innovation rates. ***Journal of Business Venturing***, 27(2): 234–247.

Tedeschi, J. T. 1981. ***Impression Management Theory and Social Psychological Research***. New York: Academic Press.

Terjesen, S., Hessels, J., and Li, D. 2016. Comparative International Entrepreneurship: A Review and Research Agenda. ***Journal of Management***, 42(1): 299–344.

Terry, D. J., Hogg, M. A., and White, K. M. 1999. The theory of planned behaviour: Self-identity, social identity and group norms. ***British Journal of Social Psychology***, 38(3): 225–244.

Thomas, A. S. and Mueller, S. L. 2000. A Case for Comparative Entrepreneurship: Assessing the Relevance of Culture. ***Journal of International Business Studies***, 31(2): 287–301.

Thompson, L. L. 2018. ***Making the Team: A Guide for Managers***. New York: Pearson Education, sixth edition edition.

Thornhill, S. and Amit, R. 2003. Learning About Failure: Bankruptcy, Firm Age, and the Resource-Based View. ***Organization Science***, 14(5): 497–509.

Tiessen, J. H. 1997. Individualism, collectivism, and entrepreneurship: A framework for international comparative research. ***Journal of Business Venturing***, 12(5): 367–384.

Timmons, J. A. 1990. ***New Venture Creation: Entrepreneurship in the 1990s***. The Irwin Series in Management and the Behavioral Sciences. Homewood, Ill: Irwin, 3. ed edition.

Tornikoski, E. 2009. Legitimating behaviors and firm emergence: A resource dependence perspective. ***International Entrepreneurship and Management Journal***, 5(2): 121–138.

Tuchman, G. 1978. ***Making News: A Study in the Construction of Reality***. New York: Free Press.

Tzabbar, D. and Margolis, J. 2017. Beyond the Startup Stage: The Founding Team's Human Capital, New Venture's Stage of Life, Founder–CEO Duality, and Breakthrough Innovation. ***Organization Science***, 28(5): 857–872.

Überbacher, F. 2014. Legitimation of New Ventures: A Review and Research Programme: Legitimation of New Ventures. ***Journal of Management Studies***, 51(4): 667–698.

Ucbasaran, D., Shepherd, D. A., Lockett, A., and Lyon, S. J. 2013. Life After Business Failure: The Process and Consequences of Business Failure for Entrepreneurs. ***Journal of Management***, 39(1): 163–202.

Uzzi, B., Mukherjee, S., Stringer, M., and Jones, B. 2013. Atypical Combinations and Scientific Impact. ***Science***, 342(6157): 468–472.

Van Lente, H., Spitters, C., and Peine, A. 2013. Comparing technological hype cycles: Towards a theory. ***Technological Forecasting and Social Change***, 80(8): 1615–1628.

Vanacker, T. and Forbes, D. P. 2016. Disentangling the Multiple Effects of Affiliate Reputation on Resource Attraction in New Firms. ***Organization Science***, 27(6): 1525–1547.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. 2017. Attention Is All You Need.

Venkataraman, S. 2019. Chapter 1 The Distinctive Domain of Entrepreneurship Research. In Katz, J. A. and Corbet, A. C.(Eds.), ***Advances in Entrepreneurship, Firm Emergence and Growth***, volume 21, 5–20. Emerald Publishing Limited.

Vesper, K. H. 1980. ***New Venture Strategies***. Englewood Cliffs, N. J: Prentice-Hall.

Von Hippel, E. 1994. “Sticky Information” and the Locus of Problem Solving: Implications for Innovation. ***Management Science***, 40(4): 429–439.

Wagner, C. S., Whetsell, T. A., and Mukherjee, S. 2019. International research collaboration: Novelty, conventionality, and atypicality in knowledge recombination. ***Research Policy***, 48(5): 1260–1270.

Wiesenfeld, B. M., Wurthmann, K. A., and Hambrick, D. C. 2008. The Stigmatization and Devaluation of Elites Associated with Corporate Failures: A Process Model. ***Academy of Management Review***, 33(1): 231–251.

Wiklund, J., Baker, T., and Shepherd, D. 2010. The age-effect of financial indicators as buffers against the liability of newness. ***Journal of Business Venturing***, 25(4): 423–437.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., Von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., Drame, M., Lhoest, Q., and Rush, A. 2020. Transformers: State-of-the-Art Natural Language Processing. In ***Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations***, 38–45, Online. Association for Computational Linguistics.

Wooldridge, J. M. 2021. Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators. ***SSRN Electronic Journal***.

Yang, L., Holtz, D., Jaffe, S., Suri, S., Sinha, S., Weston, J., Joyce, C., Shah, N., Sherman, K., Hecht, B., and Teevan, J. 2021. The effects of remote work on collaboration among information workers. ***Nature Human Behaviour***, 6(1): 43–54.

Younger, S. and Fisher, G. 2020. The exemplar enigma: New venture image formation in an emergent organizational category. ***Journal of Business Venturing***, 35(1-18): 105897.

Zavyalova, A., Pfarrer, M. D., and Reger, R. K. 2017. Celebrity *and* Infamy? The Consequences of Media Narratives About Organizational Identity. ***Academy of Management Review***, 42(3): 461–480.

Zavyalova, A., Pfarrer, M. D., Reger, R. K., and Shapiro, D. L. 2012. Managing the Message: The Effects of Firm Actions and Industry Spillovers on Media Coverage Following Wrongdoing. ***Academy of Management Journal***, 55(5): 1079–1101.

Zimmerman, M. A. and Zeitz, G. J. 2002. Beyond Survival: Achieving New Venture Growth by Building Legitimacy. ***The Academy of Management Review***, 27(3): 414–431.

Appendices

Appendix A

Data Cleaning

Merging of PitchBook and Crunchbase data. The steps applied in this paragraph apply to both the second and third study, which both base on the sample of PitchBook and Crunchbase data. We combine the startups database from PitchBook and Crunchbase following a outer join logic, i.e. we look for matches in both data sets but also add all non-matched companies. To make sure, that the match is valid, we check also for geographic equivalence. Still, for some matches we obtain ambiguous result for some describing characteristics such as closure date, country or status. In the following, we describe how we handle such cases. When matched startups are not unambiguously assigned to a particular country, we drop these (947) startups from the sample. When there are disagreements reagarding the closure date, we proceed with the earlier closure date. We drop startups from the sample if has been sold or acquired.

Improving coverage of failures. Because the number of closed startups is lower that what we would expect, we complement the data database information by manually searching for closure dates. We search for all German startups using a combination of Handelsregister, deutsche-startups.de, and additionally, if we suspect that a startup is closed, e.g. because the website does not exist anymore, a manual search using NorthData and Google. The Handelsregister is made available by "Offene Register", an initiative to make the Handelsregister accessible using an API instead of the tedious manual search needed on the official site. For the American and British startups, we make use of the OpenCorporates, which contains information from many register around the world, including US registers. When the identified closure date is earlier than the one provided by our initial data base, we proceed with the earlier date. Through this approach, we add/update:

- United States: [added 9251, corrected 2536]
- United Kingdom: [added 283, corrected 35]
- Germany: [added 206, corrected 35]

Joining CoreSignal data. To achieve more comprehensive coverage of founder and employee data, we integrated a dataset provided by CoreSignal, which leverages LinkedIn data. While the absence of a unique identifier posed a challenge, domain information was utilized as an alternative matching variable.

Given the lack of a unique identifier, the matching process involved multiple steps, each assigned a weight based on the reliability of the match. The matching criteria, listed in order of decreasing priority, were:

- Domain (URL),
- LinkedIn handle, and
- Startup name.

Matches were often established across multiple criteria, enhancing the confidence in their validity. To further refine the process, geographical information was cross-verified. Specifically, matches based solely on the startup name were accepted only if supported by geographical consistency, with a higher priority given to matches within the same city rather than the same region.

Classifying publication outlets. We enrichen the news data by adding the country of the publishing outlet. Only for a fraction, LexisNexis provides the country of publication. We rely on external data sources to match the publication outlets by name. The sources are as follows (latest accessed at 2024-12-20).

United States:

- `https://en.wikipedia.org/wiki/Category:American_news_websites`
- `https://en.wikipedia.org/wiki/News_media_in_the_United_States`
- `https://en.wikipedia.org/wiki/Category:News_agencies_based_in_the_United_States`
- `https://en.wikipedia.org/wiki/List_of_weekly_newspapers_in_the_United_States`

- `https://en.wikipedia.org/wiki/Category:24-hour_television_news_channels_in_the_United_States`
- `https://en.wikipedia.org/wiki/Category:News_magazines_published_in_the_United_States`

United Kingdom:

- `https://en.wikipedia.org/wiki/List_of_newspapers_in_the_United_Kingdom`
- `https://en.wikipedia.org/wiki/List_of_magazines_in_the_United_Kingdom`

Germany:

- `https://www.bdzv.de/zeitungen-in-deutschland`
- `https://de.wikipedia.org/wiki/Liste_deutscher_Zeitungen`
- `https://de.wikipedia.org/wiki/Liste_deutscher_Zeitschriften`
- `https://de.wikipedia.org/wiki/Liste_deutscher_Regionalzeitungen`
- `https://de.wikipedia.org/wiki/Liste_von_Wirtschaftsmagazinen`
- `https://www.deutschland.de/de/topic/wissen/nachrichten`
- `https://de.wikipedia.org/wiki/Fachzeitschrift`

To identify press releases, which we exclude from the main analysis, we relied on pattern search. Most PR agencies have recurring patterns in their names that clearly identify them. The regular expression pattern we used is:

```
"wire|release|pressemitteilung|presseportal|\bots\b|\bnews aktuell\b|\bpr\b"
```

Other press release agencies we identified are:

- US Official News
- US Fed News
- Market News Publishing
- Mynewsdesk
- European Union News

Cleaning of text. To remove irrelevant sentences, we set the minimum and maximum length of the sentence to 4 and 80 words, respectively and we manually create patterns to identify sentences containing

- more than one financial symbols: text likely to be a factual financial statement
- at least one patent symbol: sentence likely to be a patent announcement
- more than 100 capital letters: sentence likely not from news article
- e-mail address included: sentence likely not from news article
- URL included: sentence likely not from news article.

Appendix B

Analysis of Not-Yet-Treated Sample

We apply a classical ordinary least square regression and incorporate the variables and fixed effects described in the data section. We only include startups that underwent failure before the failure event.

TABLE B.1: Results of OLS Regression on Not-Yet Sample

	Avg. Sentiment	Log. # Neutral Sentences	Log. # Negative Sentences	Log. # Positive Sentences
(Media) DEU : (Startup) DEU	−0.06***	0.03	0.00	−0.15***
	(0.00)	(0.03)	(0.02)	(0.02)
(Media) DEU : (Startup) GBR	−0.07***	−0.42***	−0.10***	−0.48***
	(0.01)	(0.05)	(0.02)	(0.04)
(Media) DEU : (Startup) USA	−0.09***	−0.42***	−0.06***	−0.52***
	(0.00)	(0.02)	(0.01)	(0.01)
(Media) GBR : (Startup) DEU	0.00	−0.35***	−0.06**	−0.23***
	(0.01)	(0.03)	(0.02)	(0.03)
(Media) GBR : (Startup) GBR	0.04***	−0.14***	0.00	0.07***
	(0.00)	(0.02)	(0.01)	(0.02)
(Media) GBR : (Startup) USA	0.01**	−0.37***	−0.06***	−0.26***
	(0.00)	(0.01)	(0.00)	(0.01)
Num. of Articles	0.00*			
	(0.00)			
Age in Quarters	0.00**	−0.01***	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Num. of Founder	0.00+	0.01***	0.00*	0.01+
	(0.00)	(0.00)	(0.00)	(0.00)
Num. of Female	0.00	−0.01***	0.00*	−0.01**
	(0.00)	(0.00)	(0.00)	(0.00)
Num. of Experienced	0.00+	0.01***	0.00***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
Num. of PhD	0.00	−0.01**	−0.01***	−0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
Num. of Master	0.00+	−0.01**	0.00*	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)
Num. of Bachelor	0.00	0.00	0.00	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)
Log Raised Amount	0.01***	0.14***	0.04***	0.14***
	(0.00)	(0.01)	(0.01)	(0.01)
VC Entry	0.01***	−0.08***	−0.04***	−0.04**
	(0.00)	(0.02)	(0.01)	(0.01)
Num.Obs.	276 300	276 300	276 300	276 300
R^2	0.024	0.068	0.028	0.069
R^2 Adj.	0.023	0.068	0.027	0.069
Mean DV	0.297	1.484	0.186	0.882
Quarter FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Series FE	Yes	Yes	Yes	Yes

Notes. This table presents results from ordinary least squares regressions with various outcomes of media sentiment. The "not-yet" sample excludes never-treated startups (i.e., those that never fail). Publication quarter industry, and funding round fixed effects are included. Interaction terms distinguish publication country, denoted as (Media), and startup location, denoted as (Startup). For the countries, USA refers to United States, GBR to United Kingdom, and DEU to Germany as defined by ISO 3166-1 alpha-3. Standard errors, clustered at the company level, are reported in parentheses. Significance levels are indicated as follows: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

When atypicality is computed globally and not within industry, we get a very similar result. The results in Table B.2 are very close to the one in the main analysis in Table 4.3.

TABLE B.2: Heterogeneous Treatment Effects of Entrepreneurial Failure on Media Coverage

	Avg. Sentiment	Log. # Neutral Sentences	Log. # Negative Sentences	Log. # Positive Sentences
Atypicality $q_{[0\%-25\%]}$	-0.15 ***	0.19 ***	0.22 ***	-0.15 **
	(0.02)	(0.04)	(0.04)	(0.05)
Atypicality $q_{[25\%-50\%]}$	-0.13 ***	0.25 ***	0.24 ***	-0.07 .
	(0.01)	(0.04)	(0.02)	(0.04)
Atypicality $q_{[50\%-75\%]}$	-0.12 ***	0.16 ***	0.14 ***	-0.10 *
	(0.01)	(0.03)	(0.03)	(0.04)
Atypicality $q_{[75\%-100\%]}$	-0.11 ***	0.31 ***	0.21 ***	0.04
	(0.01)	(0.04)	(0.04)	(0.03)
Num.Obs.	282,178	282,178	282,178	282,178
Mean DV	0.294	1.488	0.191	0.88
Media Country X Quarter FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Series FE	Yes	Yes	Yes	Yes
Controls in First Stage	Yes (Funding + Founders)			

Notes. This table presents results from a model with treatment effect interactions 1 quarter of anticipation. Both main effects of publication quarter and publication country and the interaction and industry fixed effects are included. Standard errors are clustered at the company level. Control variables are included in the first stage. Stars indicate significance level: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Anticipation is set to 1.