Common Ownership and Innovation Efficiency

Xuelin Li

Tong Liu

Lucian A. Taylor*

February 18, 2021

Abstract

How does common ownership affect innovation? We study this question using project-level data on pharmaceutical startups and their venture capital (VC) investors. We find that common VC ownership reduces duplication of R&D in patent races. Specifically, common ownership leads VCs to shut down lagging drug projects, withhold funding from lagging startups, and redirect those startups' innovation. These results support theories dating back to Loury (1979). By coordinating R&D efforts across firms in a patent race, a common owner can reduce excess R&D and thereby help solve a market failure. Consistent with common ownership improving innovation efficiency, common ownership rates are positively correlated with the number of approved drugs per dollar of aggregate VC funding.

Key words: Common ownership, innovation, venture capital, healthcare

*Li is at the University of South Carolina (Darla Moore). Liu and Taylor are at the University of Pennsylvania (Wharton). Emails: xuelin.li@moore.sc.edu, tongl@wharton.upenn.edu, luket@wharton.upenn.edu. We are grateful for comments from Bruno Biais, Mike Ewens, Joan Farre-Mensa, Jillian Grennan, Ernest Liu, Adrien Matray, Martin Schmalz, Antoinette Schoar, Da Tian, and Vish Viswanathan; conference participants at the American Finance Association Meeting, IPC Spring Research Symposium, Midwest Finance Association Meeting, Northern Finance Association Meeting, University of Minnesota Corporate Finance Conference, and Workshop on Entrepreneurial Finance and Innovation; and seminar participants at IDC Herzilya, Queen's University, and the University of Pennsylvania. A previous version was titled "Do Venture Capitalists Stifle Competition?"

Introduction

Competing firms sometimes share a large investor in common. Common ownership has grown significantly in recent years, leading to debates about its welfare implications. Recent papers suggest a welfare loss due to reduced product-market competition (e.g., Azar et al., 2018). However, common ownership may also have social benefits, including through its effects on innovation.

In this paper, we study how common ownership affects innovation. We find that common ownership leads investors to shut down lagging R&D projects, restrict their funding, and encourage their firms to pivot to new R&D projects. In this manner, common owners help reduce duplication of R&D efforts in patent races. Theories dating back to Loury (1979) show that if firms compete fiercely in a patent race, they invest more in R&D than is socially optimal. A common owner can help solve this market failure by coordinating firms and reducing their excess R&D. Consistent with common ownership improving innovation efficiency, we show that common ownership is associated with more R&D output per dollar of R&D funding.

We study these issues in the venture capital (VC) setting. This setting is important, as VC-backed startups generate a large share of the innovation in our economy (e.g., Gompers and Lerner, 2004, Kaplan and Lerner, 2010). Also, the effects of common ownership should be especially strong for VC investors. VCs are sophisticated, active investors who hold concentrated portfolios, so they are likely to be attentive to spillovers across portfolio companies. VCs own large equity stakes and have strong control rights in their startups, so they have the power to influence those companies. We also find that common ownership by VCs is widespread: 39% of startups in our sample have a close competitor with a shared VC investor. For all these reasons, one might expect the effects of common ownership to be even stronger in the VC setting than, for example, in the setting of passive index funds holding modest stakes in public companies.

Our main testable prediction is from Grossman and Shapiro (1987), an extension of Loury (1979). They model firms engaged in a multi-stage patent race, and they ask how a firm responds after seeing its competitor reach an intermediate milestone. The theory implies that if both firms share the same owner, then the lagging firm's project is likely to be shut down,

because the common owner wishes to avoid the inefficient duplication of R&D efforts. If the firms instead have different owners, they fail to internalize the negative spillovers they impose on each other. The lagging project is therefore likely to continue, even if it is socially suboptimal.

Motivated by this theory, we predict that common ownership increases the likelihood that a lagging project is held back. Our ideal experiment would feature two pairs of competing startups. We would randomly assign one pair to share a common VC investor and the other pair to not. One "pioneer" startup in each pair would reach an intermediate milestone, and we would compare the outcomes of the two pairs' "lagging" startups. Following Grossman and Shapiro (1987), we predict that the lagging startup is less likely to continue progressing in the pair that shares a common VC.

To apply this ideal experiment to real data, we need to identify startups that are natural competitors, and we need publicly observable intermediate milestones. We overcome this challenge by using project-level data on pharmaceutical startups. Our data cover 1,045 Phase I drug projects conducted by 481 U.S. startups financed by 764 VC firms, from 2005 to 2018. We work at the level of VC firms, so "VC" stands for a VC firm rather than a VC fund or partner. Our data partition the pharmaceutical industry into 78 highly detailed product markets, so we can compare, for example, a pair of startups with competing influenza/pneumonia projects to a pair of startups with competing arthritis projects. Regulation by the U.S. Food and Drug Administration (FDA) provides a clear, publicly observed intermediate milestone: seeing a drug project progress from Phase I to Phase II clinical trials. Besides having these useful properties, our data cover a sector of the economy that is both highly valuable and important for social welfare.

Another challenge is finding quasi-random variation in common ownership. We apply an instrumental variable (IV) approach that exploits the local nature of VC investing. VCs tend to invest in nearby companies in order to reduce the costs of search and monitoring. Our IV for whether two startups share a common VC is based on the startups' geographic proximity. The main identification assumption is that geographic proximity affects our dependent variable—the outcome at a lagging startup after a competing pioneer startup makes progress—only through

the effect of proximity on whether the two startups share a common VC.

Our first tests examine the probability that a Phase I drug project progresses to Phase II after seeing a closely competing project—a pioneer—progress to Phase II. Consistent with the prediction above, we find that sharing a VC in common with the pioneer makes the lagging project less likely to progress to Phase II. More simply, common ownership leads VCs to hold back lagging projects. Economic significance is high: The average effect of having a shared VC is comparable to the unconditional rate of progressing to Phase II. Our result holds in ordinary least squares (OLS) regressions as well as IV and bivariate probit regressions. The result obtains even if we include fixed effects for time by drug category, which amounts to comparing how two lagging arthritis projects (for example) react differently to seeing a third arthritis project reach Phase II, depending on whether the two lagging projects share a VC with the pioneer.

How does sharing a common VC affect drug projects' outcomes? We find evidence of a VC financing mechanism. A VC can hold back a drug project by withholding follow-on funding. We predict that after a VC sees a startup make progress, the VC is less likely to extend funding to a closely competing startup if both startups are in its portfolio. The data strongly support this prediction, with high levels of statistical and economic significance. We find evidence of a VC financing mechanism even if we compare across different VCs invested in the same startup and quarter, which effectively controls for a startup's demand for funding. Specifically, we find that relative to other VCs invested in the same startup and quarter, a given VC is less likely to make a follow-on investment if that VC also owns a stake in a close competitor that has recently made progress. Further, we show that when a common VC abandons a startup, other VCs do not step in to fill the financing hole, so the startup is less likely to raise financing from *any* VC.

We find interesting heterogeneity in these financing results. Results are stronger for VCs with larger ownership stakes in the startup. Such VCs have stronger control rights, making it more feasible for them to hold back projects. We also find stronger results for less-diversified VCs. This result is consistent with Gilje et al. (2019), who show that common owners must be highly attentive in order for them to internalize spillovers within their portfolios. Results are also stronger if the lagging and pioneering project are more technologically similar or belong to

a more narrowly defined drug category. Such projects are more likely to be close competitors, so the patent-race story is more likely to hold.

We also show that common owners redirect lagging firms' innovation activities. When a startup sees a commonly owned competitor make progress in a one drug category, the startup pivots away from that category toward non-overlapping categories. Specifically, the lagging startup is more likely to initiate new projects in non-overlapping drug categories, repurpose existing projects into those categories, and form financing alliances with large pharma companies in those categories. By redirecting the lagging firms' innovation efforts in this way, the common owner again helps avoid the duplication of R&D efforts.

To the extent that common ownership reduces unnecessary duplication of R&D costs, common ownership can improve innovation efficiency. Consistent with this story, we find that common ownership rates are positively correlated with a measure of innovation efficiency in the cross section of drug categories. We measure innovation efficiency as the total number of drugs reaching FDA approval within a drug category, scaled by the total amount of VC funding provided to all startups active in that category. This efficiency proxy is strongly correlated with the drug category's common ownership rate, with high levels of statistical and economic significance. This is a simple, descriptive result that does not have a causal interpretation. The result is consistent, however, with common ownership helping to avoid excess duplication of R&D, producing more approved drugs per dollar of aggregate R&D.

This efficiency result suggests that common ownership has a social benefit. Common ownership can also impose social costs, however. By holding back lagging drug projects, common owners can reduce future product market competition, thereby raising drug prices and shrinking consumers' choice sets. Also, our tests examine a specific counterfactual in which we hold the set of projects or startups fixed while changing the ownership structure from non-common to common. Common ownership can also influence which projects or startups exist in the first place, which affects welfare.[1] Quantifying all these effects to better understand common ownership's welfare implications is an important area for future research.

[1]Newham et al. (2018) and Gerakos and Xie (2019) show that common ownership reduces entry by generic drugs. Unlike us, they focus on public firms, mature products, non-VC investors.

We address various challenges to our identification strategy. For example, one potential concern is that geographic proximity reflects the ease of poaching employees from a competing startup, which in theory could explain why a project is more likely to fail after a nearby competing project gains an edge. Inconsistent with that story, however, our results are not weaker when the startup's state strongly enforces employee non-compete agreements. We also analyze potential concerns related to technological proximity and information sharing.

Our main contribution is to document a new, important way in which common ownership affects innovation. Many papers study how common ownership affects product market competition,[2] but relatively few study its effects on innovation. He and Huang (2017) show that common ownership improves firm-level innovation productivity, which relates to the positive correlation we document between common ownership and aggregate innovation efficiency. They conjecture that innovation productivity improves because of resource sharing and R&D coordination across firms. Supporting that conjecture, Kostovetsky and Manconi (2020) find that common institutional ownership facilitates innovation diffusion among portfolio firms. The mechanism in this paper is quite different: Innovation efficiency improves because a common owner avoids duplicating R&D costs and redirects innovation at lagging firms. Cunningham et al. (2020) study "killer" acquisitions of competing drug companies. Similar to us, they show that an acquired drug project is less likely to progress if the acquirer has a similar drug project. They study direct ownership, not common ownership by a financial intermediary.

A few papers study spillovers within VC portfolios. Lindsey (2008) finds that startups sharing a common VC investor are more likely to form alliances, and González-Uribe (2019) finds they are more likely to share innovation resources. The mechanism we study is quite different, especially its focus on close competitors within a VC's portfolio. Closer to our paper, Eldar et al. (2020) ask whether common VC ownership affects startup performance. They show that common ownership improves information sharing among startups, which increases

[2]He and Huang (2017), Azar et al. (2018), and Azar et al. (2019) find the common ownership reduces product market competition. Opposing evidence comes from Gramlich and Grundl (2017), Kennedy et al. (2017), Lewellen and Lewellen (2017), Kini et al. (2018), Dennis et al. (2019), Lewellen and Lowry (2019), Gilje et al. (2019), and Koch et al. (2020). Gutiérrez and Philippon (2017) show a negative relation between common ownership and corporate investment, which relates to our finding that common ownership leads VCs to hold back projects and funding.

their growth and odds of success. Like us, they find that common VC ownership has social benefits, but the mechanism is clearly different. One advantage of Eldar et al. (2020) is that their data span all industries, making it easier to generalize results. An advantage of using only pharmaceutical industry data, however, is that it lets us better identify firms and projects that are close competitors. For example, we assume an arthritis drug only competes with other arthritis drugs. Eldar et al. (2020) instead use a coarse industry classification that, for instance, combines all pharmaceutical companies into one category.

Our results also shed light on theories of common ownership and innovation. López and Vives (2019) show that a common owner internalizes firms' positive R&D spillovers, which can increase R&D and improve welfare. They predict that common ownership increases R&D if and only if R&D spillovers are sufficiently high. Our finding that common owners hold back lagging projects, viewed through the lens of their theory, suggest that R&D spillovers are small in our setting. This result makes sense. If two startups are engaged in a winner-take-all patent race, then one startup's progress can impose even negative spillovers on the other firm. Of course, R&D spillovers can be large and positive in settings different from ours (e.g., Bloom et al., 2013, Matray, 2020). Even though common VC owners hold back R&D at lagging startups, common ownership can improve welfare by avoiding duplication of costs. That feature, which dates back to the patent-race theory of Loury (1979), is not present in López and Vives (2019). Related to López and Vives (2019), Antón et al. (2018) show theoretically that common ownership increases R&D when technological spillovers are large relative to product market spillovers, and they provide supporting evidence using panel data on public firms.

We also contribute by documenting a "horse race" investment strategy used by some VCs. In this strategy, a VC invests in closely competing startups, waits for one to gain an edge, and then reduces funding to the lagging startup while redirecting its innovation. The Appendix provides a case study about one prominent VC firm following this strategy. We argue this strategy has social benefits, but the strategy clearly hurts the lagging startup, at least ex post. Related to this strategy, Ewens et al. (2018) show that VCs increasingly follow a "spray and pray" approach, investing small amounts in many startups with the expectation that many will

be abandoned. Distinct from that strategy, VCs in our paper set up horse races between closely competing startups, and VCs help lagging startups to pivot. Stories of pivots are common among practitioners, but we are not aware of any research on VC-enabled pivots.

1. Empirical approach

1.1. Data and institutional details

Drug development in the U.S. is strictly regulated by the FDA and follows several stages: discovery and Phase I, II, and III clinical trials. The discovery stage involves pre-clinical research for drug candidates in the laboratory and animal testing for basic safety. Phase I tests the safety and determines dosage of a drug candidate on a small group of humans. Phase II tests efficacy and side effects at a larger scale. Phase III, which involves thousands of participants, tests whether there is a treatment benefit. Favorable results are required to move from one phase to the next. All three phases feature high amounts of scientific uncertainty.

We construct a sample of pharmaceutical projects initiated by U.S.-based companies that are funded by VCs. Detailed information on drug development and clinical trials comes from the Cortellis Life Sciences Healthcare Database. Cortellis obtains its data from public records, such as clinical trial registries, FDA submissions, patent filings, company press releases, and financial filings. We provide additional information on the Cortellis database in the Online Appendix. Cortellis records drug development history at the project level. A project is a sequence of trials for testing the safety and efficacy of a drug targeting a specific indication. An indication is a specific disease or medical condition. From Cortellis we collect the dates when projects progress through the various clinical phases; each project's suspension date (in the event of ultimate failure) or FDA drug approval date (in the event of success); detailed information on the company running the project, including the company's name, funding status, organization type, headquarters location, and major shareholders; information on alliance formation and drug repurposing; and detailed information on the portfolio of patents associated with the drug. We also scrape each patent's application date, grant date, and patent citation information from

Google Patent, using the patent number from Cortellis.

We obtain VC investment records from SDC Platinum VentureXpert. We use a fuzzy matching algorithm to link funding records in SDC to company names and locations in Cortellis. For each matched pair, we manually check the matching accuracy. Our sample includes startups funded by corporate VCs (CVCs), but they account for only 2% of our panel dataset.

We apply several filters to obtain our final sample. First, we require both the startup and its VC investors to be in the U.S. We apply this filter because of data availability. Fortunately, the U.S. accounts for a very large share of global pharmaceutical R&D.[3] Second, we drop projects initiated before the first quarter of 2005, because Cortellis's coverage of development histories is less reliable before then. Third, we drop projects initiated after the first quarter of 2016 if no progress or suspension is observed. We do so because insufficient time has passed for these projects to reach any outcome. Next, given our focus on VC, we only include quarters when a given project is held within the portfolio of a VC.[4]

We focus on projects in Phase I trials. A project enters our sample in the first quarter of its Phase I trials, and it exits our sample either when it progresses to Phase II, reaches the end of 2018, or is classified as abandoned. We classify a project as abandoned once it is suspended, discontinued, withdrawn, or coded as "no development" for four years. We focus on projects in Phase I for several reasons. Our startups typically raise their first VC money when their projects are in either the discovery or Phase I trials. We exclude the discovery stage, because it requires significantly less capital, tends to last a short period of time, and is often not reported to the public or Cortellis. Also, VCs have greater control over the startup during projects' early stages. By the time a project reaches Phase II or III, the startup has often gone public or been acquired, removing the VCs' influence. Since we want to study how VCs influence startups, we focus on the stage when that influence is most relevant. Finally, Phase I is inherently important, sometimes being called the "valley of death" in drug development (Seyhan, 2019). In the full Cortellis

[3]In 2016, the U.S. accounted for 58% of pharmaceutical R&D among 11 leading countries (ABPI, 2019).

[4]This period lasts from the first quarter when its drug developer receives its initial VC funding until either the company exits successfully (e.g., IPO or being acquired) or is written off. Following González-Uribe (2019), absent any exit, we say a startup is written off five years after its last VC financing round. If that date exceeds 2018Q4, we keep the project in our sample but record its final outcome as missing. Our main results are robust to using other cutoffs (e.g., 3, 7, and 10 years) to define the write-off date.

database, 55% of project failures occur during or prior to Phase I. Of the Phase I projects in our sample, only 21% progress to Phase II by the end of 2018.[5] Progressing from Phase I to Phase II increases a drug's probability of FDA approval from 10% to 31% (Thomas et al., 2016). While Phase I is traditionally viewed as testing safety and dosage, in VC-backed companies it often has more ambitious goals, such as identifying target patient populations, drug response rates, and biomarker evidence (Ivy et al., 2010; Roberts et al., 2004). The median duration of Phase I is 1.6 years, and for some projects it extends well beyond 3 years (Wong et al., 2019).

Our analysis requires identifying drug projects that are close competitors. We do so by partitioning the overall pharmaceutical industry into 78 highly detailed sub-markets. We map Cortellis indications to the second chapter level of International Classification of Diseases (9^{th} Revision, "ICD"), and we refer to each chapter as an ICD category. Examples of ICD categories include anemia, urinary system diseases, and Kaposi's sarcoma. Some ICD categories combine indications with unspecified sites, unknown causes, or uncertain effects. The diseases in these "miscellaneous" categories can be quite different from each other, so we drop these categories and the projects in them. After we apply this filter, indications in the same ICD category have a high pathological correlation, so drug projects in the same ICD category are plausibly close substitutes and hence close competitors.

Finally, we drop a project if none of its VC investors are disclosed. These filters together produce a final sample of 57,316 VC-project-quarter observations from 1,045 projects across 78 drug categories. The projects are run by 481 startup companies funded by 764 VC firms. The sample covers 2005 to 2018.

1.2. Predictions and identification strategy

Our first prediction is that common ownership increases the likelihood that a lagging competitor project is held back. We provide additional background on the theory of Grossman and Shapiro (1987), which motivates our prediction. Grossman and Shapiro (1987) model firms competing in

[5]For comparison, the FDA website reports that 70% of Phase I projects progress to Phase II, and Wong et al. (2019) report a 66% probability of transitioning from Phase I to II. The transition rate is lower in our sample because (1) many projects are still in progress at the end of our sample period, and (2) VC-backed drug ventures tend to be riskier and more innovative. For example, 32% of our sample projects are related to oncology, a notoriously risky area of drug development.

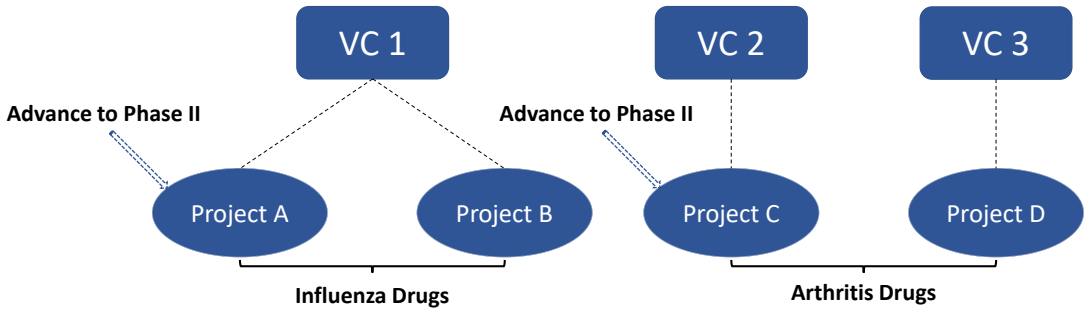
a winner-take-all patent race. They extend Loury (1979) by adding an intermediate milestone that allows one firm to gain an observable lead in the race. The theory maps well to drug development, as patent protection creates a winner-take-all dynamic, and progressing to Phase II trials is an observable intermediate milestone. Grossman and Shapiro (1987) analyze whether the lagging firm continues its R&D project after seeing its competitor reach the milestone. The theory implies that if the two firms share a common owner with strong control rights, the lagging project is more likely to be shut down, because the common owner wishes to avoid duplicating R&D costs.[6] Absent a common owner, the lagging project is less likely to shut down, because the lagging firm sees some chance of catching up in the patent race, and neither firm internalizes the other firm's R&D costs. Competition, in other words, results in over-investment in R&D, as in Loury (1979). By coordinating firms' R&D, a common owner reduces excess R&D and thereby helps solve a market failure associated with patent races.

Figure 1 illustrates the ideal experiment for testing this prediction. Suppose there are two pairs of drug projects, A-B and C-D. The projects within each pair are in the same ICD category, so each pair contains close competitors. One pair (A-B) is randomly assigned to share a common VC investor, and the other (C-D) does not. Within each pair, we shock one project with an observable intermediate success, defined as progressing from Phase I to Phase II. We label the shocked projects, A and C, the "pioneers." Our tests compare the outcomes of the non-shocked projects, B and D, which we label the "lagging" projects. Our tests do *not* compare project A to B, C to D, or A to C, so we do not require those projects to have similar quality. Instead, our main prediction is that project B is less likely than D to progress to Phase II. Having a common VC makes it more likely that the lagging project gets held back.[7]

[6]Specifically, in their application to intermediate patents, Grossman and Shapiro (1987) compare the ex post value if only the leading firm is allowed to continue (V_m in their notation) to the total ex post value if both firms continue ($V_{10} + V_{01}$ in their notation). A common owner with strong control rights can choose between those two options. Grossman and Shapiro (1987) show that $V_m > V_{10} + V_{01}$, implying a common owner would shut down the lagging project, as long as discount rates and elasticities of marginal costs are not too high. Fulghieri and Sevilir (2009) generate a similar prediction, which is that VCs have an incentive to invest in technologically similar companies, because it lets the VC shift resources from one portfolio company to another in case the first company fails.

[7]Similar to us, Krieger (2017) studies how drug projects react to the failure of a competing project. We add an extra "diff" to the analysis by comparing the reaction between pairs with and without a common VC.

Figure 1: Ideal Experiment



We can express this experiment in regression form as follows:

$$Progress_{it} = \gamma\, Shocked_{it} + \beta\, Shocked_{it} \times SharedVC_{it} + \Gamma' Controls_{it} + FEs + \eta_{it}. \qquad (1)$$

The sample includes all projects i that are in Phase I at the beginning of quarter t. $Progress_{it}$ is an indicator for whether project i progresses to Phase II in quarter t. $Shocked_{it}$ is an indicator for whether drug project i has experienced a shock, meaning a different project in the same ICD category (but different startup) has progressed to Phase II between project i's initiation quarter and quarter $t-1$. The indicator $SharedVC_{it}$ equals one if project i and the project causing the shock share a common VC in quarter t. The coefficient of interest is β, which measures how lagging projects' outcomes depend on whether there is a shared VC. We predict $\beta < 0$, meaning a project is less likely to progress to Phase II, after receiving a shock, if there is common ownership.

At first glance, Eq. (1) looks like a "triple-diff" regression. It examines projects over time (the first diff), depending on whether the project experiences a shock (the second diff), and also depending on whether the projects share a VC (the third diff). Unlike in a triple-diff, though, $SharedVC_{it}$ cannot be included as an extra regressor, because it is only defined when $Shocked_{it} = 1$. More important, $SharedVC_{it}$ is not randomly assigned, so OLS estimation of Eq. (1) does not have a causal interpretation.

We therefore take an IV approach, which requires finding quasi-random variation in $SharedVC$, the indicator of common VC ownership. Our instrument exploits the local nature of VC invest-

ing. VCs prefer investing in nearby companies to reduce the costs of screening and monitoring (Lerner, 1995). For example, Gompers et al. (2019) show that VCs rely heavily on their networks to source deals, and networks are often local. We use the geographic proximity of two startups to create an instrument for whether they share a common VC investor. Specifically, we create an instrument for the endogenous interaction term $Shocked_{it} \times SharedVC_{it}$ as follows. For project i and quarter t, we collect the set S_{it} of projects j that are in the same ICD category as i, and that progressed from Phase I to Phase II between the birth of project i and quarter $t-1$. These projects j are the pioneers that shock project i. Let d_{ij} denote the distance, in miles, between the Metropolitan Statistical Areas (MSAs) of startups' headquarters for those startups running projects i and j. We scale the distance by a constant of 2,600 miles, which is roughly the air travel distance from Boston to San Francisco. This scaling helps produce a strong first-stage regression. Given the scaled distance d'_{ij}, we compute $f\left(d'_{ij}\right) = \exp(-d'_{ij})$ as a proxy for the probability that projects i and j share a common VC. Our IV for the variable $Shocked_{it} \times SharedVC_{it}$ is then

$$Shocked_{it} \times Proximity_{it} = \sum_{j \in S_{it}} \exp(-d'_{ij}). \tag{2}$$

We set the instrument to zero if the set S_{it} is empty, meaning $Shocked_{it} = 0$. We compute a sum in Eq. (2) because common ownership is more likely if there are multiple nearby pioneer projects. Summation is not driving our results, however; our results are actually stronger if we control for the number of pioneer projects in set S_{it}. We show later that geographic proximity strongly explains common ownership, and the first stage of our IV regressions is quite strong. The relevance condition holds for this instrument, in other words.

Our main identifying assumption is that the instrument affects the outcome variable $Progress_{it}$ only through its effect on $Shocked \times SharedVC_{it}$. Why does this exclusion restriction plausibly hold? Our IV test boils down to checking whether, after being shocked, a lagging project is more likely to fail if it is geographically closer to the pioneer project. More simply, we test whether geographic proximity predicts opposite outcomes for the lagging and pioneer projects. Reverse causation is not an issue here, because the startups' locations are determined well be-

fore the projects' outcomes. The bigger concern is omitted variable bias, meaning there exists some omitted variable W that is correlated with both geographic proximity and the projects' achieving opposite outcomes. No such W is apparent to us. To reduce portfolio risk, a VC may prefer investing in projects with negatively correlated outcomes, and unobserved negative correlation could explain projects' opposite outcomes. This is exactly why we need an instrument for $SharedVC$. There is no reason why geographic proximity should coincide with a negative correlation. If anything, nearby companies are likely to have unobserved similarities, for example, because the scientists and managers have similar backgrounds. Unobserved similarity, however, predicts *similar* project outcomes. We predict *opposite* project outcomes, so this force biases us against our findings. Also, we find no support for the premise that nearby startups are more similar: In the Online Appendix, we show that when the lagging and pioneering projects are closer to each other, they are actually less likely to share patent citations, although the difference is not statistically significant. We further show that our conclusions are robust to controlling for this measure of technological similarity. Section 3 addresses other possible identification challenges.

In the Online Appendix, we show that projects B and D (i.e., the lagging projects with and without common ownership) do not differ significantly on observable proxies for quality. Proxies we consider include project age, startup age, whether the startup has advanced previous projects to Phase II or Phase III, and whether the project is located in Boston or San Francisco. The lack of correlation holds whether we compare lagging projects based on the endogenous variable $SharedVC$ or based on the instrument $Proximity$. The exclusion restriction cannot be tested, but it's comforting that neither the endogenous variable nor the instrument correlates significantly with proxies for project quality.

2. Empirical results

2.1. Summary statistics and frequency of common ownership

Figure 2 shows the geographic distribution of drug projects in our sample. As expected, there are many projects in California, Massachusetts, and New York. Approximately 46% of projects, however, are located outside these states. Many projects are in states with a strong medical research background, such as Illinois, Maryland, North Carolina, Pennsylvania, and Washington.

Table 1 shows summary statistics for our main variables in the full sample as well as the "Never Treated" and "Ever Treated" subsamples. The "Never Treated" group contains projects or companies for which $Shocked \times SharedVC$ equals zero across their entire lifespan. Remaining projects or companies are in "Ever Treated."

Panel A shows project-level variables. The full-sample mean of *Shocked* is 0.504, meaning in roughly half of our sample a project has seen a close competitor progress to Phase II. *Progress* is roughly three times higher in the "Never Treated" group, which foreshadows our result that sharing a VC with a competing, successful project significantly reduces the likelihood of progressing to Phase II.

Panel B contains startup-level variables. In a typical quarter, the average startup has 1.66 projects covering 1.42 ICD categories. A typical startup will run several projects over its lifespan, so the total number of projects per startup is much higher than 1.66. The typical startup has 4.7 VC firms invested in it at a given point in time, although there is wide variation in this number.

Next, we show that common ownership by VCs is quite common. Table 2 shows that a sizeable 39% of our startups have at least one close competitor that shares at least one VC in common, in a typical quarter. At first glance, this common ownership rate seems high given that our sample contains 764 distinct VC firms. The rate is high for a few reasons. First, the average startup has 4.7 VCs, so there is a decent chance that at least one VC is shared with a competitor's (multiple) VCs. Also, Table 2 shows that VCs specialize in specific ICD categories. While a VC could potentially invest across all 78 drug categories, the typical VC holds stakes in just 2.5 categories in the typical quarter, making it more likely that the VC holds multiple

startups in the category. Put differently, while there are 764 VCs in our full sample, the typical ICD category has only 16.4 VCs invested in it at once, increasing the chances of VC overlap. This high degree of specialization makes sense given the high level of scientific expertise needed to evaluate drug startups. Finally, some ICD categories contain many startups, making it easier to find at least one other competing startup that shares a VC. The rate of common ownership appears lower, 8.6%, if we measure it instead as the probability of a shared VC between two startups chosen at random from within an ICD category.[8]

2.2. Evidence from project outcomes

Our main results begin in Table 3, which contains results from the panel regression in Eq. (1). Except in columns 2, 3, and 5, the dependent variable is $Progress_{it}$, the indicator for whether project i progresses to Phase II in quarter t. Within each project, $Progress$ equals zero until the project's last observation, which equals either one (if the project progresses) or zero (if the project is suspended or reaches the end of the sample). We include fixed effects (FEs) for the startup, ICD category, quarter to control for unobserved heterogeneity.[9] We include three controls: the project's age ($\ln(Age)$), number of Phase I projects being developed by the startup ($NProjects$), and number of VC firms owning a stake in the startup company ($NVCs$). We find that $\ln(Age)$ is a particularly important control, as it takes time to obtain Phase I results. Standard errors are computed by two-way clustering at the ICD category and startup levels.

The first column contains OLS estimates. This regression does not have a causal interpretation, but it provides a useful description of the data. Since the dependent and independent variables of interest are indicators, the slope coefficients can be interpreted as average changes in probability. The OLS results point in the predicted direction: Sharing a common VC is associated with a 0.019 lower probability that, if a similar project makes progress, the project progresses to Phase II during the quarter. The unconditional average of $Progress$ is 0.018,

[8]These two different rates of common ownership imply that, from a given startup's point of view, it is quite unlikely that a specific competitor shares a VC, but it is easier to find at least one competitor that shares a VC.

[9]The ICD category FEs are useful because it may be more difficult to reach Phase II for certain diseases. The quarter FEs controls for aggregate changes in the FDA process, conditions in the macroeconomy or pharmaceutical industry, etc. The startup company FEs are useful, because the typical startup runs several projects over its lifespan, and some startups may be systematically better at getting projects through Phase I.

so the estimated coefficient is economically large. It is also statistically significant, with a t-statistic of -2.95. The insignificant slope on *Shocked* indicates that, absent a shared VC, seeing a related project make progress makes it neither more nor less likely that the lagging project progresses. Combining both coefficients implies that a lagging project becomes less likely to progress after being shocked, relative to non-shocked projects, only if it shares a common VC with the pioneering project.

Our first IV test is in columns 2–4. Since the focal regressor, $Shocked \times SharedVC$, is a binary variable, two stage least squares (2SLS) is generally imprecise and results in coefficients with large magnitudes. Wooldridge (2010) and Angrist and Pischke (2008) instead recommend a Probit-2SLS procedure.[10] In the first step (column 2), we estimate a probit regression of the endogenous regressor ($Shocked \times SharedVC$) on the instrument ($Shocked \times Proximity$) and exogenous controls.[11] The probit model produces predicted probabilities $\widehat{Shocked \times SharedVC}$, which we use as the instrument in the first stage of the 2SLS procedure (column 3). Probit models do not easily accommodate a large number of FEs, so we only introduce the FEs in the first stage of the 2SLS procedure.

As expected, in column 2 we see a strong, positive relation between $Shocked \times Proximity$ and $Shocked \times SharedVC$, consistent with VCs preferring to invest in nearby companies. The t-statistic is 7.2 in column 2 and 7.6 in column 3's first-stage regression. Weak instruments are clearly not a problem. The IV estimate in column 4 echoes the OLS results. Consistent with our prediction, after seeing a similar "pioneer" project progress to Phase II, a project is less likely to progress to Phase II if it shares a VC with the pioneer project. Compared to OLS, the IV results have higher statistical and economic significance.

Jiang (2017) shows that it is common for IV estimates to be much larger than their OLS counterparts. It makes sense that our IV slopes exceed the OLS slopes. Our IV picks up commonly owned startups that are geographically near each other. Our IV-compliers are therefore VCs whose investment decisions are more influenced by the monitoring costs associated with

[10] See, for example, page 191 in Angrist and Pischke (2008). This approach has been used recently by Dinc and Erel (2013), Saretto and Tookes (2013), and Ewens and Marx (2018).

[11] We omit *Shocked* from column 2 because its slope is not identified in the probit model. If $Shocked = 0$, then $Shocked \times SharedVC$ is zero with probability one, so the probit model wants to assign *Shocked* a slope of positive or negative infinity.

geographic proximity. This subgroup of VCs is likely to exhibit larger local average treatments effects (LATEs), for at least three reasons. A VC who intends to run a horse race between two startups knows it will have to monitor them closely, so the VC will want them closer together. Second, our IV-compliers arguably care more about monitoring in general, making them more aware of spillovers within the portfolio, which is necessary for our prediction to hold. Finally, the IV-compliers are arguably more cost-conscious, making them more eager to avoid duplicating R&D costs, which in turn makes them more willing to suppress a lagging project.[12]

Columns 5 and 6 show that our result is robust to using a bivariate probit model instead of 2SLS. Bivariate probit models are suited to IV settings like ours, in which the dependent and endogenous independent variables are both binary.[13] The bivariate probit estimates again support our prediction. Relative to OLS, economic significance is higher, but statistical significance is lower. An advantage of the bivariate probit model over 2SLS is that it recognizes that our dependent and independent variables are binary. A disadvantage of the bivariate probit model is that it imposes a strong distributional assumption on the data, namely, that error terms are normally distributed. A bigger disadvantage is that bivariate probit does not easily accommodate the large number of FEs in our 2SLS specification, so we omit all FEs in columns 5 and 6. Despite these shortcomings, we continue to report bivariate probit results alongside 2SLS results when possible.

One potential concern is that the pioneering shocks' strength correlates with the frequency of common ownership in a disease area. We address this concern by adding ICD-by-quarter FEs in column 7, so that our test now asks how lagging projects react to exactly the same shock. These FEs isolate, for example, variation across different Phase I influenza drug projects in a given quarter. If an influenza project progresses to Phase II in quarter $t-1$, then all remaining Phase I influenza projects in quarter t would have $Shocked = 1$. (Note the project that progresses to Phase II during quarter $t-1$ is no longer in our sample during quarter t.) The variable $Shocked \times SharedVC$ would therefore only pick up variation in $SharedVC$ across remaining

[12]Supporting evidence is in the Online Appendix. Consistent with the IV-compliers being more attentive to spillovers, IV-compliers hold less-diversified portfolios. IV-compliers are also slightly smaller, which arguably correlates with being more cost-conscious.

[13]Bivariate probit models are discussed by Angrist and Pischke (2008), pages 197–205, and Wooldridge (2010), pages 594–599.

Phase I influenza projects in quarter t, and the regression would test how those remaining projects respond differently to the same shock depending on whether they share a VC with the shocked project. Column 7 shows that including ICD-by-quarter FEs has a negligible effect on the OLS magnitude and decreases its statistical significance only slightly. We show in the Online Appendix that our IV results also continue to hold if we include ICD-by-quarter FEs.

Some projects fail to reach Phase II not because their owners actively suppress them, but simply because they have bad Phase I outcomes. If common owners play an active role in holding back lagging projects, then our results should hold even in the subset of projects with good Phase I outcomes. To test this idea, we re-estimate the previous models after excluding projects that, according to Cortellis, fail to reach the primary Phase I end point or experience an adverse Phase I event, such as a dangerous side effect or death. Results are in Table 4. Economic significance is even stronger compared to Table 3. The negative coefficients on $Shocked \times SharedVC$ remain statistically significant, even though the sample has shrunk by more than half. These results support the idea that common owners play an active role in holding back lagging projects.

Figure 4 Panel A describes these effects' dynamics. We estimate an OLS regression similar to Eq. (1), except we include multiple leads and lags of $Shocked$ and $Shocked \times SharedVC$. The figure plots the OLS coefficients on $Shocked \times SharedVC$ at the various leads and lags. The horizontal axis denotes the number of quarters after the "event" quarter, i.e., the quarter when $Shocked \times SharedVC$ first equals 1 for each project. As in our main analysis, we do not include shocks occurring before projects are initiated. The category $\tau = 12$, for example, contains projects that existed and received a shock 12 quarters previously. We see no significant coefficients in the six quarters leading up to the event, indicating no anticipation effects or pre-trends.[14] After the event, all coefficients have negative point estimates, and several are statistically significant. There is no indication of a delayed response or reversion back to zero. Of course, this figure only shows the average pattern for a treated project, and there is heterogeneity around this pattern. For example, projects will sometimes wait for Phase I results to arrive before deciding whether to progress to Phase II. Overall, the figure implies that when a commonly

[14]The absence of pre-trends is reassuring. Recall, though, that our main IV tests are not diff-in-diff regressions, so they do not require an assumption about parallel trends. As before, these OLS regressions do not have a causal interpretation.

owned competitor make progress, on average there is a prompt, persistent drop in the project's odds of progressing. To address concerns that our panel structure spuriously drives our result, we show in Section 3 that our result continues to hold in a purely cross-sectional test.

To summarize, these project-level results support the prediction that common ownership leads VCs to hold back lagging projects. Specifically, we find that common ownership makes it less likely that a lagging project progresses to Phase II after a competing project does so. The result holds in OLS, IV, and bivariate probit specifications. This result goes through whether we look at how projects respond to different shocks or the same shock. The result holds even if we study only those projects that have the potential to progress to Phase II. The effects begin soon after the shock occurs, and they are persistent.

These results support the patent race theory of Grossman and Shapiro (1987). According to their theory, common ownership makes it more likely that a lagging project is held back, because the common owner internalizes both firms' R&D costs and seeks to avoid duplicating costs. Without a common owner coordinating their R&D investments, the startups are predicted to over-invest in R&D in hopes of winning the patent race. Our results suggest that common ownership reduces the duplication of R&D, which can improve innovation efficiency.

2.3. A VC financing mechanism

How exactly can a VC hold back a lagging drug project in its portfolio? We investigate a mechanism related to VC financing. A common VC can avoid duplicating R&D in its portfolio by refusing follow-on funding for lagging projects. Startups typically have negative cash flows, so without follow-on funding they are unable to continue their projects.

For this mechanism to work, a startup that loses funding from one VC cannot easily raise funding from another source, including other VC firms. VCs' private information makes this condition likely to hold, at least in some situations. The logic comes from the relationship-lending literature (e.g., Rajan, 1992; Petersen and Rajan, 1994). A startup will typically raise multiple rounds of financing from one or more VC firms. After investing in and working with a startup, a VC gains private information about the startup. A VC's choice to stop funding

a startup can send a negative signal, leading other VCs to either avoid funding the startup or offer unfavorable terms.[15] This private information channel is not necessary, though. VCs often have contractual protections such as veto power over the startup's future financing rounds. A VC with strong enough veto power can effectively prevent the startup from fundraising.[16]

To test this mechanism, we consider a variation on our ideal experiment. As before, we consider two pairs of similar drug projects, and one project (the "pioneer") within each pair progresses to Phase II. In the ideal experiment, we would randomly assign one pair of projects to share a VC and the other pair to not share a VC. Whereas we previously studied whether the lagging projects progress to Phase II, we now study whether the startup companies running those lagging projects receive VC funding. We predict that, if there is a shared VC, then that VC is less likely to fund the lagging project's company. We can express this experiment via the following regression:

$$ExtendFunds_{ijt} = \gamma\, Shocked_{ijt} + \beta\, Shocked_{ijt} \times SharedVC_{ijt} + \Gamma' Controls_{ijt} + FEs + \eta_{ij}, \quad (3)$$

where $ExtendFunds_{ijt}$ equals one if VC j extends funding to startup i in quarter t. To create a company-level rather than project-level variable, we redefine $Shocked_{ijt}$ to equal one if at least one project in startup i has seen another project (a "pioneer") in the same ICD category (but different startup) progress to Phase II between the quarter when VC j invests in startup i and quarter $t-1$. $SharedVC_{ijt}$ equals one if VC j holds a stake in both startup i and at least one of the startups owning the pioneering projects that produced $Shocked_{ijt} = 1$. To create a company-level version of our instrument, $Shocked_{ijt} \times Proximity_{ijt}$, we modify Eq. (2) so that we sum our proximity measure over all the pioneering projects that shock any of company i's projects. We also include controls for the number of startups and ICD categories in the

[15] This mechanism might not work if outsiders perfectly infer that funding is withheld due to common ownership rather than a negative private signal. But as long as outsiders cannot perfectly distinguish between a common ownership motive and a negative private signal, the mechanism should continue to work. Ewens et al. (2016) show empirically that "inside" VC financing rounds predict worse startup outcomes, which is inconsistent with VCs using their private information to hold up portfolio companies.

[16] In theory, a lagging firm can also turn to banks or government for funding. While those sources are broadly important for funding innovation (e.g., Hombert and Matray, 2017, Hombert and Matray, 2018), they are much less relevant for pharma startups (Wessner, 2009 and Harrington, 2012).

VC's portfolio, the size of and time elapsed since the startup's previous financing round, the number of VC firms invested in the startup, and the number of projects owned by the startup. The first columns include VC firm FEs (because some VCs invest more frequently than others), startup FEs (because some startups raise more VC rounds than others for unrelated reasons), and year-quarter FEs (to soak up aggregate unobservables). We predict $\beta < 0$, meaning a VC is less likely to extend funding to a startup, after the startup sees a competitor make progress, if the VC is invested in both companies.

Support for this prediction is in Table 5. The OLS slope coefficient on $Shocked \times SharedVC$ is indeed negative and statistically significant at the 5% level. The coefficient's magnitude, -0.030, is quite large relative to the 0.065 unconditional mean of $ExtendFunds$. Column 3 shows a very strong first-stage regression, and column 4's 2SLS slope on $Shocked \times SharedVC$ is negative, large in magnitude, and statistically significant at the 5% level. Results from the bivariate probit model also strongly support the prediction, with a t-statistic of -9.2. Overall, these results support the idea that VCs withhold funding from a startup if a commonly owned close competitor has made progress.

The large negative slope on $Duration$ in Table 5 deserves comment. $Duration_{ijt}$ is the log of one plus the number of years since VC j last funded startup i. VCs—especially early-stage ones—will often fund a company for a few rounds, then stop participating in rounds once the round sizes get too large. The negative slope on $Duration$ likely reflects that, once a VC has stopped funding a company, the VC is unlikely to start funding it again.

One potential concern with the previous test is that a startup's demand for funding, an omitted variable, may differ depending on whether there is a shared VC or on geographic considerations. We address this concern by adding startup-by-quarter FEs in column 7 of Table 5. These FEs isolate variation across VCs within the same startup-quarter, and hence they effectively control for the startup's funding demand. For example, suppose startup i has two existing VC investors, A and B, and startup i sees a competing startup i' make progress. VC A is invested in both startups, but VC B is only invested in startup i. The regression tests whether VC A is less likely than B to extend funding to startup i in the given quarter. As we predict,

column 7 shows that the coefficient on $Shocked \times SharedVC$ remains negative, indicating that VC A (in our example) is indeed less likely than B to extend funding. Economic significance is similar to column 1, and statistical significance is slightly higher. The Online Appendix shows that our IV results also continue to hold after including startup-by-quarter FEs.

Panel B of Figure 4 illustrates these effects' dynamics. As in the project-level results, we see no significant pre-trends, the effects begin to appear quite soon after the shock, and the effects are highly persistent.

One potential challenge to the financing mechanism is that other VCs could step in and fill the hole left by a VC who abandons the startup. Table 6 shows this not to be the case. When there is a shock to a startup with common ownership, the startup is significantly less likely to raise money from *any* VC (column 1), and the total dollar amount raised from VCs is significantly lower (column 2). These results indicate that new VCs, or existing VCs without common ownership, do not fill the hole left by the existing VC with common ownership. In fact, column 3 shows that new VCs are significantly less likely to invest in a startup when there is a shock to a startup with common ownership. As explained earlier, this behavior can result from VCs' contractual protections or negative signals from abandonment by an earlier investor. Column 4 tests whether the existing non-common VCs offset the effects of the common VC. The regressor of interest is $Shocked \times SharedVC \times NonCommonVC$, an indicator for whether startup i is shocked by a pioneer, at least one of startup i's VCs is shared with the pioneer, and VC j is not the shared VC. We find no significant relation between this regressor and $ExtendFunds$, an indicator for whether VC j extends funds to startup i in quarter t. This result implies that existing non-common VCs do not significantly offset the common VC's reduction in investment.

We find interesting heterogeneity in the VC-funding results. Table 7 shows results from estimating the baseline VC-funding regression by OLS in various subsamples. (We estimate by OLS because the subsamples are too small to deliver significant first-stage IV regressions.) The first two columns indicate a much stronger VC-funding mechanism for lead VCs, defined as the VC with the highest total amount invested across the given startup's VCs. For non-lead VCs, the coefficient on $Shocked \times SharedVC$ is close to zero. This result makes sense,

because a lead VC arguably has stronger control rights over the startup. Also, lead VCs are expected to provide greater follow-on funding, so they would suffer more from duplication of R&D. The last two columns indicate a stronger VC-funding mechanism for less-diversified VCs. We create a simple diversification proxy equal to the number of unique startups, including non-pharmaceutical startups, the VC invests in during the previous five years.[17] The subsample with below-median diversification has a coefficient on $Shocked \times SharedVC$ that is much larger in magnitude (-0.049 versus -0.009). This result supports the finding in Gilje et al. (2019) that, in order for common ownership to have any effects, the common owners must be attentive to spillovers within the portfolio. Since less-diversified VCs can be more attentive to each portfolio company, it makes sense that common ownership has stronger effects among such VCs.

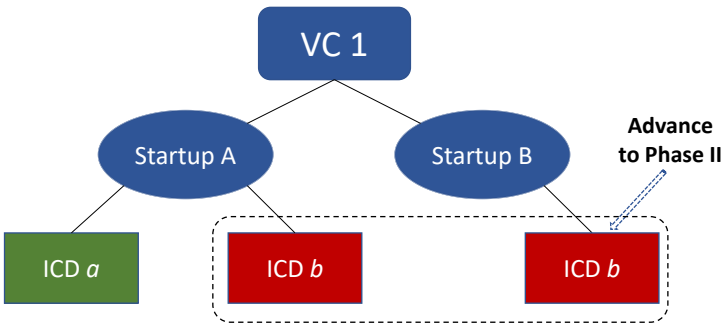
To summarize, the previous three tables are consistent with a VC financing mechanism by which common ownership affects project outcomes. An existing VC with common ownership is less likely to fund a startup that is lagging behind a (commonly owned) competitor. Making matters worse, new VCs are deterred when the existing VC abandons the firm, and other existing VCs do not offset these effects. By starving the lagging startup of capital, the common VC owner can hold back the startup's projects. The financing channel is especially strong for lead VCs and less-diversified VCs.

Besides the financing mechanism, there may be other ways that common ownership results in weaker startups being held back. A common VC owner can use their board seats, voting rights, option to replace the CEO, or other control rights. The VC can also encourage employees to leave the lagging startup and join the leading one. A VC can simply exert less effort helping a lagging startup. These mechanisms are not mutually exclusive. Our earlier results on project outcomes reflect these mechanisms' combined effects.

2.4. Redirecting innovation

Next, we explore how common owners redirect innovation at lagging firms. The figure below illustrates our testable prediction.

[17] Computing a more sophisticated diversification proxy is challenging, because portfolio holdings are not marked to market, we do not know when VCs write off failed startups, and our data do not allocate a VC syndicate's investment across the syndicate members.



Suppose VC 1 is a common investor in Startups A and B, and Startup B has a project in ICD category b that progresses to Phase II. That project corresponds to the pioneering project that "shocks" the lagging startup, denoted A. Our previous tests focus on Startup A's projects in the overlapping ICD category, b. We now examine Startup A's projects in non-overlapping categories, such as ICD a. To the extent that Startup A continues innovating, the common VC would like to redirect A's innovation efforts so they do not overlap with Startup B's pioneering project in ICD b. We therefore predict that Startup A's innovation activity increases in ICD a relative to ICD b. The common VC may redirect innovation because it wishes to avoid duplicating R&D costs in ICD b, or because it wants to reduce competition faced by Startup B. Even if the common VC stops funding the lagging startup, the VC's existing control rights may give it enough power to redirect innovation, at least in some cases.

The Cortellis data allow us to measure the redirection of innovation in four ways. Our first measure, $Initiation_{ijt}$, is an indicator for whether startup i initiates a new project in ICD category j in quarter t. $RepurposeInto_{ijt}$ is an indicator for whether the startup repurposes an existing project into ICD category j from a different ICD category j' during t.[18] In the example above, we predict that Startup A repurposes its project from ICD b into ICD a. Conversely, $RepurposeAway_{ijt}$ is an indicator for whether startup i repurposes an existing project from ICD category j into a different ICD category j' during t. In the example above, we predict that Startup A is less likely to repurpose away from ICD a. Finally, $Alliance_{ijt}$ is an indicator for whether the startup i forms an alliance with another pharma company—typically a large, public company—in ICD j and quarter t. Encouraging an alliance in ICD a is one way that VC 1 can

[18]Drug companies sometime repurpose a chemical compound from treating one disease to another. For example, Pfizer created Viagra by repurposing a PDE5 inhibitor from coronary artery disease to erectile disfunction.

redirect innovation toward ICD a without funding the innovation itself.

We regress these outcome variables on $Treated \times NonSharedICD_{ijt}$, an indicator for whether (1) startup i has a project in a different ICD category j' that experienced a shock, meaning a different project (at a different startup) in ICD j' progressed to Phase II; and (ii) startup i and the startup causing the shock share a common VC. In the figure above, $Treated \times NonSharedICD$ would equal 1 for Startup A and ICD a, and it would equal zero for Startup A and ICD b. We include controls related to those in previous tables. We also include startup-by-quarter FEs in order to measure innovation activity in ICD a relative to b, within the startup-quarter. Those FEs are useful, because we know from the previous section that Startup A is likely to lose VC funding, so startup A is likely to reduce innovation activities in all ICD categories. The interesting comparison is therefore across ICDs within a startup-quarter.

Results are in Table 8. Column 1 shows that the lagging startup is 3.4 percentage points (pp) more likely to initiate in the non-common ICD. This magnitude is roughly 40% of *Initiation*'s unconditional average (8.3 pp). Column 2 shows that the lagging startup is 2.6 pp more likely to repurpose existing projects into the non-common ICD, and column 3 shows it is 3.6 pp less likely to repurpose projects away from the non-common ICD. In column 4, we see that the lagging project is 0.9 pp more likely to form an alliance in the non-common ICD. That result is significant only at the 10% level, whereas the previous results' t-statistics range from 2.2 to 3.8 in magnitude. All four results are consistent with the lagging, shocked startup redirecting innovation toward ICD categories without common ownership. Using the example from the previous figure, we find that Startup A's innovation efforts increase in ICD a relative to ICD b. This redirection occurs through starting new projects, repurposing existing projects, and creating alliances with large pharma companies.

2.5. Implications for innovation efficiency

We show above that common ownership makes a lagging project and startup more likely to be held back. By reducing the duplication of R&D costs in this way, common ownership can improve innovation efficiency. Specifically, common ownership can reduce the total R&D cost, aggregated

across all startups and projects, per drug that makes it all the way to FDA approval. This subsection provides descriptive evidence that common ownership indeed improves innovation efficiency. Specifically, we show that common ownership rates are positively correlated to a proxy for innovation efficiency.

We perform a simple cross-sectional comparison of drug categories. We measure an ICD category's common ownership rate as the probability that a randomly selected pair of startups active in the ICD share at least one VC in common, averaged over startup pairs and the sample period. Our efficiency proxy seeks to measure the number of approved drugs per dollar of R&D spending, aggregated across all projects. The proxy equals the number of drugs within the ICD category that reach FDA during our sample period, divided by the total dollar amount of funding provided by VCs to all startups (successful and unsuccessful) that are active in the given ICD during the sample period. If a startup has projects in multiple ICD categories, we allocate the startup's VC funding to ICD categories in proportion to the startup's number of projects in each category. Using VC funding as a proxy for R&D spending is reasonable given that most pharma VC dollars go toward drug development as opposed to, say, marketing.

Figure 5 plots innovation efficiency versus the common ownership rate, with each circle representing an ICD category. Common ownership rates vary significantly across ICD categories, from zero to roughly 4.5%. Several ICD categories see no drug approvals and therefore have efficiency measures equal to zero, highlighting the challenges in drug development. Other ICD categories see efficiency measures as high as 0.005, which corresponds to 1/0.005=$200 million of aggregate VC funding per approved drug. The dashed best-fit line shows a positive relation between the efficiency measure and common ownership rate. Column 1 of Table 9 shows the underlying regression and confirms that the positive relation is highly statistically significant, with a t-statistic of 3.6. Economic significance is also quite high. A one standard deviation increase in the common ownership rate is associated with a 0.34 standard deviation increase in the efficiency measure.[19]

We perform a number of additional checks. In column 2 we control for potentially important

[19]The standard deviations of *Approval Efficiency* and *Common Ownership Rate* are 0.00111 and 0.0123, respectively, and 0.34=0.031×0.0123/0.00111.

omitted variables. To capture the idea that drug development is costlier or more prone to failure in certain ICDs, we control for the average duration between project initiation and reaching Phase III and the probability of reaching Phase III. (We compute these measures for Phase III rather than FDA approval, since many ICDs experience no FDA approvals.) One limitation of our efficiency measure is that it excludes funding raised after the VCs exit, e.g., in an IPO or after an acquisition. Although it is not clear why those exclusions would relate to common ownership rates, we control for VCs' average holding period to proxy for how early or late VCs exit their investments. None of these controls enter significantly, and the positive relation between efficiency and common ownership weakens only slightly. Column 3 shows that the positive relation also emerges if we replace the common ownership rate with a dummy for whether the rate is positive. If we exclude ICDs with zero approval efficiency rates, the relation remains positive but is not statistically significant (t =1.3, not tabulated). In the Online Appendix, we show that results are robust to measuring common ownership rates differently or measuring efficiency based on the number of projects reaching Phase III rather than FDA approval.

Even with these checks, this remains a simple, descriptive analysis. Common ownership rates are not randomly assigned to ICDs, so these results do not have a causal interpretation. Nevertheless, the results are consistent with common ownership improving innovation efficiency, arguably because common ownership helps to avoid the duplication of R&D costs.

3. Robustness and identification challenges

For our main predictions to hold, the pioneering and lagging projects should be close competitors. Next, we show that our results are stronger when that condition is more likely to hold. Our first proxy for "close competitor" is whether the pioneering and lagging projects have an overlapping patent citation, indicating technological similarity.[20] We define $Shocked \times SharedCite$ to equal one if the lagging startup's drug shares at least one outgoing citation in common with the progressing drug. We interact that term with $SharedVC$, creating a triple-interaction term,

[20]Cunningham et al. (2020) and others use an alternative proxy based on drugs' mechanism of action. We do not use that proxy because the data on mechanism of action are missing much more often for small, private companies than large, public companies.

and we include it in our project-outcome and VC-funding regressions. Results are in Table 10. To provide a reference point, columns 1 and 3 recap our previous OLS analysis from Tables 3 and 5 while adding the new control variable, $Shocked \times SharedCite$. We still find significantly negative slopes on $Shocked \times SharedVC$. Next, in columns 2 and 4, we add the triple interaction term $Shocked \times SharedVC \times SharedCite$. We predict a negative coefficient on this term, meaning the negative relation between the dependent variable and $Shocked \times SharedVC$ is even more negative if the projects' technologies are more similar. The estimated coefficients are indeed negative. When the dependent variable is project-level success ($Progress$, column 2), the slope on the triple-interaction term has only marginal statistical significance. It is very large in magnitude, however, implying the effects of $Shocked \times SharedVC$ are roughly twice as large ([0.017+0.020]/0.017]) if $SharedCite = 1$. When the dependent variable is company-level VC funding ($ExtendFunds$, column 4), the slope on the triple-interaction term is large in magnitude and statistical significance.

Our second proxy for "close competitor" is whether the projects belong to a narrowly defined ICD category. Some ICD categories have precise, narrow definitions, such as Kaposi's sarcoma. That category contains only 5 active Phase I projects in the typical quarter, and those projects are likely to be close competitors. Other ICD categories have much broader definitions, such as "hereditary and degenerative diseases of the central nervous system" (e.g., Alzheimer's, Huntington's, and Parkinson's diseases). That category contains 96 Phase I projects in the typical quarter, and it is unlikely all those projects are close competitors. We define $NarrowICD$ to be an indicator for whether the ICD contains a small number of projects in the average quarter, and we interact that variable with $SharedVC$ in our OLS regressions. Table 11 shows that the resulting triple-interaction term enters negatively, with t-statistics below -2 and high economic significance. The result holds whether we define $NarrowICD$ by counting Phase I projects, Phase I+II projects, or Phase I+II+III projects.

Together, these results imply that common ownership is especially likely to result in a project being held back if the lagging and pioneering projects are close competitors. This result helps rule out a simpler story, which is that patent races are irrelevant, yet VCs have limited capital.

In that story, success by any portfolio company results in less funding for all other portfolio companies, regardless of how closely the companies compete. As a result, the triple-interaction terms would not enter significantly. Since they do enter significantly, our interpretation about patent races between close competitors appears more plausible. Of course, capital constraints are not necessarily irrelevant, since they can influence a VC's desire to avoid duplicating R&D in a patent race.

Our main project-level tests take the form of panel regressions. For robustness, we show that results are similar when we collapse the panel into a single observation per project. We redefine $Progress_i$ to be an indicator for whether project i ultimately progresses to Phase II. $Progress_i$ equals zero if the project is eventually suspended or remains "in progress" indefinitely. We redefine $Shock_i$ to be an indicator for whether project i receives a shock at any time between the project's birth and the end of its Phase I. The results, shown in the Online Appendix, are quite similar to those in the full panel regressions. Whether we estimate by OLS, 2SLS, or bivariate probit, we find a significantly negative loading of $Progress$ on $Shocked \times SharedVC$, consistent with our prediction. The slope coefficients are considerably larger in magnitude than in the panel regressions, which makes sense given that this test aggregates all of a project's quarters. The OLS estimate, for example, implies a sizable 0.137 lower probability of ultimately progressing to Phase II if there is a shock and a shared VC.

Our results are also robust to estimating a Cox hazard model. The Online Appendix shows that the hazard rate of progressing from Phase I to Phase II, conditional on seeing a close competitor reach Phase II, is significantly lower if the pioneering and lagging projects share a common VC. Since the Cox model does not easily accommodate a large number of FEs, we focus instead on linear probability models throughout the paper.

We also find support for our project-level predictions if we replace the dependent variable $Progress_{it}$ with its opposite, $Suspend_{it}$, an indicator for whether project i is suspended in quarter t. We focus on $Progress$ in our main tests because it is difficult to accurately measure suspensions. We define $Suspend_{it}$ in two steps. First, we check whether Cortellis records a project as being explicitly discontinued, withdrawn, or out-licensed in a given quarter. Second,

there are many projects that are never officially discarded and yet continue to be listed in the drug portfolio without further trial updates. We assume these "zombie projects" are suspended three years after the first quarter when Cortellis designates them as "no development reported." For zombie projects without such designation, we assume they are suspended five years after project initiation. We predict a positive relation between $Suspend$ and $Shocked \times SharedVC$, meaning a project is more likely to be suspended if a close competitor with a shared VC makes progress. The results, shown in the Online Appendix, support this prediction. Whether we estimate by OLS or IV, the slopes on $Shocked \times SharedVC$ are significantly positive, consistent with the negative slopes we find when using dependent variable $Progress$ in Tables 3 and 4. Relative to $Progress$, $Suspend$ produces higher economic significance but somewhat lower statistical significance. Overall, our main conclusions seem robust to using this alternative outcome variable.

We show in the Online Appendix that our main results are largely robust to dropping either Boston or San Francisco, our sample's two largest MSAs. Dropping either MSA has little effect on economic significance in the OLS and IV specifications. The OLS results remain highly statistically significant, but the IV results become significant at only the 5% confidence level, due largely to a weaker first-stage regression.

Next, we address four potential identification challenges. The first challenge is that our instrument, based on geographic proximity, may simply be picking up the effects of information sharing. Some types of information are more likely to flow between startups that are nearby. This concern, however, is not likely to hold for the type of information we study, namely, whether a project progresses from Phase I to II. Progressing to Phase II is public information revealed through clinical trial registries, FDA submissions, patent filings, and company press releases. Even if this information were shared privately between nearby startups, the bias would likely work against our main result. If drug A experiences success, that sends a positive signal about the scientific underpinnings of nearby competing drug B, making it *more* likely that drug B progresses to Phase II.[21] We instead find that progress by A is *less* likely to predict progress by

[21] Supporting this logic, Krieger (2017) finds that a drug project is more likely to be terminated after a close competitor terminates its own project, implying that technological learning dominates competition effects.

B if the two companies are nearby, consistent with common owners wishing to avoid duplication or R&D.

A related concern is that nearby competitors have opposite outcomes not because of common ownership, but because success by one drug project makes its local competitors optimally "give up" earlier because they will not be first to market. This story is implausible for Phase I drug projects, for two reasons. First, as discussed above, a competitor's success sends a positive signal, not a negative signal, about the lagging project's scientific fundamentals, making the lagging project less likely to give up. Second, Phase I projects are far from reaching the market. There is still a good chance that the lagging project can catch up to the pioneer, so giving up is potentially suboptimal.

We also address this concern empirically in two ways. First, if this concern were valid, then even in the absence of a shared VC, we should see a negative relation between $Progress$ and success by a close competitor ($Shocked$). We do not see this in the data. Instead, we see a slope of basically zero on $Shocked$ in column 1 of Table 3. Second, to address the concern that local competitors are more likely to give up because they are more technologically similar, we control for overlapping patent citations. Specifically, we repeat our main panel regressions while controlling for $Shocked \times SharedCite$, the indicator for whether the shocked company's drug shares at least one citation in common with the progressing drug. OLS regressions still produce a significantly negative slope on $Shocked \times SharedVC$ when we control for $Shocked \times SharedCite$, for both project and funding outcomes (Table 10, columns 1 and 3). In the Online Appendix, we show that our IV results continue to hold after controlling for $Shocked \times SharedCite$. Also interesting, $Progress$ is positively related to $Shocked \times SharedCite$, consistent with the argument above that seeing a close competitor make progress, even absent a shared VC, is a positive signal, not a negative signal. Overall, these results indicate that our instrument is not simply picking up the omitted effects of technological similarity.

A third concern is that our proximity instrument simply picks up the ease of poaching employees from a competing startup (e.g., Hombert and Matray, 2017). For example, if startup A experiences success, it could poach employees from a nearby, competing startup B, causing B

to fail. This story can in theory explain why nearby startups are more likely to have opposite outcomes. We offer two counterarguments. First, there is an opposing force. Success by A sends a positive signal about B (explained above), making startup B and its employees less willing to separate from each other, which makes it harder for A to poach B's employees. Second, if the competing story were true, then our results should be weaker in states that strongly enforce employee non-compete agreements, because poaching is less feasible there.[22] The evidence in Table 12 does not support that prediction, however. We create a dummy variable for whether non-compete agreements are strongly enforced in startup i's state in quarter t. We then interact that non-compete dummy variable with $Shocked \times SharedVC$ in our main project-outcome regressions. If the competing story were true, we should find a significantly positive slope on the triple-interaction term in column 1. Instead, it has a slightly negative coefficient with a t-statistic of -0.3. This result indicates that employee poaching is not spuriously generating our main results.

A final concern is that our results simply reflect that VCs follow a "spray and pray" investing approach in certain markets (Ewens et al., 2018). Suppose in some product markets exactly one project will ultimately succeed, which makes projects' outcomes negatively correlated. If VCs "spray and pray" in such markets, then these markets will feature more common ownership, so common ownership will be associated with projects' reaching opposite outcomes. This story cannot explain our results, however. Even our simplest OLS regressions include ICD FEs, which difference out any effects of a negative correlation at the product-market level. The story is even less likely in our test with ICD-quarter FEs, which difference out potentially time-varying correlations within product markets. Also, this story cannot explain our IV results, because it is not clear why nearby drug startups would be more likely to belong to a winner-take-all market.

[22] Supporting this idea, Ewens and Marx (2018), Jeffers (2019), and others confirm empirically that non-compete agreements restrict worker mobility, especially in knowledge-intensive sectors.

4. Conclusions

We study the effects of common ownership on innovation. Our evidence comes from detailed project-level data on pharmaceutical startups and their VC investors. Common ownership by VCs is likely to have strong effects, as VCs have significant control rights and are likely to be attentive to spillovers within their portfolios. We examine how a startup responds after seeing a competitor make progress on a closely related drug project. If the two startups share a common VC, the lagging startup is less likely to advance its own project and obtain VC funding. Also, commonly owned startups redirect their innovation away from areas where they have fallen behind their commonly owned peers. Taken together, these results suggest that common owners reduce inefficient duplication of R&D costs, which can help solve a well known market failure in patent races (Loury, 1979). Consistent with common ownership improving innovation efficiency, common ownership is positively correlated with the number of approved drugs per dollar of aggregate VC funding.

Our results apply to the pharmaceutical industry, which features patent races, high experimentation costs, and well established business models. Eldar et al. (2020) examine a broad range of industries and find that common VCs facilitate growth. Understanding how common ownership affects innovation differently across industries is an interesting area for future work. More research is also needed on how common ownership affects welfare, including through its effects on innovation. While we emphasize a social benefit from common ownership, there can also be costs from reduced product market competition.

References

ABPI (2019). Worldwide pharmaceutical company R&D expenditure by country.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Antón, M., Ederer, F., Giné, M., and Schmalz, M. (2018). Innovation: The bright side of common ownership? *Working paper*.

Azar, J., Raina, S., and Schmalz, M. C. (2019). Ultimate ownership and bank competition. *Working Paper*.

Azar, J., Schmalz, M. C., and Tecu, I. (2018). Anticompetitive effects of common ownership. *Journal of Finance*, 73(4):1513–1565.

Bloom, N., Schankerman, M., and Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, 81(4):1347–1393.

Cunningham, C., Ederer, F., and Ma, S. (2020). Killer acquisitions. *Forthcoming, Journal of Political Economy*.

Dennis, P. J., Gerardi, K., and Schenone, C. (2019). Common ownership does not have anticompetitive effects in the airline industry. *Working Paper*.

Dinc, S. I. and Erel, I. (2013). Economic nationalism in mergers and acquisitions. *Journal of Finance*, 68(6):2471–2514.

Eldar, O., Grennan, J., and Waldock, K. (2020). Common ownership and startup growth. *Working paper*.

Ewens, M. and Marx, M. (2018). Founder replacement and startup performance. *Review of Financial Studies*, 31(4):1532–1565.

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

Ewens, M., Rhodes-Kropf, M., and Strebulaev, I. A. (2016). Insider financing and venture capital returns. *Working Paper*.

Fulghieri, P. and Sevilir, M. (2009). Size and focus of a venture capitalist's portfolio. *Review of Financial Studies*, 22(11):4643–4680.

Gerakos, J. J. and Xie, J. (2019). Institutional horizontal shareholdings and generic entry in the pharmaceutical industry. *Working Paper*.

Gilje, E. P., Gormley, T. A., and Levit, D. (2019). Who's paying attention? measuring common ownership and its impact on managerial incentives. *Journal of Financial Economics*, Forthcoming.

Gompers, P. A., Gornall, W., Kaplan, S. N., and Strebulaev, I. A. (2019). How do venture capitalists make decisions? *Journal of Financial Economics*, Forthcoming.

Gompers, P. A. and Lerner, J. (2004). *The venture capital cycle*. MIT press.

González-Uribe, J. (2019). Exchanges of innovation resources inside venture capital portfolios. *Journal of Financial Economics*, Forthcoming.

Gramlich, J. and Grundl, S. (2017). Testing for competitive effects of common ownership. *Working Paper*.

Grossman, G. M. and Shapiro, C. (1987). Dynamic r&d competition. *The Economic Journal*, 97:372–387.

Gutiérrez, G. and Philippon, T. (2017). Investmentless growth: An empirical examination. *Brookings Papers on Economic Activity*, pages 89–169.

Harrington, S. E. (2012). Cost of capital for pharmaceutical, biotechnology, and medical device firms. *The Oxford handbook of the economics of the biopharmaceuticals industry*, pages 75–99.

He, J. J. and Huang, J. (2017). Product market competition in a world of cross-ownership: Evidence from institutional blockholdings. *Review of Financial Studies*, 30(8):2674–2718.

Hombert, J. and Matray, A. (2017). The real effects of lending relationships on innovative firms and inventor mobility. *Review of Financial Studies*, 30(7):2413–2445.

Hombert, J. and Matray, A. (2018). Can innovation help us manufacturing firms escape import competition from china? *The Journal of Finance*, 73(5):2003–2039.

Ivy, S. P., Siu, L. L., Garrett-Mayer, E., and Rubinstein, L. (2010). Approaches to phase 1 clinical trial design focused on safety, efficiency, and selected patient populations: A report from the clinical trial design task force of the national cancer institute investigational drug steering committee. *Clinical Cancer Research*, 16(6):1726–1736.

Jeffers, J. (2019). The impact of restricting labor mobility on corporate investment and entrepreneurship. *Working Paper*.

Jiang, W. (2017). Have instrumental variables brought us closer to the truth. *Review of Corporate Finance Studies*, 6(2):127–140.

Kaplan, S. N. and Lerner, J. (2010). It ain't broke: The past, present, and future of venture capital. *Journal of Applied Corporate Finance*, 22(2):36–47.

Kennedy, P., O'Brien, D. P., Song, M., and Waeher, K. (2017). The competitive effects of common ownership: Economic foundations and empirical evidence. *Working paper*.

Kini, O., Lee, S., and Shen, M. (2018). Common institutional ownership and product market threats. *Working paper*.

Koch, A., Panayides, M. A., and Thomas, S. (2020). Common ownership and competition in product markets. *Journal of Financial Economics*, Forthcoming.

Kostovetsky, L. and Manconi, A. (2020). Common institutional ownership and diffusion of innovation. *Working paper*.

Krieger, J. L. (2017). Trials and terminations: Learning from competitors' R&D failures. *Working Paper*.

Lerner, J. (1995). Venture capitalists and the oversight of private firms. *Journal of Finance*, 50(1):301–318.

Lewellen, J. and Lewellen, K. (2017). Institutional investors and corporate governance: The incentive to increase value. *Working paper*.

Lewellen, K. and Lowry, M. (2019). Does common ownership really increase firm coordination? *Working paper*.

Lindsey, L. (2008). Blurring firm boundaries: The role of venture capital in strategic alliances. *Journal of Finance*, 63(3):1137–1168.

López, A. L. and Vives, X. (2019). Overlapping ownership, R&D spillovers, and antitrust policy. *Journal of Political Economy*, 127(5):2394–2437.

Loury, G. C. (1979). Market structure and innovation. *The Quarterly Journal of Economics*, 93:395–410.

Matray, A. (2020). The local innovation spillovers of listed firms. *Forthcoming, Journal of Financial Economics*.

Newham, M., Seldeslachts, J., and nol, A. B.-E. (2018). Common ownership and market entry: Evidence from the pharmaceutical industry. *Working Paper*.

Petersen, M. A. and Rajan, R. G. (1994). The benefits of lending relationships: Evidence from small business data. *Journal of Finance*, 49(1):3–37.

Rajan, R. G. (1992). Insiders and outsiders: The choice between informed and arm's-length debt. *Journal of Finance*, 47(4):1367–1400.

Roberts, T. G., Goulart, B. H., Squitieri, L., Stallings, S. C., Halpern, E. F., Chabner, B. A., Gazelle, G. S., Finkelstein, S. N., and Clark, J. W. (2004). Trends in the risks and benefits to patients with cancer participating in phase 1 clinical trials. *Journal of the American Medical Association*, 292(17):2130–2140.

Saretto, A. and Tookes, H. E. (2013). Corporate leverage, debt maturity, and credit supply: The role of credit default swaps. *Review of Financial Studies*, 26(5):1190–1247.

Seyhan, A. A. (2019). Lost in translation: The valley of death across preclinical and clinical divide– ididentification of problems and overcoming obstacles. *Tranlational Medicine Communications*, 4:1–19.

Thomas, D. W., Burns, J., Audette, J., Carroll, A., Dow-Hygelund, C., and Hay, M. (2016). Clinical development success rates 2006–2015. *BIO Industry Analysis*, https://www.bio.org.

Wessner, C. W. (2009). *Venture funding and the NIH SBIR program*. National Academies Press.

Wong, C. H., Siah, K. W., and Lo, A. W. (2019). Estimatoin of clinical trial success rates and related parameters. *Biostatistics*, 20:273–286.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Appendix: Case study

New England Associates (NEA) is a Maryland-based VC firm that, in 2018, was ranked second in Forbes' list of most-active healthcare VCs.[23] In July 2012, NEA was an investor in two Boston-based startups, Intarcia and Rhythm Pharmaceuticals. Both startups at that time had drug projects that were targeting obesity and that were in Phase I clinical trials. Rhythm's obesity-related project, setmelanotide, corresponds to the pioneer project in our paper. That project progressed from Phase I to Phase II in December 2012. This event corresponds to the "shock" in our main tests. Rhythm eventually went public in 2017, shortly after beginning Phase III trials, and setmelanotide obtained FDA approval in 2020. Intarcia's obesity-related project, named "weight regulating human endocrine peptide" in the Cortellis database, corresponds to the lagging project in our paper. That project never progressed to Phase II, and updates on the project stopped in 2016, indicating the project was abandoned. NEA's last investment in Intarcia occurred in November 2012, so NEA stopped funding Intarcia after the "shock" from Rhythm. Since then, Intarcia has mainly focused on its diabetes pipeline related to the drug exenatide. In 2015 Intarcia redirected some of its earlier obesity-related research into two new projects, "exenatide + optimized peptide 1" and "exenatide + optimized peptide 2," both of which target "type 2 diabetes and/or obesity."

[23] https://www.forbes.com/sites/michelatindera/2018/12/27/these-10-vc-firms-made-the-most-investments-in-healthcare-startups-this-year/?sh=5f4321485498

Figure 2: Geographic Distribution of Drug Projects

This figure shows the number of drug projects in our sample in each state.

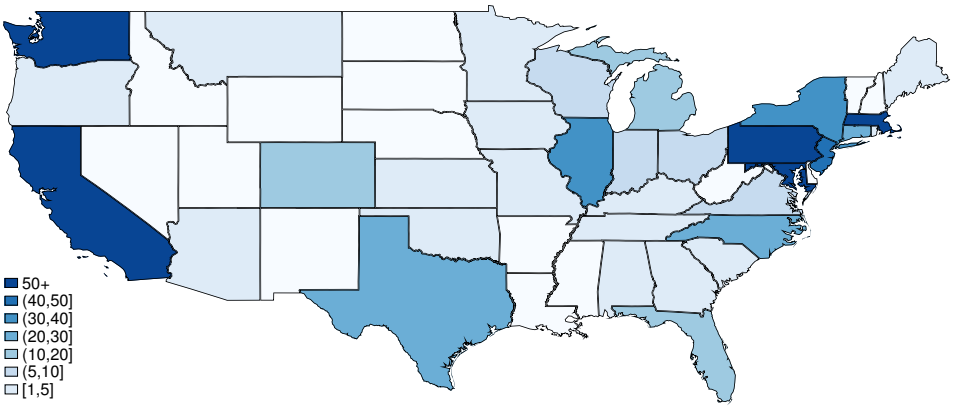


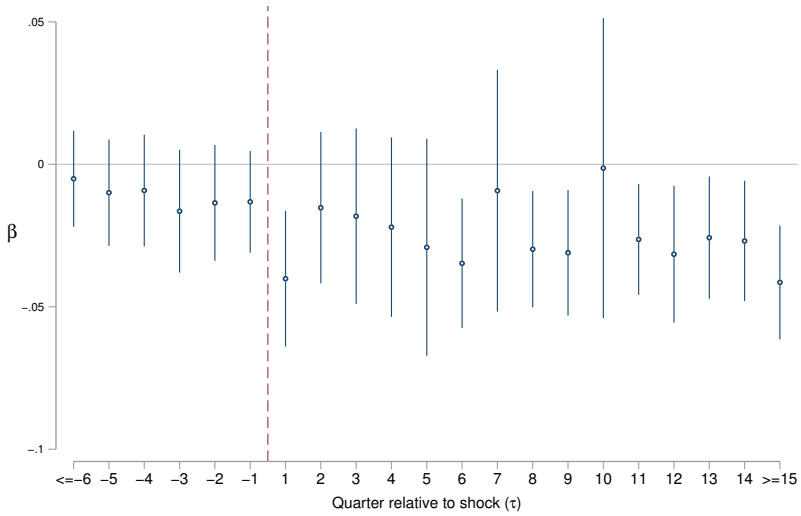
Figure 4: Coefficient Dynamics

This figure presents the coefficients dynamics in an event-study framework. Panel A contains results for project outcomes. It plots the OLS coefficients β_τ from the following regression:

$$Progress_{i,t} = \sum_{\tau=-6}^{15+} \gamma_\tau Shocked_{i,\{t-\tau_0=\tau\}} + \sum_{\tau=-6}^{15+} \beta_\tau Shocked \times SharedVC_{i,\{t-\tau_0=\tau\}} + FEs + Controls + \varepsilon_{i,t}.$$

where τ_0 is the first quarter when $Shocked \times SharedVC = 1$ for project i. The x-axis denotes τ, the number quarters after $Shocked \times SharedVC$ first equals one for a given project. $\tau = 0$ is the excluded category. For any event time that is more than 6 quarters before the shock, they are binned into the $-6th$ quarter and denote as $\tau <= -6$ in the figure. For any event time that is more than 15 quarters after the shock, they are binned into the $15th$ quarter and denoted as $\tau >= 15$ in the figure. Panel B plots results for VC financing outcomes. To create Panel B, we replace $Progress_{it}$ by $ExtendFunding_{ijt}$ in the regression above and plots the resulting estimates β_τ. Remaining details are the same as in Tables 3 and 5.

Panel A: Project Outcomes



Panel B: VC Financing Outcomes

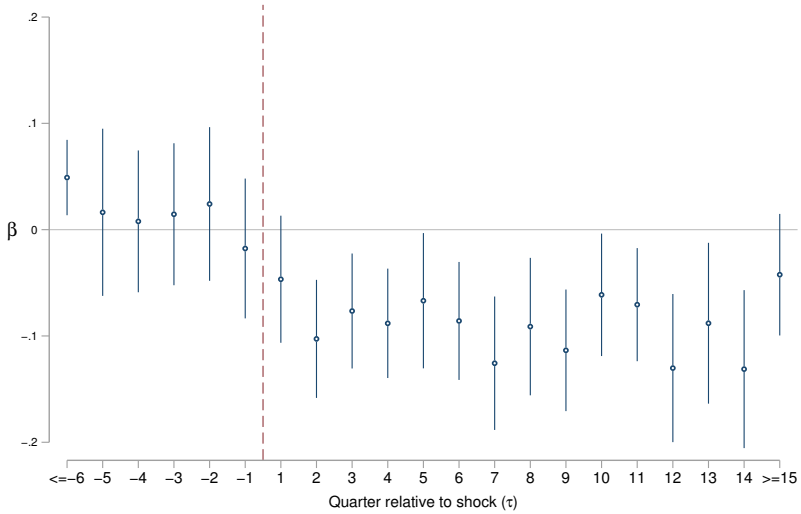


Figure 5: Innovation Efficiency

This figure shows the relation between innovation efficiency and the common ownership rate. Each circle corresponds to an ICD category. An ICD category is included in the sample if it has at least one drug initiated later than 2005Q1 by a VC-funded startup. We exclude ICD categories with unspecified sites, unknown causes, or uncertain effects. The y-axis denotes the efficiency proxy, defined as number of FDA approved drugs divided by total VC funding (millions of dollars), where the numerator and denominator are both aggregated across startups and over time within the ICD. We include VC funding through all drug development phases, not just Phase I, which extends sample of ICDs to 94 (rather than 78 in our main regression sample). If a startup is operating in multiple ICDs, we allocate the startup's VC funding to each ICD proportionally based on the number of its projects in that ICD. The x-axis denotes the ICD's *Common Ownership Rate*, defined as the probability that a randomly chosen pair of startups within the ICD shares at least one VC in common. The dashed line denotes the best-line; the underlying regression is in column 1 of Table 9.

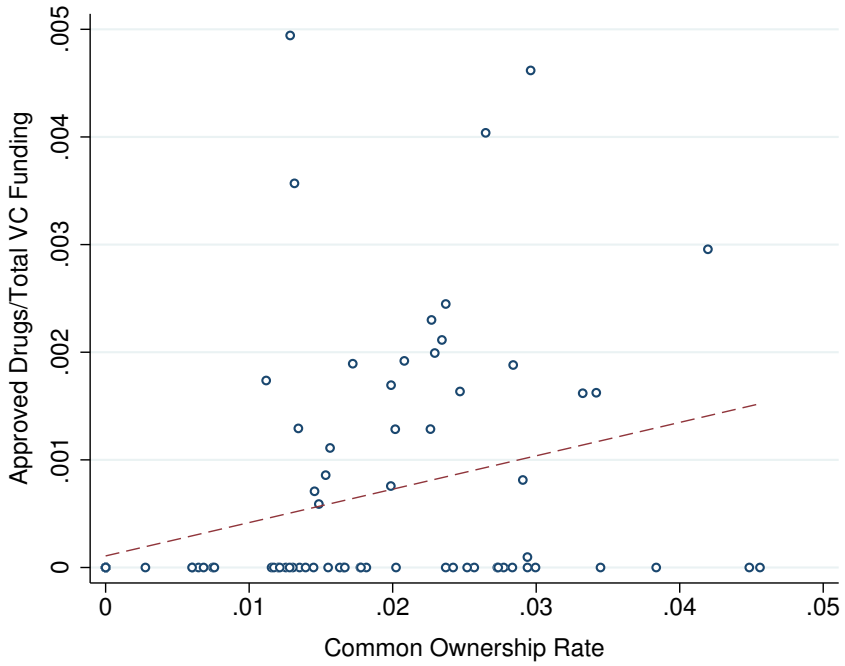


Table 1: Summary Statistics

This table contains summary statistics for our sample of Phase I drug projects in U.S. VC-backed startups. The unit of observation is the project×quarter in Panel A, and the startup company×quarter in Panel B. We show summary statistics for the full sample and the subsamples of observations that are "Never Treated" versus "Ever Treated." The last column reports the difference between those subsamples. Projects/startup companies are categorized as "Never Treated" if the variable $Shocked \times SharedVC$ is equal to zero across its lifespan, and "Ever Treated'" otherwise. $Shocked$ is an indicator for whether a different project (at a different startup) in the same ICD category progresses from phase I to II between the given project's initiation quarter and the previous quarter. $Progress$ is an indicator for whether a project progresses to Phase II in a given quarter. Age equals the number of quarters since the project's initiation. $Number\,of\,Projects$ is the number of projects being developed within the startup company at all clinical stages in the given quarter. $Number\,of\,ICD\,Categories$ is the number of distinct ICD categories covered by those projects. $Number\,of\,VCs$ is the number of VC firms invested in the startup. Ln($Size\,of\,Last\,VC\,Round$) is the log dollar amount that the startup raised in its most recent VC financing round. $*$, $**$, and $***$ denote statistical significance at the 10%, 5%, and 1% level, respectively, two-way clustering at ICD category and startup company levels.

	Full Sample						Never Treated	Ever Treated	
Variable	*Mean*	*Std Dev*	*P25*	*Median*	*P75*	*Obs*	*Mean*	*Mean*	*Difference*
Panel A: Project-Level Variables									
Shocked	0.504	0.500	0.000	1.000	1.000	12,481	0.460	0.762	−0.302***
Progress	0.018	0.132	0.000	0.000	0.000	12,481	0.020	0.005	0.015***
Age	13.444	9.266	6.000	12.000	19.000	12,481	13.305	14.267	−0.962
Panel B: Startup-Level Variables									
Number of Projects	1.659	1.374	1.000	1.000	2.000	7,745	1.617	2.292	−0.675*
Number of ICD Categories	1.419	0.882	1.000	1.000	2.000	7,745	1.395	1.781	−0.386
Number of VCs	4.688	3.924	2.000	4.000	6.000	7,745	4.698	4.537	0.161
Ln(*Size of Last VC Round*)	8.992	1.522	8.161	9.313	10.096	6,690	9.007	8.780	0.227

Table 2: Frequency of Common Ownership by VCs

Row 1 reports the percent of startup-quarter observations in which the startup has a close competitor. Startup i is defined as having a close competitor in a quarter if there is another startup in the same quarter developing a drug in Phase I in the same ICD category as one of i's Phase I drugs. Row 2 contains the fraction of startup-quarter observation in which the startup has a close competitor, and that close competitor shares at least one VC in common with the startup in question. Row 3 reports the average number of distinct VCs that own stakes in Phase I drug projects in a given ICD category, averaging across drug categories and quarters. Row 4 reports the average number of distinct ICD categories in which a typical VC owns a stake in a startup with a Phase I project, averaging across VCs and quarters. Row 5 contains the percent of competing startup pairs that share a common VC. Specifically, we count the unique pairs of startups in a given ICD category and quarter that share a VC, and we divide it by the number of possible unique pairs of startups in the ICD category and quarter; we then average this ratio across quarters and ICD categories.

Percent of startups with a close competitor	93.2%
Percent of startups with a close competitor held by same VC	38.6%
Average number of VCs per drug category	16.4
Average number of drug categories per VC	2.52
Percent of competing startup pairs with a common VC	8.6%

Table 3: Project Outcomes

This table contains estimates of regression (1). The unit of observation is the project by quarter. The sample contains all project-quarters that are in Phase I trials. Dependent variables are indicated in the column titles. $Progress_{it}$ is an indicator for whether project i progresses to Phase II in quarter t. $Shocked_{it}$ is an indicator for whether another project (at a different startup) in the same ICD category as project i progresses from phase I to II between project i's initiation quarter and quarter $t-1$. $SharedVC_{it}$ is an indicator for whether that progressing project shares a VC with project i in quarter t. $\ln(Age_{it})$ is the log of one plus the number of quarters since the project's i initiation. $NProjects_{it}$ is the number of Phase I projects being developed in project's i's company during quarter t. $NVCs_{it}$ is the number of VC firms that own a stake in project i's startup company in quarter t. $NProjectsperICD_{it}$ is the number of Phase I projects being developed in project's i's ICD during quarter t. Column 1 reports results from OLS regressions. Column 2 reports results from estimating a probit model with dependent variable $Shocked \times SharedVC$ and independent variable $Shocked \times Proximity$, which is defined in equation (2) and captures the distance between project i and the project causing the shock. The probit model's predicted probabilities, denoted $\widehat{Shocked \times SharedVC}$, form the instrument in the first-stage regression (column 3). Columns 5 and 6 report results from the bivariate probit method, with standard errors clustered by ICD category. Column 7 reports results from the OLS regression that includes $ICD \times Quarter$ FEs. Standard errors are computed by two-way clustering at the ICD category and startup company levels in the OLS and 2SLS regressions. t-statistics are in parentheses. FEs are noted in the bottom row. ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS *Progress*	**Probit** *Shocked× SharedVC*	**1st Stage** *Shocked× SharedVC*	**2SLS** *Progress*	**Biprobit** *Shocked× SharedVC*	**Biprobit** *Progress*	**OLS** *Progress*
$Shocked \times SharedVC$	-0.019***			-0.070***		-0.038*	-0.018**
	(-2.95)			(-3.82)		(-1.90)	(-2.37)
$\widehat{Shocked \times SharedVC}$			0.721***				
			(7.55)				
$Shocked \times Proximity$		0.447***			0.445***		
		(7.17)			(7.05)		
$Shocked$	0.000		0.101***	0.007		0.000	
	(0.16)		(3.40)	(1.56)		(0.04)	
$\ln(Age)$	0.009***	0.208**	-0.016**	0.010***	0.212**	-0.053	0.006*
	(2.89)	(2.06)	(-2.17)	(3.03)	(2.05)	(-1.28)	(1.74)
$NProjects$	0.002**	0.001	0.005	0.002**	0.000	-0.022	0.002**
	(2.20)	(0.04)	(1.47)	(2.48)	(0.02)	(-1.35)	(2.06)
$NVCs$	0.000	0.084***	-0.012*	0.000	0.084***	0.023**	-0.001
	(0.10)	(7.26)	(-1.88)	(0.05)	(7.24)	(2.26)	(-0.43)
$NProjectsperICD$	-0.034**	-0.061	-0.008	-0.034**	-0.060	-0.094**	
	(-2.53)	(-0.62)	(-0.20)	(-2.49)	(-0.61)	(-2.47)	
Startup FE	Yes	No	Yes	Yes	No	No	Yes
Yr-Qtr FE	Yes	No	Yes	Yes	No	No	No
ICD FE	Yes	No	Yes	Yes	No	No	No
ICD × Qtr. FE	No	No	No	No	No	No	Yes
N	12,469	12,481	12,469	12,469	12,481	12,481	11,507
Adj. R^2	0.073		0.553				0.078

Table 4: Project Outcomes by Excluding Ones with Adverse Readouts in Phase I

This table repeats the estimation in Table 3, except now we restrict the sample to drug projects that have no adverse readouts from Phase I clinical trials. Specifically, we match each drug project in the sample with detailed clinical trials information from the Clinical Trials Database of Cortellis. We exclude projects with adverse events (e.g., death of trial participants) in the process of clinical trials and those failed to reach their primary endpoints. In total, 6,880 observations involving 525 drug projects are dropped. All other details are the same as the previous table.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS *Progress*	***Probit*** *Shocked× SharedVC*	***1st Stage*** *Shocked× SharedVC*	***2SLS*** *Progress*	**Biprobit** *Shocked× SharedVC*	**Biprobit** *Progress*	***OLS*** *Progress*
$Shocked \times SharedVC$	-0.039**			-0.207**		-0.135*	-0.042**
	(-2.36)			(-2.61)		(-1.76)	(-2.19)
$Shocked \times \widehat{SharedVC}$			0.838***				
			(3.34)				
$Shocked \times Proximity$		0.308***			0.301***		
		(3.45)			(3.14)		
$Shocked$	0.004		0.114***	0.025		0.002	
	(0.57)		(3.22)	(1.66)		(0.03)	
$\ln(Age)$	0.029***	0.224**	-0.013	0.030***	0.241**	0.003	0.025**
	(3.99)	(2.10)	(-1.13)	(4.41)	(2.12)	(0.04)	(2.50)
$NProjects$	0.003	0.014	-0.004	0.002	0.012	-0.015	0.002
	(1.47)	(0.56)	(-1.59)	(1.17)	(0.51)	(-1.08)	(0.97)
$NVCs$	0.008	0.091***	-0.014**	0.008	0.092***	0.035**	0.005
	(1.21)	(5.00)	(-2.47)	(1.11)	(5.06)	(2.00)	(0.64)
$NProjectsperICD$	-0.082***	-0.005	0.002	-0.079**	-0.002	-0.039	
	(-2.82)	(-0.05)	(0.03)	(-2.58)	(-0.03)	(-0.95)	
Startup FE	Yes	No	Yes	Yes	No	No	Yes
Yr-Qtr FE	Yes	No	Yes	Yes	No	No	No
ICD FE	Yes	No	Yes	Yes	No	No	No
ICD × Qtr. FE	No	No	No	No	No	No	Yes
N	5,592	5,601	5,592	5,592	5,601	5,601	4,578
Adj. R^2	0.088		0.563				0.098

Table 5: VC Funding Outcomes

This table reports results from regression (3). The sample includes all pairs of VC firms j that invest in startups i. For each pair $\{i, j\}$, we include quarters t when the startup has at least one project in Phase I trials and the quarter is between j's first investment in i and i's exit (e.g., IPO, trade sale). The dependent variable, $ExtendFunds_{ijt}$, is an indicator for whether VC j invests in startup i in quarter t. $Shocked_{it}$ equals one if at least one project in startup i has seen another project (a "pioneer") in the same ICD category (but different startup) progress to Phase II between the former project's inception and quarter $t-1$. $Shocked \times SharedVC_{ijt}$ equals one if VC j holds a stake in both startup i and at least one of the startups owning the pioneering projects that produced $Shocked_{it} = 1$. $SelfProgress_{ijt}$ equals 1 if startup i has progressed at least one drug project to the next stage of clinical trials between the quarter when VC j first invested in i and quarter $t-1$; we include this variable as a proxy for startup performance. $NCats_{jt}$ is the number of ICD categories covered by all of VC j's portfolio companies during quarter t; this variable measures the VC's diversification. $Duration_{ijt}$ is the log of one plus the number of years since startup i last received funding from VC j. $NProjects_{it}$ is the number of projects under development at startup i in quarter t. $NVCs_{it}$ is the number of VC firms that own a stake in startup i at the beginning of quarter t. $PortfolioSize_{jt}$ is the number of startups in VC j's portfolio in quarter t. $PrevRoundSize_{it}$ is the log dollar amount that startup i raised in its most recent VC financing round before quarter t. Column 1 contains results from OLS. Columns 2–4 contain results from a probit-2SLS specification similar to those in previous tables. Columns 5 and 6 report results from the bivariate probit model. Column 7 reports results from the OLS regression by controlling $Startup \times Quarter$ FEs. Standard errors are clustered at the startup company level. We report $t-$statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS *ExtendFunds*	**Probit** *Shocked× SharedVC*	1*st* **Stage** *Shocked× SharedVC*	**2SLS** *ExtendFunds*	**Biprobit** *Shocked× SharedVC*	**Biprobit** *ExtendFunds*	**OLS** *ExtendFunds*
$Shocked \times SharedVC$	-0.030**			-0.133**		-0.141***	-0.032***
	(-2.26)			(-2.38)		(-9.20)	(-2.77)
$Shocked \times \widehat{SharedVC}$			0.808***				
			(6.00)				
$Shocked \times Proximity$		0.269***			0.269***		
		(8.83)			(8.83)		
$Shocked$	0.098***		0.018***	0.099***		0.235***	0.275***
	(9.91)		(3.21)	(9.86)		(15.55)	(12.00)
$SelfProgress$	0.022*	0.238	0.001	0.022*	0.238	1.700***	0.144***
	(1.75)	(1.43)	(0.10)	(1.75)	(1.43)	(7.84)	(5.16)
$NCats$	-0.005***	-0.004	0.006**	-0.005***	-0.004	-0.018	-0.006***
	(-3.03)	(-0.13)	(2.00)	(-2.69)	(-0.13)	(-0.71)	(-4.25)
$NProjects$	0.001	0.003	-0.006	0.002	0.003	-0.061	
	(0.31)	(0.09)	(-1.48)	(0.38)	(0.09)	(-0.76)	
$NVCs$	0.002	0.014	-0.000	0.002	0.015	0.016	
	(0.93)	(1.10)	(-0.19)	(0.89)	(1.10)	(1.39)	
$PortfolioSize$	0.005**	0.096***	-0.011**	0.005**	0.096***	0.075***	0.005***
	(2.22)	(2.91)	(-2.24)	(2.26)	(2.93)	(2.69)	(2.77)
$Duration$	-0.213***	0.018	-0.007	-0.214***	0.017	-20.008***	-0.272***
	(-27.00)	(0.31)	(-1.28)	(-26.82)	(0.31)	(-24.49)	(-28.11)
$PrevRoundSize$	-0.001	0.006	-0.007	-0.002	0.006	-0.105***	
	(-0.32)	(0.12)	(-1.54)	(-0.54)	(0.12)	(-2.90)	
VC Firm FE	Yes	No	Yes	Yes	No	No	Yes
Startup FE	Yes	No	Yes	Yes	No	No	No
Yr-Qtr FE	Yes	No	Yes	Yes	No	No	No
Startup× Qtr. FE	No	No	No	No	No	No	Yes
N	32,537	31,313	31,298	31,298	31,313	31,313	34,414
Adj. R^2	0.344		0.342				0.595

Table 6: Aggregate Funding Outcomes for Lagging Firms

The unit of observation in columns 1–3 is the startup (i) by quarter (t). The dependent variable in column 1 is $ExtendFundAgg$, an indicator for whether any VC invests in startup i in quarter t. The dependent variable in column 2 is $AmountRaised$, the percentage increase in cumulative VC funding during quarter t for startup i. Specifically, $AmountRaised$ equals $\log(1+NewFunds_{it}/CumFunds_{i,t-1})$, where $NewFunds_{it}$ is the dollar amount raised by startup i from all VCs in quarter t, and $CumFunds_{i,t-1}$ is the cumulative amount of VC funding raised by firm i through the end of quarter $t-1$. The dependent variable in column 3 is $NewInvestors$, an indicator for whether startup i raises funding from any new VC during quarter t. In columns 1–3, $Shocked_{it}$ equals one if at least one project in startup i has seen another project (a "pioneer") in the same ICD category (but different startup) progress to Phase II between the former project's inception and quarter $t-1$, and $Shocked \times SharedVC_{it}$ equals one if there is any VC that holds a stake in startup i and at least one of the startups owning the pioneering projects that produced $Shocked_{it} = 1$. Control variables $SelfProgress_{it}$, $Duration_{it}$, $NProjects_{it}$, $NVCs_{it}$, and $PrevRoundSize_{it}$ are defined as in Table 5, except we collapse them to the startup-quarter level in columns 1–3. Column 4 resembles the specification of column 1 in Table 5, which is at VC-startup-quarter level. The dependent variable is $Extendfunds_{ijt}$, an indicator for whether VC j invests in startup i in quarter t. One distinction is that our independent variable $Shocked \times SharedVC_{it}$ is collapsed to the startup-quarter level, as in columns 1–3. Additionally, we interact our focal variables with $NonCommonVC_{ijt}$, a dummy for whether VC j is not a common owner of a pioneering project, if there is any, competing with startup i in quarter t. We estimate all models by OLS. Standard errors are clustered at the startup company level. FEs are noted in the bottom row. ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) $ExtendFundsAgg$	(2) $AmountRaised$	(3) $NewInvestors$	(4) $ExtendFunds$
$Shocked \times SharedVC$	-0.086***	-0.075***	-0.024*	-0.059***
	(-3.42)	(-4.60)	(-1.70)	(-3.09)
$Shocked \times SharedVC \times NonCommonVC$				0.018
				(0.96)
$Shocked$	0.112***	0.015	-0.018*	0.085***
	(6.62)	(0.99)	(-1.72)	(4.28)
$Shocked \times NonCommonVC$				-0.011
				(-0.60)
$SelfProgress$	0.008	0.019	-0.033**	0.017
	(0.35)	(1.06)	(-2.22)	(1.62)
$NProjects$	0.012	-0.003	0.001	0.002
	(1.51)	(-0.43)	(0.23)	(0.37)
$NVCs$	0.030***	0.039***	0.021***	0.014***
	(4.55)	(4.39)	(4.83)	(5.17)
$Duration$	-0.230***	-0.042***	-0.047***	-0.122***
	(-23.79)	(-4.87)	(-8.64)	(-19.28)
$PrevRoundSize$	-0.013**		0.014***	-0.009**
	(-2.01)		(4.10)	(-2.49)
$\log(CumFunds_{i,t-1})$		-0.276***		
		(-4.35)		
$NCats$				-0.003**
				(-2.05)
$PortfolioSize$				0.002
				(1.16)
VC Firm FE	No	No	No	Yes
Startup FE	Yes	Yes	Yes	Yes
Yr-Qtr FE	Yes	Yes	Yes	Yes
N	7,704	7,640	7,704	34,045
Adj. R^2	0.371	0.290	0.079	0.182

Table 7: VC Influence and Diversification

This table presents results from estimating the VC-funding regression (3) in subsamples based on either the VC's financial commitment to the startup or a measure of the VC's portfolio diversification. The dependent variable is an indicator for whether VC j extends funding to startup i in quarter t. Columns 1 and 2 compare results across subsamples of non-lead and lead VCs, where a lead VC is defined as the VC whose total amount invested to date is the highest across all the startup's VCs. Our data report the amount invested by each VC syndicate but not by each syndicate member. In cases where these missing data create ambiguity about the lead-VC measure, we assume all syndicate members invest equal amounts. Columns 3 and 4 compare results across subsamples of low- versus high-diversification VCs, where a high-diversification VC is defined as one whose average active portfolio size is above the sample median (19.6). Active portfolio size is the number of unique startups (all industries) invested by the VC in the past five years. This table reports OLS estimates. Remaining details and variable definitions are the same as in Table 5.

	(1)	(2)	(3)	(4)
	Non-Lead	**Lead**	**Low Diversification**	**High Diversification**
$Shocked \times SharedVC$	-0.005	-0.083***	-0.049*	-0.009
	(-0.40)	(-3.35)	(-1.87)	(-0.86)
$Shocked$	0.052***	0.099***	0.039***	0.079***
	(8.23)	(8.49)	(6.17)	(10.19)
$SelfProgress$	0.012*	0.004	0.019***	0.011
	(1.76)	(0.22)	(2.83)	(0.93)
$NCats$	-0.002	-0.006	0.000	-0.003*
	(-1.28)	(-1.55)	(0.04)	(-1.70)
$NProjects$	0.000	0.006	0.001	0.001
	(0.22)	(0.88)	(0.26)	(0.22)
$NVCs$	0.011***	0.022***	0.012***	0.014***
	(6.58)	(2.97)	(4.18)	(6.45)
$PortfolioSize$	-0.001	0.010***	-0.002	0.002
	(-0.28)	(3.20)	(-0.53)	(0.95)
$Duration$	-0.099***	-0.175***	-0.072***	-0.156***
	(-17.86)	(-13.83)	(-12.99)	(-16.82)
$PrevRoundSize$	-0.009***	-0.006	-0.003	-0.012***
	(-2.74)	(-0.59)	(-0.67)	(-2.78)
VC Firm FE	Yes	Yes	Yes	Yes
Startup FE	Yes	Yes	Yes	Yes
Qtr FE	Yes	Yes	Yes	Yes
N	24,532	7,988	16,726	15,804
Adj. R^2	0.139	0.229	0.105	0.217

Table 8: Redirecting Innovation

The unit of observation is the startup (i) by ICD category (j) by quarter (t). For each startup and quarter, the sample includes all ICD categories in which the startup has an active Phase I project at some point during the quarter. In column 1, the dependent variable $Initiation_{ijt}$ is an indicator for whether startup i initiates any new project in ICD category j during quarter t. In column 2, the dependent variable $RepurposeInto_{ijt}$ is an indicator for whether startup i repurposes one of its existing drugs into ICD category j during quarter t. In column 3, the dependent variable $RepursposeAway_{ijt}$ is an indicator for whether startup i repurposes a drug project from ICD category j into a different ICD category during quarter t. In column 4, the dependent variable $Alliance_{ijt}$ is an indicator for whether startup i forms an alliance with other pharmaceutical companies in ICD area j during quarter t. The independent variable of interest is $Treated \times NonSharedICD_{ijt}$, an indicator for whether (1) startup i has a project X in a *different* ICD category $j' \neq j$ that has experienced $Shocked = 1$, meaning a different project (from a different firm) in ICD category j' progresses to Phase II between the project X's initiation and quarter $t-1$; and (2) startup i and the startup that caused $Shocked = 1$ share a VC investor in common. $ShockedICD_{jt}$ is an indicator for whether another project (from a different startup) in ICD category j progressed to Phase II between startup i's earliest initiation in ICD j and $t-1$. $SelfICDProgress_{ijt}$ is an indicator for whether startup i has a project in ICD j progressing to Phase II in quarter t. $NICDProjects_{ijt}$ is the number of active projects that startup i has in ICD category j in quarter t. $\ln(ICDAge_{ijt})$ is defined as the logarithm of one plus quarters from startup i's earliest initiation in ICD j to t . All models are estimated by OLS, and standard errors are clustered at the startup level. FEs are noted in the bottom panel. ***, **, and * indicate statistics significance at the 1%, 5%, and 10% levels, respectively.

	(1) *Initiation*	(2) *RepurposeInto*	(3) *RepurposeAway*	(4) *Alliance*
$Treated \times NonSharedICD$	0.034**	0.026***	-0.036***	0.009*
	(2.19)	(3.10)	(-3.79)	(1.73)
$ShockedICD$	0.016	0.017**	-0.019**	-0.001
	(1.28)	(2.21)	(-2.02)	(-0.27)
$SelfICDProgress$	-0.016	-0.084***	0.091***	-0.003
	(-0.37)	(-3.00)	(2.83)	(-0.48)
$NICDProjects$	0.003***	0.002***	-0.001***	-0.000
	(3.99)	(4.52)	(-4.14)	(-0.22)
$\ln(ICDAge)$	-0.122***	-0.080***	0.069***	-0.002
	(-8.46)	(-7.65)	(6.33)	(-0.82)
Startup-Qtr FE	Yes	Yes	Yes	Yes
ICD FE	Yes	Yes	Yes	Yes
N	5,374	5,374	5,374	5,374
Adj. R^2	0.313	0.209	0.437	0.119

Table 9: Innovation Efficiency

We estimate OLS regressions where the unit of observation is the ICD category. The dependent variable is our efficiency proxy, the number of approved drugs scaled by aggregate VC funding. *Common Ownership Rate* is the probability that a randomly chosen pair of startups within an ICD shares at least one VC in common. In column 2, we add following variable as controls: *Duration to Phase III* is the logarithm of one plus the average number of quarters between project initiation date and reaching Phase III, averaging across all projects within the ICD that reach Phase III. We measure the duration through Phase III rather than FDA approval since some ICDs never see a project reach approval. *Prob. Reach Phase III* is the probability that a project initiated within this ICD eventually reaches Phase III. *Num. VCs per Startup* is the average number of VC firms per startups active in the ICD. *VC Holding Duration* is the logarithm of one plus average number of quarters between a VC firm's first investment and exit from the startup, averaging across all startups within the ICD. In column 3, instead of *Common Ownership Rate*, we use *Common Ownership Dummy*, an indicator for whether there exists any pair of startups that share a common VC investor at the same time within the ICD, as the independent variable.***, **, and * indicate statistics significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)
Common Ownership Rate	0.031***	0.025**	
	(3.62)	(2.35)	
Duration to Phase III		-0.000	
		(-0.98)	
Prob. Reach Phase III		0.001	
		(0.37)	
Num. VCs per Startup		0.000	
		(0.38)	
VC Holding Duration		-0.000	
		(-0.14)	
Common Ownership Dummy			0.001***
			(3.40)
N	94	94	94
Adj. R^2	0.115	0.093	0.102

Table 10: Technological Similarity

This table reports how the results from regressions (1) and (3) vary depending on technological similarity between the lagging project and the pioneering project. Columns 1 and 2 report the project-level results. Columns 3 and 4 report the VC-funding results. The proxy for technological similarity is based on patent citations, which we collect from the Google Patents database. $Shocked \times SharedCite$ equals one if the pioneering drug project and the shocked drug project have at least one outgoing citation in common. $Shocked \times SharedVC \times SharedCite$ equals one if the shocked company's drug shares at least one citation in common with at least one progressing drug of a different company, and both those companies share a common VC. To provide a reference point, Columns 1 and 3 estimate the regressions without a triple-interaction term. Remaining details and variable definitions are the same as in Table 3 and 5.

	(1) *Progress*	(2) *Progress*	(3) *ExtendFunds*	(4) *ExtendFunds*
$Shocked \times SharedVC$	-0.020***	-0.017**	-0.021*	-0.009
	(-3.03)	(-2.50)	(-1.85)	(-0.89)
$Shocked \times SharedVC \times SharedCite$		-0.020*		-0.103***
		(-1.84)		(-3.60)
$Shocked$	-0.000	-0.000	0.069***	0.069***
	(-0.06)	(-0.08)	(6.79)	(6.77)
$Shocked \times SharedCite$	0.010	0.012	-0.003	0.001
	(1.00)	(1.12)	(-0.16)	(0.08)
$NProjects$	0.002**	0.002**	0.001	0.001
	(2.20)	(2.24)	(0.26)	(0.27)
$NVCs$	0.000	0.000	0.012***	0.012***
	(0.13)	(0.11)	(4.37)	(4.34)
$\ln(Age)$	0.009***	0.009***		
	(2.85)	(2.82)		
$NProjectsperICD$	-0.034**	-0.034**		
	(-2.58)	(-2.58)		
$SelfProgress$			0.019*	0.019*
			(1.70)	(1.67)
$NCats$			-0.003**	-0.004**
			(-2.16)	(-2.20)
$PortfolioSize$			0.003	0.003
			(1.32)	(1.25)
$Duration$			-0.121***	-0.122***
			(-19.31)	(-19.52)
$PrevRoundSize$			-0.009**	-0.009**
			(-2.46)	(-2.53)
VC Firm FE	No	No	Yes	Yes
Startup FE	Yes	Yes	Yes	Yes
Yr-Qtr. FE	Yes	Yes	Yes	Yes
ICD FE	Yes	Yes	No	No
N	12,469	12,469	32,537	32,537
Adj. R^2	0.073	0.073	0.179	0.180

Table 11: Comparing Narrow and Broad Drug Categories

This table reports how the results from regression (1) vary depending on the narrowness of ICD definition. $NarrowICD_{it}$ is a dummy equal to one if the number of drug projects within i's ICD in quarter t is below the medium number of drug projects of all ICD categories in that quarter. In columns 1 to 3, we vary the method to count projects to define $NarrowICD_{it}$. Specifically, column 1 corresponds to a definition of $NarrowICD_{it}$ based on the number of Phase I projects within an ICD in each quarter. The average number of projects in ICDs of $NarrowICD_{it} = 1$ is 10.2. Column 2 corresponds to a definition of $NarrowICD_{it}$ based on the number of Phase I and II projects within an ICD in each quarter. The average number of projects in ICDs of $NarrowICD_{it} = 1$ is 24. Column 3 corresponds to a definition of $NarrowICD_{it}$ based on the number of Phase I, II and III projects within a ICD in each quarter. The average number of projects in ICDs with $NarrowICD_{it} = 1$ is 29.3. Remaining details are the same as in Table 3.

	(1) *Progress*	(2) *Progress*	(3) *Progress*
$Shocked \times SharedVC$	-0.017**	-0.017**	-0.017**
	(-2.53)	(-2.57)	(-2.56)
$Shocked \times SharedVC \times NarrowICD$	-0.046**	-0.047**	-0.048**
	(-2.10)	(-2.07)	(-2.05)
$Shocked$	-0.001	-0.001	-0.001
	(-0.20)	(-0.18)	(-0.22)
$Shocked \times NarrowICD$	0.013	0.012	0.012
	(0.83)	(0.75)	(0.77)
$NarrowICD$	-0.007	-0.011	-0.023**
	(-0.75)	(-1.39)	(-2.60)
$\ln(Age)$	0.009***	0.009***	0.009***
	(2.91)	(2.89)	(2.87)
$NProjects$	0.002**	0.002**	0.002**
	(2.18)	(2.18)	(2.15)
$NVCs$	0.000	0.000	0.000
	(0.11)	(0.16)	(0.20)
$NProjectsperICD$	-0.036**	-0.037***	-0.041***
	(-2.58)	(-2.69)	(-2.92)
Startup FE	Yes	Yes	Yes
Yr-Qtr. FE	Yes	Yes	Yes
ICD FE	Yes	Yes	Yes
N	12,469	12,469	12,469
Adj. R^2	0.073	0.073	0.073

Table 12: Enforcement of Employee Non-Compete Agreements

This table examines whether our results differ across states with strong versus weak enforcement of employee non-compete agreements. Column 1 repeats our OLS regression of project outcomes, from column 1 of Table 3, except we include interactions terms $Shocked \times SharedVC \times Noncompete$ and $Shocked \times Noncompete$. $Noncompete_{it}$ is an indicator for whether startup i's state in year t strongly enforces non-compete agreements. We obtains states' current non-compete statutes from Beck Reed Riden LLP (`http://www.beckreedriden.com/50-state-noncompete-chart-2/`). We record states' historical statute changes following Ewens and Marx (2018) and Jeffers (2019). We then use the U.S. Department of Treasury's report (`https://www.treasury.gov/resource-center/economic-policy/Documents/UST%20Non-competes%20Report.pdf`) to classify the enforceability of non-compete agreements into 5 levels: not enforced, undecided, red pencil, blue pencil, and reformation. We classify blue pencil and reformation as having a high degree of enforceability. Column 2 repeats our OLS regression of VC funding outcomes, from column 7 of Table 5, with a similar triple-interaction term. Standard errors are computed by two-way clustering at the ICD category and startup company levels. FEs are noted in the bottom row. ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) *Progress*	(2) *ExtendFunds*
$Shocked \times SharedVC$	-0.015**	-0.021**
	(-2.15)	(-2.04)
$Shocked \times SharedVC \times Noncompete$	-0.002	-0.008
	(-0.28)	(-0.48)
$Shocked$	0.001	0.265***
	(0.11)	(7.26)
$Shocked \times Noncompete$	0.000	0.013
	(0.06)	(0.30)
$Noncompete$	0.008	
	(1.01)	
$\ln(Age)$	0.010***	
	(2.98)	
$NProjects$	0.002**	
	(2.23)	
$NVCs$	0.000	
	(0.14)	
$SelfProgress$		0.091***
		(3.61)
$NCats$		-0.003**
		(-2.13)
$PortfolioSize$		0.002
		(1.36)
$Duration$		-0.172***
		(-18.96)
Startup FE	Yes	No
Startup× Qtr. FE	No	Yes
VC Firm FE	No	Yes
Yr-Qtr FE	Yes	No
ICD FE	Yes	No
N	12,469	34,414
Adj. R^2	0.071	0.465