Excess Commitment in R&D*

Marius Guenzel

Tong Liu

January 2024

Abstract

We document a form of "excess" commitment to R&D projects and examine the consequences for innovation outcomes and consumer welfare, using detailed data on pharmaceutical firms' clinical trial projects. Plausibly-exogenous delays in the completion of the preceding trial-phase, empirically uncorrelated with various project-quality measures, substantially reduce firms' subsequent project suspension propensity. This excess project commitment intensifies when the CEO has higher stock-price–compensation sensitivity and is personally responsible for the project's initiation. While delay-driven project commitment induces investment crowd-out, it does not predict increased adverse effects in marginally-launched drugs and predicts continuation of drugs targeting diseases lacking existing treatments (orphan diseases).

Keywords: R&D, (excess) commitment, consumer outcomes, welfare, healthcare

*Guenzel is at the Wharton School, University of Pennsylvania. Liu is at the MIT Sloan School of Management. Emails: **mguenzel@wharton.upenn.edu**, **tongl@mit.edu**. We are grateful for comments from Tania Babina (discussant), Amber Batata, Alice Bonaime (discussant), Hui Chen, Taha Choukhmane, Luca Fare (discussant), Johan Hombert, Josh Krieger (discussant), Ashley Litwin, Andrew Lo, Deborah Lucas, S. Katie Moon (discussant), David Robinson (discussant), Antoinette Schoar, Jerome Taillard, Richard Thakor (discussant), and David Thesmar, as well as seminar and conference participants at Binghamton University, Maastricht University, MIT Sloan, Peking University, Nanjing University, Rotterdam University, Wharton, Virtual Corporate Finance Seminar, 2023 Kentucky Finance Conference, 2023 LBS Summer Finance Symposium, 2023 ENTFIN Conference, 2023 EFA Meeting, 2023 NFA Meeting, 2023 HEC Paris Entrepreneurship Workshop, and 2024 AFA. We thank Yubo Wei and Tim Zhang for excellent research assistance. All errors and omissions are our own.

1 Introduction

Research and development (R&D) is a primary driving force behind innovation. One defining feature of R&D is the frequent need for long-term commitment to a given course of action to achieve scientific breakthroughs. For example, in the realm of new drug development—the focus of this paper—it typically takes 10 to 15 years for a new drug to reach the market, involving extensive development and testing (Wong et al., 2014). This may suggest a positive influence of commitment—i.e., the sustained effort and resources dedicated to maintaining a particular R&D strategy over time—on fostering innovation in general.[1]

However, seminal prior work, in particular Staw (1976)'s evidence from the lab, finds that decision-makers frequently remain committed to R&D projects they have previously allocated resources to even when confronted with costly setbacks and the availability of superior alternatives. Such behavior, which one may succinctly label *excess* commitment to ongoing R&D activities, can arise from market frictions, informational frictions, as well as more psychological frictions, and can have profound consequences for both firms and consumers. For one, the high-effort, high-cost nature of R&D implies constraints on the number of projects a firm can pursue simultaneously. Additionally, excess commitment can affect not only which drugs firms develop and progress, but may also induce firms to pursue inferior or even less safe innovation relative to a frictionless counterfactual.

In this paper, we provide a detailed analysis of how firms' commitment to ongoing R&D projects varies with the amount of resources the firm has already expended on the project, and the extent to which firms' R&D commitment may be characterized as excessive. Specifically, we study how firms' propensity to continue versus abandon an existing R&D project is affected by incurred project delays that are unanticipated, plausibly-exogenous, and costly. Our setting enables us to investigate in detail the resulting effects on a wide range of firm and consumer R&D outcomes, such as investment crowd-out and

[1] We note that our notion of commitment, encompassing the (potentially time-varying) inclination of firms to maintain a particular R&D strategy, is slightly different from the term's usage in the contracting literature. In contracting, commitment typically pertains to the ex-ante promise and adherence to contracts in multi-period principal-agent frameworks (e.g., Rey and Salanie, 1990). Aligned with our notion of commitment, R&D firms such as Merck and Pfizer frequently emphasize, for example, their "long-term commitment to research and development strategies," along with their endeavors in specific R&D areas such as a "long-term commitment to oncology" (https://www.sec.gov/Archives/edgar/data/310158/000130817918000164/lmrk2018_def14a.htm; https://www.pfizer.com/news/press-release/press-release-detail/pfizer_to_showcase_diverse_and_growing_oncology_portfolio_at_the_american_society_of_clinical_oncology_asco_2016_annual_meeting).

innovation novelty and quality.

Studying how firms' decision-making related to ongoing R&D activities depends on prior temporal or financial investments made, and may involve excess commitment, is challenging, for at least two key reasons. First, it requires granular, project-level data on R&D activities, including observable information on R&D milestone achievements and final R&D outcomes. Second, it requires not only reliable measures capturing the degree of commitment to R&D projects over their life, but also a setting in which there is plausibly exogenous variation in the intensity of initial project-specific temporal or financial investments made by firms.

Focusing on the important setting of new drug development and clinical trials by pharmaceutical firms allows us to overcome these twin challenges related to data and identification. We assemble a new dataset of clinical trials initiated by U.S.-based companies, comprised of more than 10,000 clinical trials initiated between 1985 and 2019. The underlying data come from a variety of sources, including Cortellis Clinical Trial Intelligence, web-scraped and hand-matched data from ClinicalTrials.gov, and the FDA Adverse Event Reporting System database.

Crucial for identification, we observe in our data both anticipated timelines of clinical trials and each trial phase, as filed by firms on ClinicalTrials.gov, as well as realized timelines. This allows us to isolate unanticipated delays incurred by firms in already-completed trial phases, i.e., variation in unanticipated temporal investments made in the project. Further, our setting equips us with observable project decisions that directly capture firms' project commitment (in particular, whether they advance or terminate a given clinical trial at given trial stage), as well as detailed observable and quantifiable R&D outcomes (such as patient outcomes in ultimately approved drugs).

To guide our empirical analysis, we develop a simple conceptual framework that allows us to establish a benchmark for the relation between unexpected project delays and continuation versus suspension, and to specify our notion of excess commitment. We consider a manager who is in charge of a multi-phase project and who can be of high or low ability ("good" vs. "bad"). A bad manager has a lower instantaneous probability of project phase completion. Both manager types can choose to suspend the project and start employment with a new employer (managers' outside option). This basic setup, paired with full information about managerial ability, yields that delay in completing the first project phase will not affect

the manager's probability to suspend the project subsequently. This result motivates us to label any positive relation between project delay and continuation probability as *excess* commitment, also reflecting that the majority of projects in our data experience a positive (i.e., > 0) delay (see Table 1).

We organize our empirical analysis into four main parts. In the first part, we establish a new baseline fact: a strong positive link between unanticipated delays that firms experience in completing the preceding clinical trial phase and their subsequent commitment to the clinical trial project, i.e., their propensity to subsequently continue the project through the next trial phase rather than terminate it. We document, leveraging the information on anticipated trial end dates, that the average trial phase in our sample is completed with a delay of nearly one year, with the 25th (75th) percentile trial phase being completed with no delay (a delay of 1.6 years). A one standard deviation increase in completion delay of the preceding trial phase is associated with a large reduction in the subsequent propensity to terminate the project of 4 percentage points, or 15% relative to the baseline suspension probability of 31%.

In the second part, we relate this baseline finding to our notion of excess commitment. A key question is whether unanticipated project delays might produce information that increases the value of projects and therefore should lead firms to increase their commitment and propensity to continue delayed projects. At odds with this, experienced delay is empirically uncorrelated with various independent project quality measures obtained from GlobalData as well as the the extent of drug dosage experimentation (i.e., learning-by-doing) in the preceding phase.

To probe even deeper, we implement an instrumental variables (IV) approach, instrumenting for unanticipated trial phase completion delay with a measure capturing *clinical trial site congestion*. The IV strategy exploits that trial sites have limited capacity to accommodate clinical trials, with increases in the number of trials taking place simultaneously in a given location increasing the likelihood of bottlenecks. We construct the congestion measure to capture *changes* in site busyness relative to the trial start, to account for the fact that firms do not select trials at random, and find a strong positive relation between trial site congestion increase and trial phase completion delay, i.e., a strong first stage.

The identifying assumption of the IV approach is that changes in trial site congestion affect firms' decision to continue developing new drug candidates only through the effect on trial phase completion delay. We argue this exclusion restriction may plausibly hold since the congestion measure is empirically

uncorrelated with a broad array of potentially confounding factors. These factors include, as above, drug quality and experimentation measures, as well as detailed trial site quality measures assessing hospital care quality in a given region. In the second stage, congestion-induced delay strongly predicts project continuation, similar in magnitude to the OLS estimates, reinforcing the excess commitment interpretation of the findings, i.e., that the delays and associated increased temporal (and financial[2]) project investments themselves induce firms to continue projects due to underlying frictions.

In the third part, we provide evidence on underlying mechanisms, which further substantiates the excess commitment interpretation. Our findings are not driven by firms lacking alternative drugs (Guedj and Scharfstein, 2004) or being unable to pivot due to financial constraints. Differential sorting by firms into trials with different lengths or delays based on initial expectations is equally unlikely to be an underlying mechanism, since our data allow us to control for trial length expectations. Moreover, the documented empirical patterns are unlikely to reflect a real options channel, since drug delays are not correlated with the expected drug sales upon drug approval and launch.

Instead, the effect of unexpected trial delays on subsequent trial commitment is amplified in the presence of management-related frictions. Delay is substantially more predictive of continuation among chief executive officers (CEOs) with greater wealth exposure to stock price changes (delta), consistent with these CEOs being particularly concerned that investors will interpret the suspension of a delayed drug project as a negative signal (about the firm or the CEO). Additionally, delay is significantly more predictive of project continuation when the CEO heading the firm remains the same between trial phase start and end, consistent with personal responsibility for initiating the project, and project phase, playing an important mediating role.

In the final part, we evaluate the implications of our findings for firm investment and consumer outcomes. Whereas unanticipated trial phase completion delay increases the firm's subsequent commitment to that specific project, we find that it crowds out the initiation of new drug development projects, in particular projects targeting other diseases. In the long run, this may negatively affect drug choices available to consumers. Regarding consumer outcomes observable in the more near-term, we find that

[2]Drug development delays are inherently linked to increases in project costs (Wong et al., 2014). For example, longer timelines generate higher labor costs, as investigators and staff must be compensated for longer, as well as increased patient monitoring and care costs.

marginally-approved drugs due to delay-induced drug project commitment are associated with economically modest but statistically insignificant increases in harmful adverse events in patients. At the same time, we find that delay-induced increased commitment to trials increases the probability that firms ultimately bring the drug to market among the important subgroup of drugs designated as orphan drugs, i.e., drugs targeting diseases with a lack of existing treatment options.[3] Taken together, these findings suggest a nuanced relation between firms' excess commitment and welfare implications for consumers.

Broadly, our findings complement a large literature and active debate about the extent to which there is adequate innovation activity in the economy. Seminal prior work has, for instance, uncovered many reasons for why there might be aggregate under-investment in innovation, including low spillover effects, low competition, patent protection, infringement lawsuits, and risk aversion (Hall and Lerner, 2010). Our contribution to this debate is unique in that we examine the commitment with respect to *existing, in-process* R&D endeavors, rather than assessing the adequacy of R&D adoption at the extensive margin.

Our paper makes three main contributions to the literature. First, we contribute to the literature on the economics of innovation. Moser (2016) and Bryan and Williams (2021) provide comprehensive surveys of recent research on the market failures in the markets for innovation and intellectual property. In the context of pharmaceutical innovations, Acemoglu and Linn (2004) investigate the effect of potential market size on drug innovations. Budish et al. (2015) show that short-termism and the fixed patent term distort private research investments. Azoulay et al. (2019) study the impact of scientific grant funding at the NIH on pharmaceutical innovations. Krieger et al. (2022a) study how risk aversion leads pharmaceutical firms to underinvest in radical innovation. Thakor and Lo (2022) investigate how product market competition affects R&D investment and financing. Bonaime and Wang (2023) shed light on how mergers affect pharmaceutical innovation. Li et al. (2023) show that common ownership among venture capital firms could be beneficial to innovation efficiency.[4] Departing from the previous work, this paper studies how firms' (excess) commitments to the R&D process, an organizational response within firms,

[3] See, e.g., https://www.fda.gov/news-events/fda-voices/rare-disease-day-2021-fda-shows-sustained-support-rare-disease-product-development-during-public.

[4] With respect to biomedical innovation, Lo and Thakor (2022) and Lo and Thakor (2023) review the recent literature on how external financing frictions affect drug development, including Lerner et al. (2003) on equity financing cycles, Robinson and Stuart (2007) on financial contracting in strategic alliances, Lerner and Malmendier (2010) on contractibility in research agreements, Cunningham et al. (2021) on the M&A market, Aghamolla and Thakor (2022) and Liu (2021) on the IPO decisions of biotech, and Krieger et al. (2022b) on profit erosion of existing products.

affect innovation efforts and outcomes. In particular, we discuss the potential welfare consequences, and challenge the conventional wisdom that friction-induced R&D decisions by firms are unequivocally bad for consumers.

Second, we contribute to the literature in economics and finance on the effects of top management on firm investment. Guenzel (2023) documents distortions in firm investment due to managerial sunk cost effects (which can be thought of as one manifestation of "path-dependent decision-making," similar to excess commitment) in the context of mergers and acquisitions, with exogenously more expensive acquired businesses being less likely to be abandoned through divestiture. Related to aggregate efficiency and welfare, Barrero (2022) and Ma et al. (2020) study the macro and equilibrium effects of managerial overconfidence and diagnostic expectations, but do not consider, e.g., consumer-relevant effects on product quality or variety as a result of distorted firm investment. Compared to this prior work, we focus on an entirely different investment context, R&D, extend the scope of distortive factors by studying the effect of project delays, and quantify various welfare-relevant consumer outcomes. Despite its importance, studying consumer and welfare implications of management-related firm decisions is still mostly absent in the literature on nonstandard firm decision-making (see Malmendier (2018) and Guenzel and Malmendier (2020) for recent literature surveys).

Third, we contribute to a literature in economics and related fields on the effects of commitment and escalation of commitment. As alluded to above, in an influential laboratory study on commitment and path dependence, Staw (1976) finds that subjects are more committed to a chosen (hypothetical) R&D project if they were personally responsible for the initial action and, consistent with an escalation mechanism, in response to negative signals about the investment. Subsequent work has provided empirical evidence consistent with escalating commitment in various contexts including "problem loan" write-offs (Staw et al., 1997) and NBA draft picks (Staw and Hoang, 1995; Camerer and Weber, 1999). A common problem in empirical studies is the lack of plausibly-exogenous variation in initial investments made, which opens the door for alternative explanations related to information, beliefs, or selection to explain subsequent commitment. We advance this literature by deliberately focusing on how *unanticipated* factors, project delays, affect subsequent commitment.[5] We also advance this literature by studying not solely the

[5]We note two papers that use price variation induced by auctions to study commitment in *consumer*, rather than firm, settings (Augenblick, 2016 in consumers' penny auction behavior; Ho et al., 2018 in consumers' car usage behavior). Guenzel (2023)

immediate effects of commitment on the decisions and decision-makers in focus, but also the ensuing consequences for related decisions and agents (i.e., investment crowd-out and effects on consumers).

The remainder of the paper is structured as follows. Section 2 introduces a simple theoretical framework that discusses the notion of excess commitment. Section 3 provides details on institutional background and data. Section 4 introduces the empirical approach. Sections 5 to 7 discuss the results and welfare implications. Section 8 concludes.

2 Theoretical Framework

This section introduces a simple theoretical framework to derive a benchmark for the relation between project delays and the decision to continue versus suspend it. We sketch the framework here, and present the full version of the model in Section I of the Online Appendix.

There are two types of managers: high ability ("good") and low ability ("bad") managers. Each firm hires a manager to implement one project (which differs in their profitability across firms), and managers are randomly matched with firms. Projects need to complete two phases (Phases I and II) before profits are realized. Compared to a good manager, a bad manager has a lower instantaneous probability to complete each project phase. After completing Phase I, both managers can choose to suspend the project and start employment with a new employer, which corresponds to their outside option. Managers' compensation and outside option are increasing functions of their (perceived) ability by the market.

This simple setup suffices to characterize our notion of excess commitment. Specifically, in the base case of full information about managerial ability, the framework predicts that delay in completing Phase I will be *unrelated* to the manager's probability to suspend the project subsequently.

Prediction: *In the benchmark model with full information about managers' ability, unexpected delays in the already-completed trial stage do not affect the manager's propensity to suspend the project subsequently.*

This result motivates our terminology of referring to the positive relation between project delay and

exploits post-acquisition-agreement aggregate stock market fluctuations to obtain quasi-random variation in sunk acquisition costs. We further note the existence of a literature on R&D focusing on how commitments by some firms can lead to "path dependence" in a *different* sense, namely creating barriers to operate for *other* firms (e.g., Manez et al., 2009). In contrast to such studies, we focus on *within-firm* commitment effects.

continuation probability observed empirically as *excess* commitment, also considering that the majority of projects in our data experience a positive (i.e., > 0) delay (see Table 1).

A positive relation between project delay and continuation is only predicted if we enrich the framework with additional frictions. As we derive in Section I of the Online Appendix, one plausible such friction is information asymmetries about managerial ability. Our empirical evidence in Section 5.4 is consistent with the predictions of this extended model.

Our framework abstracts from other underlying frictions, such as more psychological frictions (e.g., those grounded in cognitive dissonane avoidance (Festinger, 1962); see Section 5.4.2 for a more detailed discussion), which could also generate a positive relation between experienced delay and subsequent project continuation.

3 Data and Institutional Background

3.1 Institutional Background

The pharmaceutical industry is highly R&D intensive. In 2019, the whole industry spent $83 billion dollars on R&D, and its R&D intensity, defined by R&D spending as a share of net revenues, reached 25 percent, which is higher than the software and semiconductor industries (Austin and Hayford, 2021). Among a variety of activities, a large portion of the R&D spending in the pharmaceutical industry is devoted to conducting clinical trials.

Clinical trials are regulated by the FDA and typically involve three stages: Phase I, II, and III trials. Phase I tests the safety of a drug candidate on a small group of humans and determines how and where it distributes within the body. Phase II involves a larger number of patients, aiming to determine the dosage and effectiveness of the drug. Phase III primarily focuses on the safety and efficacy for a wide variety of populations by testing the drug candidate on a much larger group of patients. Obtaining favorable trial results in a given trial phase is required to be allowed to move to the next phase. After completing all three phases, a drug developer (sponsor) can submit the new drug application to the FDA to seek the market launch permission. For each stage of trials, a sponsor needs to design a strict protocol—a scientific plan of action. The protocol explains how many patients will be recruited, the inclusion/exclusion criteria for

eligible patients, which tests will be performed, how results will be measured and collected, and why the trial may be discontinued. Clinical trials are often conducted in different settings and locations. These locations are across cities, states, and countries.

Clinical trials are costly. Recent studies reveal that it costs over $1 billion to bring a new drug to the market (DiMasi et al., 2003; DiMasi et al., 2016; Wouters et al., 2020). A variety of factors that significantly affect the final cost of a clinical trial include clinical procedure costs, administrative staff costs, site monitoring costs, site retention costs, and central laboratory costs. The total costs for a trial can be grouped into per-patient costs and per-site costs. Per-patient costs contain the costs for patient retention, payments to nurses and physicians, lab costs, clinical procedure costs, etc. Per-site costs are approximately the sum of site recruitment costs, site retention costs (per months), administrative staff costs (per month), and site monitoring costs (per day) (Wong et al., 2014).

Notably, these cost factors heavily hinge on the length of time to complete a trial. Longer trial timelines increase development costs (Wong et al., 2014). For example, longer timelines mean higher labor costs as investigators and staff must be compensated for longer hours. Longer timelines also imply more devoted hours to take care of trial participants, which leads to higher costs to drug developers. A recent study by Shadbolt et al. (2023) reveals that extensive delays in surgical clinical trials are associated with inefficiency, and the economic impact of delays to trial timelines is undoubtedly substantial.

In terms of the typical length of clinical trials, Wong et al. (2019) document that the median duration of Phase I, II, and III are 1.6, 2.9, and 3.8 years, respectively.[6] In our sample, whose construction we describe next, the median trial phase (comprised of Phase I and II trials) similarly takes 2.4 years to complete (Table 1; see Section 3.3 for further summary statistics).

3.2 Data and Sample Construction

Clinical Trials Data. We construct a sample of clinical trials initiated by U.S.-based companies. The trials data come from Cortellis Clinical Trial Intelligence, and allow us to observe detailed trial

[6]However, the time to complete a trial varies widely across indications. For example, the median duration of Phase III oncology trials is about six years. Table OA.1 of the Online Appendix examines the determinants of a trial's anticipated duration in our sample by gradually adding variables of clinical trial characteristics and comparing the change in adjusted R-squared. The analysis suggests that the disease field that a drug project targets, the time when a trial is initiated, and the identity of drug developing company are three important factors affecting the expected trial duration. We account for all these variables in our empirical estimation below.

information, including the trial title and phase, start and completion dates, number of sites and site countries, trial protocol, eligibility criteria, interventions, adverse outcome (if applicable), and whether the endpoints are achieved; information on drug candidates, drug indications, drug dosages and associated technology; and information on the drug developer ID and name, its role in the trial, funder type, and whether the trial is commercial. Our analysis focuses on the trial by drug indication level (hereafter referred to as a project). To obtain the outcome of trials and corresponding drug projects, we merge the trials data with drug development data from Cortellis Competitive Intelligence. We classify a project as suspended at a certain trial phase if it is explicitly coded as suspended, discontinued, withdrawn, or coded as "no development" for over three years, or coded as staying in Phase I, II, or III for over five, eight, or ten years, respectively, in the drug development data (Li et al., 2023).

Data Filters. We apply several filters to construct our final analysis sample. First, we focus on projects with already-completed Phase I or II trials and non-missing records on their outcomes after these trials, i.e., suspension or continuation. Second, we drop any projects whose Phase I or II trials were completed after 2020, because these projects would have insufficient time to reach any outcomes. Third, since our key empirical test is to ask whether firms advance their projects, which necessitates that they can do so, we drop any projects in Phase I or II trials which did not achieve their endpoints. This ensures that suspension, if and when it occurs, is not due to unfavorable readouts in the previous trial phase. Next, we require a project to have a realized trial start and completion date and hence drop those with estimated completion dates. These filters yield a sample of 11,228 projects initiated between 1985 and 2019 (completed between 1991 and 2020), with trial sites spanning across 91 countries including the U.S.

Match With ClinicalTrials.gov. We supplement our trials data with a variety of other data sources. To obtain detailed address information for each trial site (street, city, state, country, and zip code), we manually match our sample with trial data scraped from ClinicalTrials.gov maintained by U.S. National Library of Medicine. For each trial, ClinicalTrials.gov contains the initial submission as well as the updated trial reports submitted by the drug developer afterwards. We match our sample with all report versions, which also allows us to construct one of the key variables in our analysis, the length of trial delay, as the difference of the anticipated trial completion date in the drug developer's initial submission to ClinicalTrials.gov and the realized trial completion date. Details of the sample matching procedure are

provided in Section III.A in the Online Appendix.

Adverse Events Data. To shed light on welfare implications, we manually link the approved drugs in our trials sample to the FDA Adverse Event Reporting System database (2012-2022) and its predecessor, the Legacy Adverse Event Reporting System (2004-2012). The adverse event database records adverse drug reactions in the U.S. reported to the FDA. We focus on events in which a drug is listed as the "primary suspect" (Cohen et al., 2021). Online Appendix Section III.B describes the corresponding matching procedure.

Other Data. Lastly, we obtain financial accounting and management information for a subset of the drug developers that are public pharmaceutical companies traded on U.S. exchanges from Compustat, BoardEx, and Execucomp. We obtain exchange rate data from Bloomberg. We obtain expected revenues of drugs once launched in the market from Cortellis Competitive Intelligence. We acquire drug quality proxies, including the drug-specific likelihood of approval (LOA) and the drug's phase transition success rate (PTSR) from the GlobalData Pharma database. We collect data on the quality of U.S. hospitals, including 30-day mortality rates, 30-day readmission rates, and patient safety indicators, from various CMS-managed quality programs. We collect clinical hold records from the ClinicalTrials.gov.

3.3 Summary Statistics

Table 1 presents summary statistics for our main sample comprised of completed Phase I or II clinical trial projects. About one in three projects is suspended, i.e., not advanced through the next phase (Phase II or III), respectively. The average trial takes approximately three years to complete. (We discuss further variables related to trial duration, as well as those related to exchange rates below in Sections 4.1 to 4.3, when we introduce the empirical design.) The average trial in our data is conducted at 14 different sites, though there is significant variation with respect to number of trial sites, and the average trial has 100 participants. Just under 60% of trials in our sample are Phase II trials, the remainder are Phase I trials. The average firm in our sample has 136 drug projects (drug by ICD indications[7]) under development, i.e., our sample firms are *not* single-product firms as the firms analyzed in Guedj and Scharfstein (2004). We also note that the number of observations in the regressions will sometimes differ from the summary

[7]Throughout out analysis, we use the most granular, level-three categories of ICD classification codes. We sometimes refer to them as ICD categories for conciseness.

statistics numbers in Table 1, depending on the availability of the duration variables.

Figure 1 shows the geographic distribution of clinical trial sites in our sample. Clinical trials have gone global in the recent decades, with trial sites being widely spread across countries. Though a large share of clinical trials are clustered in the U.S, accounting for 63% of the total trial sites (59,617 out of 94,261 clinical trial sites), other important countries including Canada, Germany, France and Japan contribute another 12% of the total trial sites. This pattern is robust to tabulating the number of trials, rather than trial sites, across countries (Figure OA.1 of the Online Appendix). In terms of number of trials, the U.S. accounts for 45% of observations (6,847 out of 15,222 clinical trials). Canada, Germany, France, and Japan contribute 16%.

4 Empirical Approach

Our empirical approach centers on examining pharmaceutical firms' decision to continue versus suspend drug projects. Specifically, we test whether plausibly-exogenous delays in clinical trial phase completion induce firms to be more likely to "stay committed" to a project and advance the clinical trial through the next clinical trial phase. Figure 2 visualizes the key elements of the project timeline and empirical design.

4.1 Delay in Trial Phase Completion and Project Continuation Versus Suspension

To examine how firms factor unexpected delays in clinical trial phase completion into their subsequent R&D decision-making regarding continuation of the trial through the next trial phase, we estimate:

$$\text{Suspension}_i = \beta \text{DelayInTrialCompletion}_i + \gamma \text{Controls}_i + \text{FEs} + \eta_i. \tag{1}$$

We estimate Equation (1) on the sample of all projects i that have completed clinical trial Phase I or II with detailed clinical trials records. Suspension_i is an indicator variable for suspension of project i, i.e., equals one if a completed Phase-I (Phase-II) trial is subsequently suspended and does not advance through Phase II (Phase III). The independent variable of interest, $\text{DelayInTrialCompletion}_i$, is the difference between actual trial completion date and the anticipated completion date as indicated by the firm at the

start of the trial. In our sample, the average trial completion is nearly one year later relative to the initially anticipated end date (Table 1). There is also substantial heterogeneity in the degree of delay experienced, which is useful for identification. The 25th-percentile trial phase is not completed with delay, whereas the 75th-percentile trial phase experiences a delay of approximately one and a half years.

It is important to emphasize that there exist clear incentives and institutional guidance for firms to truthfully report anticipated trial end dates. For one, firms must update the anticipated completion date on ClinicalTrials.gov whenever it changes, implying additional compliance costs if firms misreported expected trial end dates at the outset. Additionally, reporting guidelines for ClinicalTrials.gov explicitly stress that "it is important these dates are accurate."[8] We also note that our IV strategy (discussed next in Section 4.2) based on trial site congestion, and specifically changes in crowdedness over the course of the trial, is further suited to address possible concerns around any remaining strategic reporting of anticipated trial end dates by firms, as long as such remaining incentives are uncorrelated with the congestion instrument.

In the regressions corresponding to Equation (1), we will employ various levels of fixed effects (FEs) to control for unobserved heterogeneity, including FEs for the year-by-quarter of trial phase start date, for the drug company, for the drug's ICD category, and completed trial phase (Phase I or II). We will also include a vector of control variables, the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). We use two-way clustered standard errors clustered at the ICD category and quarter levels.

4.2 Increase in Trial Site Congestion as an Instrument for Delay in Trial Completion

Our *Delay in Trial Completion* measure captures the unanticipated component of the length of a given clinical trial, as the difference between realized end date and anticipated end date filed by the firm with ClinicalTrials.gov at trial start. To further isolate the effects of plausibly-exogenous variation in trial length on firm continuation-versus-suspension decision-making, and assess the degree of excess commitment in R&D decision-making, we introduce an instrumental variables (IV) strategy that uses *trial*

[8] See, e.g., https://www.dfhcc.harvard.edu/crs-resources/ODQ_Documents/02_CT.GOV_CTRP/DFHCC_ClinicalTrialsGov_Results_Reporting_Training.pdf.

site congestion as an instrument for trial completion delay. More precisely, we focus on the *change* in trial site crowdedness over the course of the trial, to account for the fact that trial sites (and therefore initial trial busyness) are not chosen randomly by firms. This instrument is based on the notion that trial sites have limited capacity to accommodate clinical trials. If a trial site hosts too many trials, it is challenging to recruit patients on time and monitor all trials simultaneously, which increases the likelihood of a trial slowdown.

To construct our instrument, we download the universe of trial records across all phases (all Phase I, II, and III trials) from ClinicalTrials.gov and standardize the zip codes of the 20 most frequent countries where these trials are conducted. The trials in the top 20 countries cover about 90% of total trials at the ClinicialTrials.gov. We define each unique zip code in a country as one trial location, dropping any trials with missing addresses, number of participants, and start and completion dates. In a first step, we then construct a congestion measure at the zip code level as follows. For a trial i conducted in N_i sites with a start year τ_1 and a completion year τ_2, we compute the average patient enrollments per year per site E_i as trial i's total number of enrolled patients divided by the number of sites N_i and the number of years to complete the trial, $(\tau_2 - \tau_1)$. The the congestion measure of year t for location z, denoted by G_{zt}, is the sum of E_i for all trials that have a site located in z and year t lying between the trial start year τ_1 and the trial completion year τ_2:

$$G_{zt} = \sum_{\{i \in I : N_i \cap z = z,\, t \in [\tau_1, \tau_2]\}} E_i$$

where I is the full set of all clinical trials. The key intuition is that a larger G_{zt} implies that location z hosts more trial participants in year t. In the second step, we calculate the *change* in congestion measure at location z between start year τ_1 and completion year τ_2 for trial i. We scale this change by the mean level of the congestion measure of location z across years, $\overline{G}_z$, so we interpret it as the relative change in congestion at location z. Our instrument for delay is then

$$z_i = Congestion_i = \max_{z \in N_i} \{ \frac{G_{z\tau_2} - G_{z\tau_1}}{\overline{G}_z} \}.$$

The instrument takes the maximum across trial i's sites because a clinical trial is not completed unless the last trial site is closed. For the average trial in the sample, the increase in trial congestion between trial

start and end, normalized by the mean congestion, is slightly more than five percent (Table 1).

The key identifying assumption is that the instrument affects the outcome variable $Suspension_i$ only through its effect on DelayInTrialCompletion$_i$. The IV test boils down to examining whether firms are more likely to advance the trial through the next phase if the trial site of the already-completed previous-stage trial became more crowded. Why does the exclusion condition plausibly hold? We argue that the change in trial site crowdedness of a zip code (macro conditions) is unlikely to be correlated with unobserved shocks to firms' decisions to advance a specific clinical trial (firm-level condition). A possible concern is that a trial site with more participants may imply better service quality, such as improved site monitoring, data collection, and implementation of clinical protocols that it could potentially offer during trials. Clinical trials conducted in these crowded sites, though with a delayed schedule, would then tend to have trial results with higher quality, and firms would be more likely to advance the trials through the next phase. We find no evidence that the trial site congestion measure G_{zt} is (positively or negatively) correlated with trial site quality, proxied by the average clinical score/evaluation across hospitals in a zip code. Figure 3 shows the correlation coefficients of 26 hospital quality measures devised by the CMS between 2005 to 2017. We also find that the instrument is empirically uncorrelated with various project quality and experimentation measures, which further helps alleviate concerns related to the exclusion restriction. We discuss these additional (non-)correlations in Section 6.

4.3 Exchange Rate Variation in Foreign Trials

As a validation exercise, we will also employ a second measure of plausibly-exogenous prior investments made to a clinical trial that is independent of the time-based delay and congestion measures, and instead captures variation in financial investments made. Specifically, we leverage the institutional feature described above that a large portion of clinical trials are conducted outside of the U.S. For a trial involving multiple sites internationally, the contracts and actual payments are conducted in other currencies rather than the U.S. dollar. Consequently, when foreign currencies become more expensive relative to the U.S. dollar, this yields unexpected costs to drug developers due to the exchange rate volatility.[9]

[9] See, e.g., https://www.appliedclinicaltrialsonline.com/view/mastering-currency-fluctuation, highlighting that a "study's financial obligations may include absorbing fluctuations in exchange rates to meet the established contractual requirements" and that payments occur "typically [in] the currencies of the countries where the trials are held."

To measure the extent to which foreign currency becomes more expensive (or cheaper) across the whole duration of a clinical trial, we take a moving-window average of changes in exchange rates. Specifically, we define $\Delta FX^c_{[\tau]}$ as the percentage change in the exchange rate (unit of dollars in exchange for one unit of foreign currency in country c) τ years after the clinical trial was initiated. Then we construct an average exchange rate change for clinical trial i as

$$\Delta FX_i = \sum_c \frac{n_{ic}}{N_i} \sum_{\tau=1}^{T} \Delta FX^c_{[\tau]} \tag{2}$$

where n_{ic} is the number of trial sites in country c, N_i is the total number of trial sites for i, and $\Delta FX^c_{[T]}$ is the percentage change in exchange rate across the whole duration of the clinical trial. Essentially, ΔFX_i measures the number-of-site-weighted rolling window exchange rate changes for trial i, with a higher ΔFX_i indicating higher exchange-rate-driven costs to the drug developer. If a trial is conducted exclusively in the U.S, then $\Delta FX_i = 0$. 35% of trials in our sample involve foreign trial sites (Table 1). Among foreign trials, the interquartile range in exchange variation is slightly above six percentage points.

Similar to Equation (1), we will then estimate:

$$\text{Suspension}_i = \beta \Delta FX_i + \gamma \text{Controls}_i + \text{FEs} + \eta_i \tag{3}$$

to study how firms' trial continuation decisions are affected by prior exchange-rate-induced investments made in the clinical trial project.

5 Results

5.1 Delay in Trial Completion and Project Continuation Versus Suspension

Table 2 presents the baseline results for how firms' decision to continue versus suspend a clinical trial after a given completed phase depends on the experienced delay in trial phase completion. Column (1), comparing trials initiated in the same quarter, finds a strong negative effect of delay on the likelihood of suspension, both economically and statistically speaking. A one standard deviation increase in delay in the preceding trial phase reduces the suspension probability by 4.4 percentage points, or 15% relative to

the baseline suspension probability of 31% (Table 1).

The effect of trial completion delay on trial continuation remains unchanged in Column (2) when we add firm fixed effects, as well as in Column (3) when we further add fixed effects for the drug's ICD category and for the trial phase (Phase I versus II trials). Finally, the effect remains unchanged in Column (4) when we add control variables for the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs).[10]

Figure 4 visualizes the negative relationship between trial completion delay and subsequent trial suspension, residualizing both variables with respect to the control variables and fixed effects in the final column of Table 2. The figure reveals an almost monotonic effect of delay on subsequent trial suspension.

Overall, the results in Table 2 and Figure 4 provide first evidence that prior investments by firms to a given clinical trial, in the form of time it takes to complete the trial *beyond* what firms anticipated at trial start, increases subsequent project commitment by firms as measured by the decision to subsquently suspend versus continue the trial.

5.2 Trial Site Congestion as an Instrument for Delay in Trial Completion

Table 3 presents the IV results, studying the effect of trial phase completion delay due to (changes in) trial site congestion on the decision to subsequently continue versus suspend the trial. Columns (1) and (2) show the first-stage results, i.e. the relation between delay and change in trial site congestion as detailed in Section 4.2, with and without the controls included in the last column of Table 2. The columns reveal a strong association between congestion and delay. Economically, a one standard deviation increase in the congestion instrument is associated with an increase in trial completion delay of close to four months. Statistically, the Kleibergen and Paap (2006) $F-$statistic is clearly above the common threshold for weak instruments of 10 (Stock et al., 2002). In Figure OA.2 of the Online Appendix, we include a binscatter plot visualizing the near-monotone relationship between trial site congestion increase and the resulting unanticipated clinical trial delays.

[10]The coefficients on controls (unreported) are intuitive. More projects in the pipeline and more competing projects are positively (but insignificantly) associated with suspension. More trial sites increases the likelihood of suspension, whereas more trial participants reduces it.

Columns (3) and (4) show the second-stage results. Delay in trial completion, instrumented with trial site congestion, continues to significantly affect the decision to continue versus suspend the trial. The economic magnitude of the effect is in the same ballpark as the OLS estimates in Table 3, and the coefficients on instrumented delay remain significant at 5% or 1%. (We continue to use two-way clustered standard errors, now corrected for using a two-stage estimation approach.) The IV results strengthen the interpretation that unanticipated, plausibly-exogenous delays in trial phase completion induce excessive project commitment by firms with respect to their subsequent project continuation propensity.

5.3 Exchange Rate Variation in Foreign Trials

Table 4 shows the validation exercise results for how firms' decision to continue versus suspend a clinical trial after a given completed phase depends on plausibly-exogenous exchange rate fluctuations experienced in foreign-based trials between trial start and completion, rather than for how project continuation depends on temporal investments made. Columns (1) and (2) are estimated on the subsample of foreign-based trials, with Column (2) adding controls as in Column (5) of Table 2. Columns (3) and (4) are estimated on the full sample including U.S.-based trials, and interact the exchange rate fluctuation variable, as well as controls and the drug indication and year-quarter fixed effects with an indicator for foreign-based trials.[11] All columns include a control for a trial's total duration, given that total trial duration determines the period over which exchange rate fluctuations can affect financial obligations.

Across columns, the coefficient on exchange rate changes is negative, implying that exchange-rate-driven cost increases make trial suspension less likely, and statistically significant. In terms of magnitudes, a one standard deviation increase in trial costs through exchange rates is estimated to reduce the likelihood of subsequent trial suspension by 1.75 percentage points, or 6% relative to the baseline. These estimated magnitudes are slightly lower compared to the effects of unexpected delays on suspension decisions, but still economically meaningful.

To further probe the robustness of the exchange rate results, in the Internet Appendix we conduct a placebo test that randomly assigns trial countries to trials in the sample and re-runs the full-sample regressions in Table 4. Figure OA.3 of the Online Appendix plots the distribution of the estimated

[11] We note that since $\Delta FX = 0$ for domestic trials, it is in fact the case that $\Delta FX = \Delta FX \times$ *Foreign Trial*, i.e., we adopt the interaction term notation solely for visual purposes.

coefficients for ΔFX_i by repeating the estimation 1,000 times. In contrast to the results in Table 4, the placebo test coefficients are not significantly different from zero, and not as large in economic magnitudes.

Overall, the results exploiting exchange rate variation provide additional evidence in favor of the interpretation that plausibly-exogenous increases in the amount of resources firms have expended on a clinical trial project increase firms' subsequent (excess) commitment to continue the project.

5.4 Underlying Channels

In this section, we examine various potential underlying channels that might lead to (excess) commitment and project continuation in response to trial phase completion delays. In Table 5 we examine agency conflicts inside the firm, initial firm expectations about trial length, financial constraints, and heterogeneities by trial phase. In Table 6 we discuss real options as a potential underlying mechanism. In addition to these firm- and project characteristics-related channels, in Table 7 we examine mechanisms pertaining to the ultimate decision-maker responsible for the R&D investment, the drug developing firm's CEO.

5.4.1 Firm- and Project Characteristics-Related Channels

Agency Conflicts—Lack of Other Viable Drug Candidates? In influential prior work, Guedj and Scharfstein (2004) find that early-stage bio-pharmaceutical firms are reluctant to abandon their only viable drug candidate, distorting trial continuation decisions. To check whether such early-stage-firm agency problems can explain our findings, we restrict the sample to trials of firms with more than ten drug projects in the company's pipeline at trial completion, i.e., to firms with other investment opportunities.

Doing so, we conclude that a mechanism related to an aversion to "start over or liquidate" does not explain our results. For one, the many-projects-in-the-pipeline restriction only drops relatively few observations from the sample (less than 25%). That is, the findings in the previous sections predominantly come from large firms with many projects to begin with, leaving little room for an early-stage-firm channel as in Guedj and Scharfstein (2004). Additionally, and consistent with the preceding point, Column (1) of Table 5 shows that unexpected delays continue to predict trial continuation decisions with very similar magnitude among firms with more than ten drug projects in the pipeline. Thus, the effects we find are not

driven by early-stage firms with lack of alternative drugs as in Guedj and Scharfstein (2004).

Expectations of Trial Duration. Another possible channel is a selection-based mechanism based on initial expectations and differential willingness to undertake lengthy trials. In the data, we observe a positive correlation between a trial's anticipated duration (anticipated end date minus start date) and the trial's delay (end date minus anticipated end date). Thus, it could in principle be the case that our delay results reflect initially optimistic drug developers selecting into trials with a longer anticipated duration *and* a longer delay period on average.

However, the empirical correlation between anticipated duration and delay, while positive, is relatively small ($\rho = 0.07$). Furthermore, when we directly control for anticipated duration in Column (2) of Table 5, we obtain very similar results.[12] In light of this evidence, differential sorting by firms into trials with different lengths or delays based on initial expectations is unlikely to be an underlying mechanism of our findings.

Financial Constraints. Next, we test whether the commitment in drug-related decision-making may be driven by financial constraints of the drug-developing firms. Intuitively, if firms are or become financially constrained, they may be more likely to stick to a chosen course of action. To investigate this, we augment our specifications with a series of financial constraints measures.[13] Specifically, the *Constrained* variable in the third column of Table 5 is an indicator variable that equal one if the firm's Whited and Wu (2006) (WW) index at trial end is in the top quartile of the index's sample distribution. Column (3) shows that the effect of delay on trial suspension remains unchanged with the added constraints measure. Thus, the path dependence in decision-making is not merely the result of constrained firms not being able to pivot (also in combination with the fact that most firms in our sample have many other viable drug candidate projects).[14]

Phase I Versus II Trials. The final two columns in Table 5 separate the effect by whether a trial is a Phase I or Phase II trial. As discussed in Section 3.1, the primary endpoint of Phase I trials is safety, whereas that of Phase II trials is drug efficacy, and Phase II trials tend to be longer and involve more

[12] In unreported tests, we include second- or third-order polynomials for anticipated duration and find nearly identical results compared to those reported in Table 5.

[13] We note that adding the financial constraints measures restricts the sample to include public firms only.

[14] We show robustness to other measures of financial constraints, in particular the Hadlock and Pierce (2010) (HP) and Kaplan and Zingales (1997) (KZ) indices, as well as to constraint measures defined as of the time of trial start in Appendix Table OA.2. Irrespective of the measure we use, we continue to find a strong effect of trial delay on subsequent suspension probabilities.

participants. The effect of trial delay on subsequent continuation versus suspension is robust and large for both Phase I and II trials. Thus, our findings cannot be not explained by any trial-phase-specific channels or confounders, and instead unveil a robust and wide-ranging phenomenon in the realm of new drug development.

Real Options. The R&D process involves a sequence of actions and hence involves option values at each decision-making node, in particular when past experience might reveal information about likely future values (Pindyck, 1991; Dixit et al., 1994). McAfee et al. (2010) propose a stylized real option model in which more past investments might reveal better prospects about the option value of further investment, which in turn might be related to why unexpected delay in a trial is positively correlated with the likelihood of the firm to advance it through the next phase trial. If trial delays indeed reveal valuable information about further investment, we would expect the delay to be positively correlated with proxies for the prospect of a drug project: either higher expected sales after launching the drug in the market (higher revenues conditional on getting the FDA approval) or fewer adverse reactions among trial participants, implying a a higher chance for the drug to survive all stipulated phases of clinical trials (i.e., a higher probability of obtaining the FDA approval).

To test for the possibility of a potential real options channel, we examine if delays in trial completion predict higher expected drug sales conditional on drug launch or fewer adverse reactions among subsequent trial participants. We adopt a specification following Equation (1). Panel A of Table 6 examines the outcome variables pertaining to expected drug sales. We use four different measures, including the expected first-year sales, the sum of first three-year sales, the sum of first five-year sales, and the average annual sales after the drug is approved by the FDA. All sales are million U.S. dollars in 2017 adjusted by GDP deflators and logged. Contrary to the real option hypothesis, delays in trial phase completion do not predict expected drug sales, neither economically nor statistically speaking. Panel B examines the outcome variables pertaining to adverse reactions during clinical trials. Columns (1) and (2) use the total and maximum percentage of raw adverse reaction counts (an enrolled participant might have multiple reports) relative to the total number of enrolled patients across various adverse symptoms. Instead of a negative correlation, we find a significantly positive relation between delays in trial completion and the number of adverse reaction reports. We note that the two measures include many light adverse reactions,

such as headache, nausea, and fatigue, which are of less importance for drug evaluation by the FDA. In columns (3) and (4), we only include life-threatening adverse reactions, such as acute myelogenous leukemia, basal cell carcinoma, and brain hemorrhage. Again, the results are inconsistent with the real option mechanism as the coefficients on delay in trial completion are economically and statistically insignificant. Taken together, these results suggest that a real option channel is unlikely to explain the delay-induced continued commitment to R&D projects in our setting.

5.4.2 CEO-Related Channels

If drug project decisions are affected by firms' past actions and investments, it is natural to relate this behavior to the responsible decision-makers inside the firm. We examine how the effect of drug delays on suspension decisions varies with CEOs' pay sensitivity to stock price changes as well as with their personal responsibility for the project as gauged by whether they led the firm at trial start.

CEO Pay Sensitivity to Firm Value Changes. We first study how CEO pay-performance sensitivity (delta) mediates the relationship between trial delays and continuation versus suspension decisions.[15] This test is motivated by our theoretical framework from Section 2. Specifically, the extended framework with asymmtric information about managerial ability predicts a negative interaction effect between the sensitivity of CEO compensation with respect to the firm's stock price and experienced unexpected delay. Consistent with this, Column (1) of Table 7 finds that the effect of trial delay on the decision to suspend versus continue is substantially more pronounced among CEOs with greater exposure to stock price changes.

CEO Changes. Second, we study how firms' project continuation decisions after trial delays vary with CEOs' personal responsibility for having initiated the project. In the context of mergers and acquisitions, Guenzel (2023) shows that managers engage in sunk-cost thinking with respect to acquired targets, distorting their subsequent decisions to abandon acquired businesses through divestiture. Guenzel (2023) finds that the sunk cost effects are driven by the CEO who made the initial acquisition. To test for

[15] We obtain the CEO delta data from Lalitha Naveen's website at `https://sites.temple.edu/lnaveen/data/` (Coles et al. 2006; Core and Guay 2002). We use the natural logarithm of delta in the analysis in Table 7 as in Coles et al. (2006), and standardize the variable to have a mean of zero and standard deviation of one for expositional purposes. Our findings below are similar when using other related measures, in particular the scaled wealth-performance sensitivity measure by Edmans et al. (2009).

a similar CEO effect in our setting, we create a *Same CEO* indicator that equals one if the firm's CEO is the same at trial start and completion, and then interact the CEO variable with our delay measure. We emphasize the test of interest is *not* whether an incumbent is more or less likely to suspend a trial, but how the effect of delay varies depending on the CEO regime.

In Column (2) of Table 7, we find that the commitment effects we uncover are more pronounced when the CEO at trial end is the same CEO that was at the helm at trial start, consistent with a personal responsibility channel. The effect of unexpected trial phase completion delay on subsequent commitment to the trial is about three times as large when there is no CEO change between trial start and end. In Column (3), we further find that these effects are not more pronounced for younger CEOs who have longer careers ahead of them. This suggests that the managerial responsibility channel likely arises not only from externally-oriented career/reputation concerns (which are arguably stronger for younger CEOs; Fama 1980; Holmström 1982; Gibbons and Murphy 1992), but also from internally-oriented, psychological motives such as avoidance of cognitive dissonance (Festinger 1962) or realization of losses (Kahneman and Tversky 1979).

Overall, the results in Table 7 support management-induced frictions as underlying mechanisms for the documented delay-induced (excess) commitment to R&D projects within firms.[16]

6 Identification Challenges

In this section, we conduct multiple robustness checks to address potential remaining identification concerns. First, one concern might be that delay in clinical trials arises because of regulatory reasons, e.g., FDA clinical holds. We provide evidence that this does not explain our results. Second, one might be concerned that clinical trials experience delays because developing firms experiment with multiple drug doses. Drug developers may thus obtain better information on the optimal dosage and may be more willing to continue the project. We implement several tests that address this alternative explanation. Third, another important threat to identification is unobserved drug quality. We use direct drug quality

[16] It is also useful to discuss how our findings relate to managerial overconfidence (Malmendier and Tate, 2005, 2008; Galasso and Simcoe, 2011; Hirshleifer et al., 2012). Overconfident managers might systematically underestimate project duration while overestimating project prospects, in line with our OLS-based findings in Table 2. At the same time, only the commitment mechanism, but not an overconfidence mechanism, predicts the IV-based findings in Table 3 (as long as overconfidence is uncorrelated with changes in trial site congestion).

proxies and show that these are uncorrelated with our delay measures. Finally, we discuss additional miscellaneous robustness tests.

FDA Clinical Trial Holds. First, despite the corroborating trial-cost-based evidence above, one potential concern regarding the main, delay-based results is that trial completion delays might be caused by FDA-imposed clinical holds. An FDA clinical hold is an order issued by FDA to the sponsor of the drug development to delay a proposed clinical investigation. Common reasons for clinical holds include the presence of impurity profile indicative of a potential health hazard, missing supporting information on the anticipated dose or exposure, or insufficient information on the clinical protocol (Lapteva and Pariser, 2016). After the hold, drug developers would work with the FDA to improve the clinical protocol and address all deficiencies until the hold is lifted. A trial that experienced a hold may have a prolonged trial duration, but the FDA may potentially help improve the trial quality throughout the hold period. This could also explain a relation between trial completion delays and continuation decision, but would be due to external, and potentially project-improving factors.

We note that this concern applies mostly to the results in Table 2, rather than the IV-based results in Table 3. To allay the concern further, Online Appendix Table OA.3 shows that there is no statistically significant relationship between the presence of clinical holds and trial duration, trial delays, or trial site congestion measures (our instrument for delays). We conclude that it is unlikely that clinical holds are a driver of any of our results.

Drug Experimentation. Second, another potential identification threat is that drug developers experiment with multiple dosages in one clinical trial. As a consequence, it could be more likely for the trial to become delayed due to the more challenging trial management across multiple tracks. In the meantime, the drug developer might obtain more scientific data points on the optimal dosage and the drug's efficacy via larger scale of dosage experiments. Better knowledge on how the drug works in the human body could ultimately make the developer more likely to continue the project, which could potentially explain the observed patterns.

To alleviate this concern, we construct a variable, *Num of Doses in a Trial*, defined as the number of dosages tested in a clinical trial and examine its correlation with our delay measures. Table OA.4 in the Online Appendix reports the results. Column (1) shows that the number of dosages experimented with

in a trial is not significantly correlated with trial delay. Also, column (2) confirms that our instrument, *Trial Site Congestion*, is barely correlated with the number of doses either. To further show the dosage experimentation mechanism does not explain our results, Table OA.5 in the Online Appendix replicates the regressions in Table 3 by adding *Num of Doses in a Trial* as an additional control variable. The main conclusions remain unchanged and the estimated coefficients on *Num of Doses in a Trial* are statistically insignificant across all columns.

Drug Quality. Third, one may still be concerned about unobserved drug quality. Even though our IV strategy can potentially alleviate such concerns, if (changes in) trial site busyness are uncorrelated with drug quality, we provide more evidence to further allay this concern. Specifically, we use two objective measures of drug quality from the GlobalData Pharma database. One is the drug-specific likelihood of approval (LOA) and the other is the drug's phase transition success rate (PTSR). GlobalData utilizes a machine learning algorithm combined with historical data to determine the probability of successful progression and market authorization for pipeline drugs from their current clinical trial stage (Wong et al., 2019). For example, the Phase-III drug candidate, *canagliflozin*, developed by Johnson & Johnson for diastolic heart failure, is estimated to have a 35% drug-specific LOA and a 56% drug-specific PTSR. For each drug project in our sample, we match it with the GlobalData Pharma database by drug names and compute the average LOA and PTSR across companies and indications as the proxies for drug quality.

Both drug quality proxies are uncorrelated with our delay measures. Table OA.6 of the Online Appendix reports the regression results. Columns (1) to (4) examine the correlations by regressing trial delay on LOA and PTSR. The estimated coefficients are both economically and statistically insignificant. Columns (5) to (8) report the results when regressing our instrument, *Trial Site Congestion*, on LOA and PTSR. Again, all estimates are insignificant, indicating no significant correlation of trial delays or trial site congestion with the drug quality proxies. We conclude that unobserved drug quality is unlikely to drive our results.

Miscellaneous Tests. Finally, we implement additional miscellaneous robustness tests. In Table OA.7, we show that our results are driven by trials experiencing a positive (i.e., ≥ 0) delay, whereas negative delay does not significantly predict the decision to continue versus suspend projects.[17] In Table

[17] Only about 15% of trials in our sample are completed prior to the originally anticipated completion date.

OA.8, we first show that our results remain robust when we exploit variation in trial completion delay within trials started in the same quarter and targeting the same ICD category (i.e., time-by-ICD-category fixed effects; Column (1)). Further, countries differ in the regulatory environment and requirements they impose on clinical trial activities, as well as in associated cost structures. We therefore re-estimate our results separately by geography. The results are robust to being estimated on US trials only (Column (2)) and non-US trials only (Column (3)). Similar to the robustness across trial phases in Table 5, these results highlight the general applicability of our findings.

7 Implications for Firm Investment and Consumer Welfare

In the final part of the paper, we investigate the resulting effects of delay-induced commitment to drug projects on a wide range of firm and consumer R&D outcomes.

7.1 Outcomes Variables Related to Firm Investment

To explore the implications of delay-induced commitment to drug projects on firm investment, we study spillover effects on other investments made by firms. Specifically, we examine the take-up of new drug projects, and test whether delay-induced commitment leads to crowding out of other projects. To this end, we compute three outcomes variables. First, we compute the total number of new projects (in logs) initiated by company within one year of completing the focal project phase (i.e., the project included in our main sample). Second, we compute the total number of new projects that share the ICD indication as the focal project. Third, we compute the total number of new projects that do not share the same ICD indication. As throughout our analysis, projects are defined as a drug by indication, and we include as newly initiated projects those that are initiated with development status "Discovery" or "Preclinical." Any potential crowd-out effects in firm investment may worsen the firm's future drug pipeline and revenue prospects and, ultimately, may also have negative implications for consumers in terms drug choices available to them.

7.2 Outcome Variables Related to Patient Welfare

To explore the implications of elevated commitment in pharmaceutical R&D for direct consumer outcomes observable in the near term, we consider two distinct outcome variables. First, at the extensive margin, an approved new drug can save lives if there is no alternative drug that exists in the market to target the disease. To measure this effect, we examine drug approval outcomes of so-called orphan drugs targeting infrequent diseases with typically few available treatment options. Second, at the intensive margin, drugs might yield adverse reactions among users once approved, which is indirectly related to the safety of the drug and relevant for consumer welfare. To gauge the impact along this dimension, we use the count of adverse events occurring after the drug is launched in the marketplace.

7.2.1 Econometric Framework

Unlike the OLS and 2SLS approaches adopted in Section 4, one complication in examining consumer outcomes, and in particular the drug's adverse reactions, is endogenous sample selection. Adverse events only occur conditional on the drug being approved and launched in the market. To account for this selection, we apply a parametric framework for the tests related to adverse events. This framework addresses sample selection with endogenous explanatory variables, and modifies the classic Heckman selection model (e.g., Heckman, 1979; Wooldridge, 2010). Alternatively, it would be possible to adopt non-parametric identification strategies to estimate bounds for average treatment effects (e.g., Lee, 2009; Huber, 2014; Bartalotti et al., 2021). We pursue the parametric approach since it is easy to implement, flexible with respect to continuous treatments, and able to circumvent the complication of estimates for observed-only-when-treated sub-populations. We lay out our framework as follows.

Let i denote a drug project and $\mathbf{X}_i$ denote all observed drug and developer characteristics except our proxy for unexpected delays, or "temporal costs incurred," c_i. When this influences drug developers' decisions to continue project i, we would expect that the drug approval outcome, Y_i^{app}, depends on c_i. Y_i^{app} is an indicator equal to one if project i is approved by FDA after completing all required clinical trials and proving that the drug is safe and effective. We define

$$\tilde{Y}_i^{app} = \beta_1 c_i + \beta_2' \mathbf{X}_i + u_i \tag{4}$$

as the latent variable that maps into potential outcomes Y_i^{app} such that $Y_i^{app} = \mathbb{1}\left\{\tilde{Y}_i^{app} > 0\right\}$. u_i is a random variable representing characteristics unobserved by the econometrician or the idiosyncratic shock affecting drug safety or expected profitability.

We define other outcomes variables, in particular adverse events, as $\tilde{Y}_i^{out}$. Assume $\tilde{Y}_i^{out}$ is a linear function of observables $\mathbf{X}_i$ and c_i so that

$$\tilde{Y}_i^{out} = \alpha_1 c_i + \alpha_2' \mathbf{X}_i + v_i \tag{5}$$

where v_i is the residual term. The sample selection is reflected in the fact that we can observe $\tilde{Y}_i^{out}$ only if drug project i is approved by the FDA. Therefore, we define the observable outcome $Y_i^{out} = \tilde{Y}_i^{out} \times \tilde{Y}_i^{app}$.

In our setting, the endogenous sample selection is embodied in the possible correlation between c_i and u_i while the endogenous explanatory variable (or endogenous treatment) is captured by the fact that c_i might be correlated with v_i. Following a similar strategy as in Section 4.2, we use an exogenous variable to instrument for c_i. Specifically, assume c_i is generated according to the model

$$c_i = \gamma_1 z_i + \gamma_2' \mathbf{X}_i + \xi_i \tag{6}$$

where z_i is the instrumental variable.

To close the model, we impose the following parametric assumption on the error terms so that

$$\begin{pmatrix} u_i \\ \xi_i \\ v_i \end{pmatrix} \sim Normal \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix} \right] \tag{7}$$

where (u_i, ξ_i, v_i) is jointly normally distributed and independent of all observables z_i and $\mathbf{X}_i$. Without loss of generality, we normalize $var(u_i) = 1$ as Y_i^{app} is a binary variable. Our setting and assumptions are similar to the sample selection model with endogenous explanatory variables in Chapter 19 of Wooldridge (2010). The key difference is that we allow the endogenous variable to enter the sample selection Equation (4). To discipline the model, we hence impose extra structure on the residuals through Equation (7). Note that Equation (7) is a stronger version of the restriction imposed on the residuals. Our estimation can

also be obtained under a weaker set of assumptions where (u_i, ξ_i) is jointly normally distributed and $E[v_i \mid \beta_1 \xi_i + u_i] = \alpha_3(\beta_1 \xi_i + u_i)$.

We detail our estimation procedure as follows:

Step one: Obtain $(\hat{\beta}_1, \hat{\beta}_2', \hat{\gamma}_1, \hat{\gamma}_2', \hat{\sigma_{12}}, \hat{\sigma_2^2})$ from an IV-probit model of Y_i^{app} on c_i and $\mathbf{X}_i$ where c_i follows Equation (6). Compute the estimated inverse Mills ratios, $\hat{\lambda}_i = \lambda(\frac{\hat{\beta}_1 \cdot \hat{\gamma}_1 z_i + (\hat{\beta}_1 \cdot \hat{\gamma}_2' + \hat{\beta}_2')\mathbf{X}_i}{\sqrt{\hat{\beta}_1^2 \hat{\sigma_2^2} + 2\hat{\beta}_1 \hat{\sigma_{12}} + 1}})$ where $\lambda(\cdot) = \frac{\phi(\cdot)}{\Phi(\cdot)}$ with $\phi(\cdot)$ as the probability density function of the standard normal distribution and $\Phi(\cdot)$ as the cumulative density function of the standard normal distribution.

Step two: Use the selected sample with Y^{out} being observed and estimate the equation

$$Y_i^{out} = \alpha_1 c_i + \alpha_2' \mathbf{X}_i + \alpha_3 \hat{\lambda}_i + error_i$$

by 2SLS, using instruments $(z_i, \mathbf{X}_i, \hat{\lambda}_i)$.

Our estimation procedure is similar to that proposed in Wooldridge (2010). We provide a detailed derivation in the Online Appendix (OA Section III.C). Standard errors and test statistics are corrected for the generated regressor problem by bootstrapping.

7.3 Effects on Firm Investment

We first assess the effects of delay-induced drug continuation on other firm investments, and specifically possible crowd-out effects of other drug projects. As alluded to above, for this part we can continue to use the IV strategy from Section 4.2. That is, we use (change in) trial site congestion as an instrument for trial completion delay, and study the resulting effect on the number of newly initiated drug projects in the year after the focal project (i.e., the trial phase in our main sample) is completed.

We present the second-stage estimation results of this IV approach in Table 8, with the first-stage Kleibergen and Paap (2006) $F-$statistics on the instrument shown at the bottom. (The sample size is somewhat reduced relative to Table 3 since we require one-year forward-looking information on each project for this test, thus losing observations at the end of our sample period.) As Columns (1) and (2) reveal, congestion-induced trial completion delay significantly reduces the number of newly initiated projects. A one-month increase in (instrumented) delay is associated with a 1% reduction in new projects. A one-year delay is associated with approximately one fewer new project started, relative to a median

of four newly launched projects. As the remaining columns show, this effect is concentrated among new projects targeting other diseases (i.e., drugs with non-overlapping ICD indications), implying that crowding out appears to occur mainly with respect to unrelated drugs. In the long run, these crowd-out spillover effects on other new drug development may negatively affect both firm profits and drug choices available to consumers.

7.4 Effects on New Drug Launches With Few Existing Medications (Orphan Drugs)

We next assess more directly observable consumer welfare implications, and first as they pertain to the development and launch of orphan drugs—drugs for infrequent diseases for which there are commonly no other or only few other treatment alternatives. For this part, we can also rely on the IV strategy from Section 4.2. Specifically, *conditioning* on clinical trials involving orphan drugs, we examine how the likelihood of eventual orphan drug approval by the FDA (i.e., launch of the drug in the market) varies with congestion-induced delay in the completion of Phase I or Phase II clinical trials.

We present the corresponding results in Table 9. Similar to Table 3, Column (1) presents the first-stage results. Despite the reduction in sample size (due to the restriction to orphan drug clinical trials), the first stage remains strongly significant. Column (2) then presents the second-stage results, with the indicator variable for FDA orphan drug approval as the dependent variable of interest. We find a significant positive effect of instrumented delay on the probability that orphan drugs are ultimately launched in the marketplace. For example, a six-month congestion-induced delay in trial completion is associated with a 1.5 percentage point increase in the probability of orphan drug market launch. This is an economically large effect, given the relatively low baseline probability of eventual FDA approval of orphan-drug-related clinical trials of just above 10% in the sample.

7.5 Effects on Adverse Events Related to Approved Drugs

Finally, we assess welfare implications as they pertain to adverse effects in patients of approved drugs. For this part, we make use of the econometric sample-selection framework detailed in Section 7.2.1. Table 10 shows the results, with standard errors bootstrapped using 100 iterations. Column (1) shows the first stage results, i.e., the first part of the second step outlined in Section 7.2.1. Similar to Table 9, we continue

to observe a strong relation between trial site congestion and delay despite the reduction in sample size. (In Table 10, the sample size reduction is due to the fact that we focus on outcomes of approved drugs.)

Columns (2) and (3) show the second-stage results, examining adverse events either over a one-year or three-year horizon since drug launch. Importantly, we continue to include ICD code fixed effects (as well as other fixed effects) in this analysis, so all results come from within-drug-category comparisons. Within one and three years since drug launch, the median number of reported adverse events in our sample is 37 and 345, respectively. The point estimates in Columns (2) and (3) point to modest increases in adverse events after congestion-induced delay but are notably insignificant. In both columns, a one-month instrumented increase in delay is estimated to increase the prevalence of adverse events by about 1%.

Taken together, the results in Sections 7.3 to 7.5 pinpoint the intricacies and complexities of how delay-driven drug continuation can affect (consumer) welfare. While the effects we uncover increase the breadth of treatments for infrequent diseases for which there are often no other existing medications, they might also lead to the launch of drugs with more adverse effects, as well as a reduction in the take-up of other drug projects. More broadly, the results in this paper highlight that for non-firm parties, firms' friction-induced R&D decisions do not need to exclusively entail welfare costs; instead, they can, and in fact may frequently, induce positive externalities on (subsets of) other parties.

8 Conclusion

In this paper, we study the extent to which firms' commitment to ongoing R&D projects is influenced by the amount of temporal (and financial) resources the firm has already expended on the project, and examine resulting welfare implications. We study these questions using project-level R&D data from the pharmaceutical industry, namely clinical trials data with granular information on many important dimensions including project timelines and project outcomes. Clinical trials that firms complete but that experienced unanticipated and plausibly exogenous delays are significantly more likely to be advanced by firms through the next trial phase, even though delays are empirically uncorrelated with various project quality measures. The economic magnitudes we uncover are substantial. Firms' delay-induced commitment to drug projects crowds out the initiation of other drug projects. The resulting implications for consumers are nuanced. On the one hand, delay-induced commitment to drug projects increases the

likelihood of drug launches for diseases with few existing treatment options (orphan drugs). On the other hand, marginally-approved drugs (not restricted to orphan drugs) in response to delays may be associated with modestly higher adverse event counts.

Our findings suggest a variety of potential directions for future research. One promising area for future work would be to investigate heterogeneities along several dimensions. For example, it would be interesting to explore how the magnitude of excess R&D commitment varies with the organizational structure (e.g., top-down versus decentralized decision-making) and the organizational culture (e.g., the degree of tolerance for failure), and whether there are heterogeneous effects on consumer welfare outcomes (e.g., by severity of adverse effects, or by various consumer demographics). Additionally, it would be interesting to study competitors' responses to firms' excess commitment to R&D projects and general equilibrium effects.

References

ACEMOGLU, D. AND J. LINN (2004): "Market size in innovation: theory and evidence from the pharmaceutical industry," *Quarterly Journal of Economics*, 119, 1049–1090.

AGHAMOLLA, C. AND R. T. THAKOR (2022): "IPO peer effects," *Journal of Financial Economics*, 144, 206–226.

AUGENBLICK, N. (2016): "The sunk-cost fallacy in penny auctions," *Review of Economic Studies*, 83, 58–86.

AUSTIN, D. H. AND T. HAYFORD (2021): "Research and development in the pharmaceutical industry," Government report, Congress of the United States, Congressional Budget Office.

AZOULAY, P., J. S. GRAFF ZIVIN, D. LI, AND B. N. SAMPAT (2019): "Public R&D investments and private-sector patenting: evidence from NIH funding rules," *Review of Economic Studies*, 86, 117–152.

BARRERO, J. M. (2022): "The micro and macro of managerial beliefs," *Journal of Financial Economics*, 143, 640–667.

BARTALOTTI, O., D. KÉDAGNI, AND V. POSSEBOM (2021): "Identifying marginal treatment effects in the presence of sample selection," *Journal of Econometrics*.

BONAIME, A. A. AND Y. E. WANG (2023): "Mergers, product prices, and innovation: Evidence from the pharmaceutical industry," *The Journal of Finance, forthcoming*.

BRYAN, K. A. AND H. L. WILLIAMS (2021): "Innovation: market failures and public policies," in *Handbook of Industrial Organization*, Elsevier, vol. 5, 281–388.

BUDISH, E., B. N. ROIN, AND H. WILLIAMS (2015): "Do firms underinvest in long-term research? Evidence from cancer clinical trials," *American Economic Review*, 105, 2044–85.

CAMERER, C. F. AND R. A. WEBER (1999): "The econometrics and behavioral economics of escalation of commitment: A re-examination of Staw and Hoang's NBA data," *Journal of Economic Behavior and Organization*, 39, 59–82.

COHEN, L., U. G. GURUN, AND D. LI (2021): "Internal Deadlines, Drug Approvals, and Safety Problems," *American Economic Review: Insights*, 3, 67–82.

COLES, J. L., N. D. DANIEL, AND L. NAVEEN (2006): "Managerial incentives and risk-taking," *Journal of Financial Economics*, 79, 431–468.

CORE, J. AND W. GUAY (2002): "Estimating the value of employee stock option portfolios and their sensitivities to price and volatility," *Journal of Accounting Research*, 40, 613–630.

CUNNINGHAM, C., F. EDERER, AND S. MA (2021): "Killer acquisitions," *Journal of Political Economy*, 129, 649–702.

CZIRAKI, P. AND D. JENTER (2022): "The market for CEOs," Working paper, European Corporate Governance Institute.

DiMasi, J. A., H. G. Grabowski, and R. W. Hansen (2016): "Innovation in the pharmaceutical industry: new estimates of R&D costs," *Journal of Health Economics*, 47, 20–33.

DiMasi, J. A., R. W. Hansen, and H. G. Grabowski (2003): "The price of innovation: new estimates of drug development costs," *Journal of Health Economics*, 22, 151–185.

Dixit, R. K., A. K. Dixit, and R. S. Pindyck (1994): *Investment under uncertainty*, Princeton university press.

Edmans, A., X. Gabaix, and A. Landier (2009): "A multiplicative model of optimal CEO incentives in market equilibrium," *Review of Financial Studies*, 22, 4881–4917.

Fama, E. F. (1980): "Agency problems and the theory of the firm," *Journal of Political Economy*, 88, 288–307.

Fee, C. E. and C. J. Hadlock (2004): "Management turnover across the corporate hierarchy," *Journal of Accounting and Economics*, 37, 3–38.

Fee, C. E., C. J. Hadlock, and J. R. Pierce (2018): "New evidence on managerial labor markets: An analysis of CEO retreads," *Journal of Corporate Finance*, 48, 428–441.

Festinger, L. (1962): *A theory of cognitive dissonance*, vol. 2, Stanford University Press.

Galasso, A. and T. S. Simcoe (2011): "CEO overconfidence and innovation," *Management Science*, 57, 1469–1484.

Gibbons, R. and K. J. Murphy (1992): "Optimal incentive contracts in the presence of career concerns: Theory and evidence," *Journal of Political Economy*, 100, 468–505.

Guedj, I. and D. S. Scharfstein (2004): "Organizational scope and investment: Evidence from the drug development strategies and performance of biopharmaceutical firms," .

Guenzel, M. (2023): "In Too Deep: The Effect of Sunk Costs on Corporate Investment," *Journal of Finance, forthcoming*.

Guenzel, M. and U. Malmendier (2020): "Behavioral Corporate Finance: The Life Cycle of a CEO Career," *Oxford Research Encyclopedia of Economics and Finance, September 2020*.

Hadlock, C. J. and J. R. Pierce (2010): "New evidence on measuring financial constraints: Moving beyond the KZ index," *Review of Financial Studies*, 23, 1909–1940.

Hall, B. H. and J. Lerner (2010): "The financing of R&D and Innovation," in *Handbook of the Economics of Innovation*, Elsevier, vol. 1, 609–639.

Heckman, J. J. (1979): "Sample selection bias as a specification error," *Econometrica*, 153–161.

Hirshleifer, D., A. Low, and S. H. Teoh (2012): "Are overconfident CEOs better innovators?" *Journal of Finance*, 67, 1457–1498.

Ho, T.-H., I. P. Png, and S. Reza (2018): "Sunk cost fallacy in driving the world's costliest cars," *Management Science*, 64, 1761–1778.

HOLMSTRÖM, B. (1982): "Managerial incentive schemes—a dynamic perspective," *Essays in economics and management in honour of Lars Wahlbeck, Helsinki: Swenska Handelshögkolan*.

HUBER, M. (2014): "Treatment evaluation in the presence of sample selection," *Econometric Reviews*, 33, 869–905.

KAHNEMAN, D. AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–291.

KAPLAN, S. N. AND M. SORENSEN (2021): "Are CEOs Different?" *Journal of Finance*, 76, 1773–1811.

KAPLAN, S. N. AND L. ZINGALES (1997): "Do investment-cash flow sensitivities provide useful measures of financing constraints?" *Quarterly Journal of Economics*, 112, 169–215.

KLEIBERGEN, F. AND R. PAAP (2006): "Generalized reduced rank tests using the singular value decomposition," *Journal of Econometrics*, 133, 97–126.

KRIEGER, J., D. LI, AND D. PAPANIKOLAOU (2022a): "Missing novelty in drug development," *Review of Financial Studies*, 35, 636–679.

KRIEGER, J. L., X. LI, AND R. T. THAKOR (2022b): "Find and replace: R&D investment following the erosion of existing products," *Management Science*.

LAPTEVA, L. AND A. R. PARISER (2016): "Investigational New Drug applications: a 1-year pilot study on rates and reasons for clinical hold," *Journal of Investigative Medicine*, 64, 376–382.

LEE, D. S. (2009): "Training, wages, and sample selection: Estimating sharp bounds on treatment effects," *Review of Economic Studies*, 76, 1071–1102.

LERNER, J. AND U. MALMENDIER (2010): "Contractibility and the design of research agreements," *American Economic Review*, 100, 214–46.

LERNER, J., H. SHANE, AND A. TSAI (2003): "Do equity financing cycles matter? Evidence from biotechnology alliances," *Journal of Financial Economics*, 67, 411–446.

LI, X., T. LIU, AND L. A. TAYLOR (2023): "Common ownership and innovation efficiency," *Journal of Financial Economics*, 147, 475–497.

LIU, T. (2021): "Financing Experimentation," Working paper, MIT Sloan School of Management.

LO, A. W. AND R. T. THAKOR (2022): "Financing biomedical innovation," *Annual Review of Financial Economics*, 14, 231–270.

——— (2023): "Financial intermediation and the funding of biomedical innovation: a review," *Journal of Financial Intermediation*, 101028.

MA, Y., T. ROPELE, D. SRAER, AND D. THESMAR (2020): "A quantitative analysis of distortions in managerial forecasts," Working paper, National Bureau of Economic Research.

MALMENDIER, U. (2018): "Behavioral Corporate Finance," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 1, 277–379.

MALMENDIER, U. AND G. TATE (2005): "CEO overconfidence and corporate investment," *Journal of Finance*, 60, 2661–2700.

——— (2008): "Who makes acquisitions? CEO overconfidence and the market's reaction," *Journal of Financial Economics*, 89, 20–43.

MANEZ, J. A., M. E. ROCHINA-BARRACHINA, A. SANCHIS, AND J. A. SANCHIS (2009): "The role of sunk costs in the decision to invest in R&D," *Journal of Industrial Economics*, 57, 712–735.

MCAFEE, R. P., H. M. MIALON, AND S. H. MIALON (2010): "Do sunk costs matter?" *Economic Inquiry*, 48, 323–336.

MOSER, P. (2016): "Patents and innovation in economic history," *Annual Review of Economics*, 8, 241–258.

PINDYCK, R. S. (1991): "Irreversibility, Uncertainty, and Investment," *Journal of Economic Literature*, 29, 1110–1148.

REY, P. AND B. SALANIE (1990): "Long-term, short-term and renegotiation: On the value of commitment in contracting," *Econometrica: Journal of the Econometric Society*, 597–619.

ROBINSON, D. T. AND T. E. STUART (2007): "Financial contracting in biotech strategic alliances," *Journal of Law and Economics*, 50, 559–596.

SHADBOLT, C., E. NAUFAL, S. BUNZLI, V. PRICE, S. RELE, C. SCHILLING, S. THURAISINGAM, L. S. LOHMANDER, Z. J. BALOGH, P. CLARKE, ET AL. (2023): "Analysis of Rates of Completion, Delays, and Participant Recruitment in Randomized Clinical Trials in Surgery," *JAMA network open*, 6, e2250996–e2250996.

STAW, B. M. (1976): "Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action," *Organizational behavior and human performance*, 16, 27–44.

STAW, B. M., S. G. BARSADE, AND K. W. KOPUT (1997): "Escalation at the credit window: A longitudinal study of bank executives' recognition and write-off of problem loans." *Journal of Applied Psychology*, 82, 130.

STAW, B. M. AND H. HOANG (1995): "Sunk costs in the NBA: Why draft order affects playing time and survival in professional basketball," *Administrative Science Quarterly*, 474–494.

STOCK, J. H., J. H. WRIGHT, AND M. YOGO (2002): "A survey of weak instruments and weak identification in generalized method of moments," *Journal of Business & Economic Statistics*, 20, 518–529.

THAKOR, R. T. AND A. W. LO (2022): "Competition and R&D financing: Evidence from the biopharmaceutical industry," *Journal of Financial and Quantitative Analysis*, 57, 1885–1928.

VYTLACIL, E. (2002): "Independence, monotonicity, and latent index models: An equivalence result," *Econometrica*, 70, 331–341.

WHITED, T. M. AND G. WU (2006): "Financial constraints risk," *Review of Financial Studies*, 19, 531–559.

WONG, C. H., K. W. SIAH, AND A. W. LO (2019): "Estimation of clinical trial success rates and related parameters," *Biostatistics*, 20, 273–286.

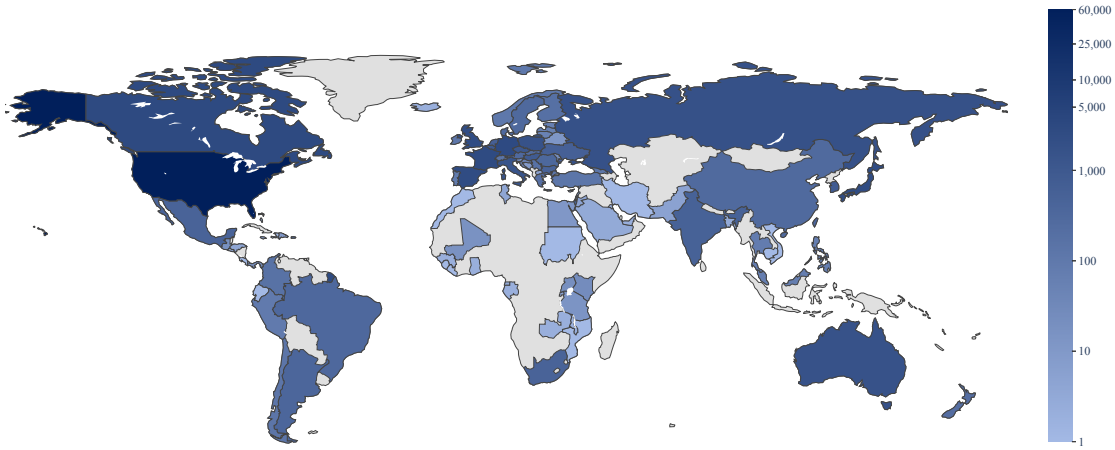
WONG, H.-H., A. JESSUP, A. SERTKAYA, A. BIRKENBACH, A. BERLIND, AND J. EYRAUD (2014): "Examination of clinical trial costs and barriers for drug development final," *Reports Submitted to Office of the Assistant Secretary for Planning and Evaluation*.

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

WOUTERS, O. J., M. MCKEE, AND J. LUYTEN (2020): "Estimated research and development investment needed to bring a new medicine to market, 2009-2018," *Journal of the American Medical Association*, 323, 844–853.

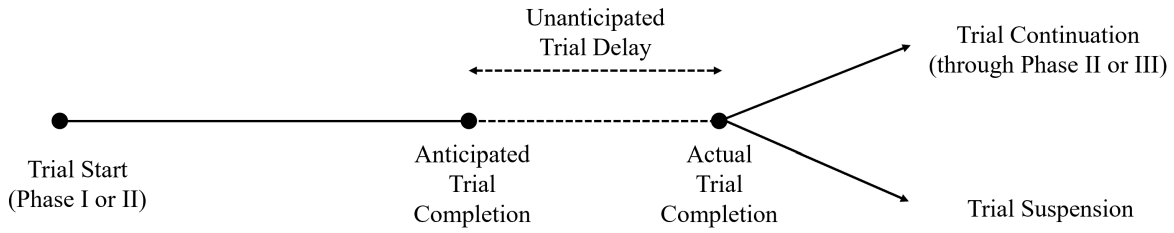
Figures

Figure 1: Geographic Distribution of Trial Sites



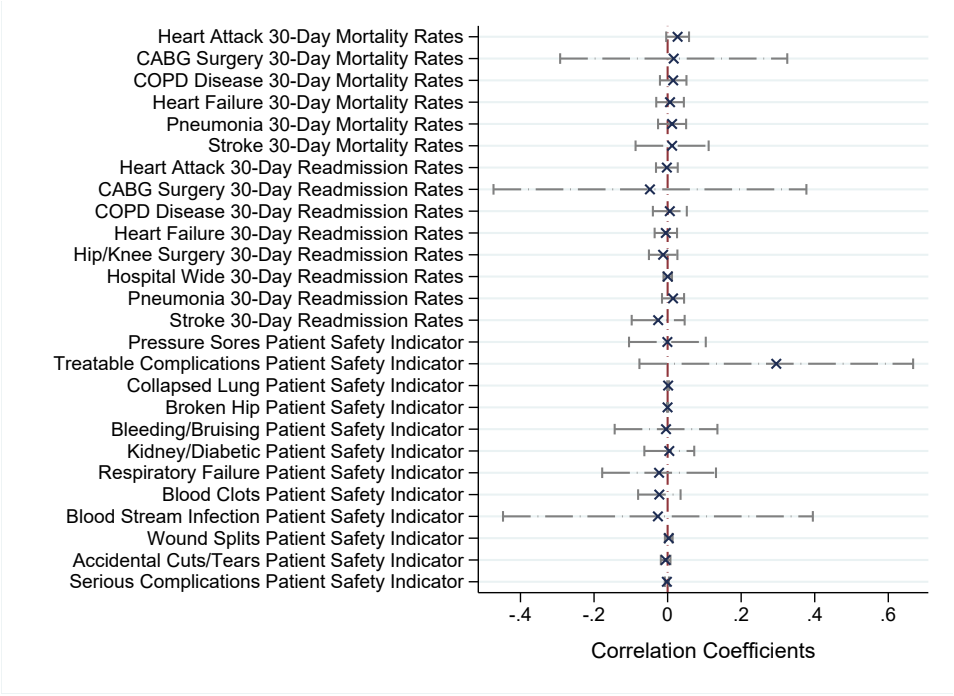
This figure exhibits the trial-site geographic distribution of clinical trials included in the analysis sample. For each country, we calculate the natural logarithm of the number of trial sites.

Figure 2: Project Timeline



This figure visualizes the project and decision timeline with respect to the key aspects of the empirical design, the relationship between the trial continuation decision and unanticipated delays in the completion of the previous trial phase.

Figure 3: Correlation between Trial Site Congestion and Hospital Care Quality

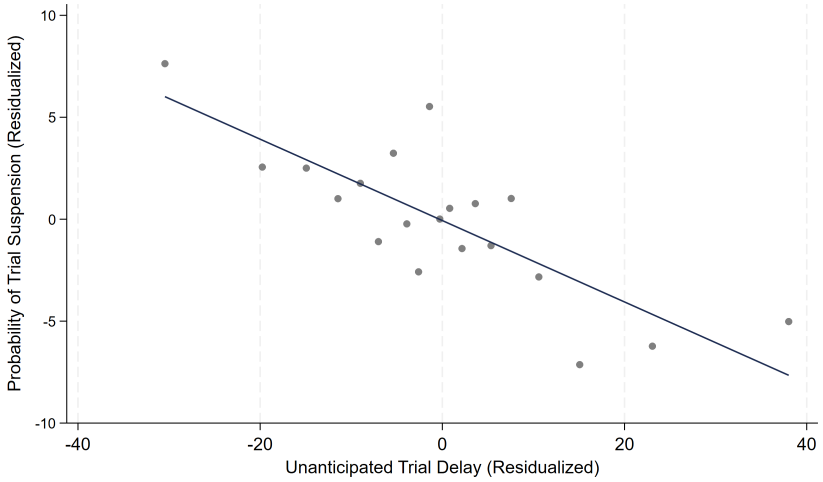


This figure exhibits the correlation coefficients between the trial site congestion measure and the average hospital care quality measures in a zip code. The correlation coefficients are estimated by the regression

$$Q_{zt} = \beta G_{zt} + \gamma_z + \tau_t$$

where Q_{zt} is the average quality measure of hospital care at zip code z in year t. G_{zt} is the congestion measure at zip code z in year t. γ_z and τ_t are zip code and year fixed effects. β represents the correlation coefficients. The y-axis denotes the names of various quality measures. All standard errors are two-way clustered at the zip code and year levels. Capped spikes represent 95% confidence intervals.

Figure 4: Delays in Trial Completion and Project Continuation Versus Suspension



This figure shows a binscatter plot visualizing the relationship between unanticipated clinical trial delays and firms' subsequent decision to suspend versus advance the trial through the next trial phase. Both the delay and suspension indicator variables are residualized with respect to the control variables and fixed effects in the final column of Table 2.

Tables

Table 1: Summary Statistics

This table reports summary statistics. The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records.

	N	Mean	SD	P25	Median	P75
Suspension	11,228	0.31	0.46	0.00	0.00	1.00
Trial Duration (in months)	11,228	35.65	29.45	12.00	29.00	51.00
Anticipated Trial Duration	9,948	22.83	17.79	9.00	20.00	32.00
Delay in Trial Completion	9,948	11.78	19.19	0.00	3.00	19.00
Number of Clinical Trial Sites	11,179	13.65	40.62	1.00	3.00	13.00
Trial Participants (in hundreds)	11,195	1.00	1.62	0.25	0.49	1.06
Clinical Trial Phase (I or II)	11,228	1.59	0.49	1.00	2.00	2.00
Number of Drug Projects	10,998	136.14	179.07	11.00	41.00	230.00
Number of Competing Drug Projects (in logs)	11,182	5.16	1.34	4.39	5.33	6.01
IV for Delay based on Congestion Measures at Zipcode Level	9,948	5.28	10.04	0.00	0.93	4.19
Foreign Trial	11,228	0.35	0.48	0.00	0.00	1.00
$\Delta FX \mid$ *Foreign Trial*	3,918	-1.40	6.35	-4.60	-0.92	1.58

Table 2: Delays in Trial Completion and Project Continuation Versus Suspension

This table reports the OLS estimation results of Equation (1). The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records. The dependent variable, $Suspension_i$, is an indicator variable (multiplied by 100 for ease of exposition) for whether project i was suspended after completing Phase I or II clinical trials. The independent variable, $Delay in Trial Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD category (taking logs). Standard errors are two-way clustered at the ICD category and quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: Suspension			
	(1)	(2)	(3)	(4)
Delay in Trial Completion	-0.229^{***}	-0.210^{***}	-0.206^{***}	-0.200^{***}
	(-5.77)	(-6.92)	(-5.70)	(-5.41)
Controls	N	N	N	Y
Year × Quarter FE	Y	Y	Y	Y
Firm FE	N	Y	Y	Y
ICD FE	N	N	Y	Y
Trial Phase FE	N	N	Y	Y
Observations	9,912	9,433	9,397	9,161
Adj. R-squared	0.1558	0.2711	0.3085	0.3098

Table 3: Delays in Trial Completion and Project Continuation Versus Suspension: Trial Site Congestion as an Instrument for Delay

This table reports the IV estimation results of Equation (1), using trial site congestion as an instrument for delay in clinical trial completion. The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records. The second-stage dependent variable, $Suspension_i$, is an indicator variable (multiplied by 100 for ease of exposition) for whether project i was suspended after completing Phase I or II clinical trials. The instrument, $TrialSiteCongestion_i$, is the normalized change in the number of participants in the universe of trials on ClinicalTrials.gov between trial start and end date, taking the maximum across all trial sites of a given trial in the sample. The second-stage independent variable, $Delay in Trial Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD category (taking logs). Standard errors are two-way clustered at the ICD category and quarter levels.We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	DV: Delay First Stage		**DV: Suspension** Second Stage	
	(1)	(2)	(3)	(4)
Trial Site Congestion	0.360^{***}	0.358^{***}		
	(4.26)	(4.37)		
Delay in Trial Completion			-0.337^{**}	-0.530^{***}
			(-2.33)	(-2.97)
Controls	N	Y	N	Y
Year × Quarter FE	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y
F-stat	18.13	19.13	–	–
Observations	9,397	9,161	9,397	9,161

Table 4: Exchange Rate Variation and Project Continuation Versus Suspension

This table reports the OLS estimation results of Equation (3). The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records. The dependent variable, $Suspension_i$, is an indicator variable (multiplied by 100 for ease of exposition) for whether project i was suspended after completing Phase I or II clinical trials. The independent variable, ΔFX_i, is the average exchange rate change for clinical trial i defined in Equation (2). $Trial\,Duration_i$, is the length of time (in months) for which it takes project i to complete Phase I or II clinical trials. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD category (taking logs). Standard errors are two-way clustered at the ICD category and quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: Suspension			
	(1)	(2)	(3)	(4)
ΔFX	-0.311^{**}	-0.270^{*}		
	(-2.04)	(-1.76)		
$\Delta FX \times$ *Foreign Trial*			-0.251^{*}	-0.249^{*}
			(-1.74)	(-1.74)
Trial Duration	-0.200^{***}	-0.208^{***}	-0.203^{***}	-0.189^{***}
	(-4.20)	(-5.40)	(-6.79)	(-5.45)
Controls	N	Y	N	Y
Year × Quarter FE	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y
Observations	3,587	3,524	10,521	10,284
Adj. R-squared	0.2995	0.2990	0.3198	0.3198

Table 5: Firm- and Project Characteristics-Related Channels

This table investigates underlying firm- and project characteristics-related channels. The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records. The dependent variable, $Suspension_i$, is an indicator variable (multiplied by 100 for ease of exposition) for whether project i was suspended after completing Phase I or II clinical trials. The independent variable, $Delay\,in\,Trial\,Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). See Section 5.4 for further details. Standard errors are two-way clustered at the ICD category and quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: Suspension				
	Has Other Inv. Opportunities	Versus Expectations	Versus Fin. Constr.	Phase I Only	Phase II Only
	(1)	(2)	(3)	(4)	(5)
Delay in Trial Completion	-0.208^{***}	-0.224^{***}	-0.196^{***}	-0.152^{**}	-0.199^{***}
	(-5.15)	(-5.87)	(-4.08)	(-2.63)	(-5.08)
Anticipated Trial Duration		-0.112^{**}			
		(-2.57)			
Constrained (WW)			0.238		
			(0.06)		
Controls	Y	Y	Y	Y	Y
Year × Quarter FE	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y	Y
Observations	7,052	9,161	4,746	3,356	5,414
Adj. R-squared	0.2725	0.3106	0.2701	0.3653	0.3292

Table 6: Real Options Channel

This table investigates real options as an underlying channel. Panel A examines the relationship between the log of expected drug sales (Million U.S. dollars adjusted to 2017 by GDP delfators) after obtaining the FDA approval and $Delay in Trial Completion_i$, the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Columns (1) to (4) use four different sales measures (all take logs), including the expected first-year sales after launch, the sum of first three-year sales, the sum of first five-year sales, and the average annual sales. Panel B examines the relationship between the percentage of adverse reactions among trial participants for project i that already completed Phase I or II clinical trials and $Delay in Trial Completion_i$. Columns (1) to (4) use four different dependent variables, including the total percentage of adverse reaction reported in the completed trial (a participant might experience multiple adverse events such as headache, nausea, fatigue) relative to the total number of enrolled participants, the maximum percentage of adverse reaction among various adverse symptoms, the total percentage of severe adverse reaction (life-threatening reaction, for example, acute myelogenous leukemia, basal cell carcinoma, brain hemorrhage, etc.), and the maximum percentage of severe adverse reaction among various adverse symptoms. We adopt the specification following Equation (4.1). Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD category (taking logs). Standard errors are two-way clustered at the ICD category and quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Expected Drug Sales

	First-year Sales	**Three-year Sales**	**Five-year Sales**	**Average Annual Sales**
	(1)	(2)	(3)	(4)
Delay in Trial Completion	−0.000	−0.000	−0.000	0.000
	(−0.12)	(−0.58)	(−0.11)	(0.41)
Controls	Y	Y	Y	Y
Year × Quarter FE	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y
Observations	3,949	3,847	3,609	3,949
Adj. R-squared	0.4370	0.5591	0.5577	0.5722

Panel B: Adverse Reaction in Trials

	Total Reaction	**Max Reaction**	**Total Severe Reaction**	**Max Severe Reaction**
	(1)	(2)	(3)	(4)
Delay in Trial Completion	0.726***	0.132***	0.001	0.003
	(5.46)	(6.12)	(0.40)	(0.86)
Controls	Y	Y	Y	Y
Year × Quarter FE	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y
Observations	9,159	9,159	9,159	9,159
Adj. R-squared	0.2178	0.1891	−0.0551	−0.0560

Table 7: CEO-Related Channels

This table investigates underlying CEO-related channels. The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records. The dependent variable, $Suspension_i$, is an indicator variable (multiplied by 100 for ease of exposition) for whether project i was suspended after completing Phase I or II clinical trials. The independent variable, $Delay in Trial Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. *CEO Delta* is the natural logarithm of delta as in Coles et al. (2006), standardized to have a mean of zero and standard deviation of one for expositional purposes. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). See Section 5.4 for further details. Standard errors are two-way clustered at the ICD category and quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: Suspension		
	Full CEO Sample		Same-CEO Subsample
	(1)	(2)	(3)
Delay in Trial Completion	-0.163^{***}	-0.111^{*}	-0.411^{***}
	(-2.97)	(-1.82)	(-3.13)
CEO Delta	0.373		
	(0.27)		
Delay × *CEO Delta*	-0.081^{**}		
	(-2.36)		
Same CEO		6.940^{***}	
		(2.64)	
Delay × *Same CEO*		-0.224^{*}	
		(-1.93)	
Young CEO			-10.070^{**}
			(-2.50)
Delay × *Young CEO*			0.273
			(1.33)
Controls	Y	Y	Y
Year × Quarter FE	Y	Y	Y
Firm FE	Y	Y	Y
ICD FE	Y	Y	Y
Trial Phase FE	Y	Y	Y
Observations	2,913	3,400	2,075
Adj. R-squared	0.2587	0.2691	0.3085

Table 8: Effects on Other Firm Investments

This table investigates the effects of delay-induced R&D investment distortions on new drug project initiation, using trial site congestion as an instrument for delay in clinical trial completion. The unit of observation is a drug project. The sample contains all orphan-drug-designated projects that completed Phase I or II trials with detailed clinical trials records. The second-stage dependent variable is the total number of new projects (in logs) initiated by company within one year of completing the focal project (i.e., the project included in our main sample). As throughout out analysis, projects are defined as a drug by indication, and we include as newly initiated projects those that are initiated with development status "Discovery" or "Preclinical." Columns (3) to (4) (Columns (5) to (6)) calculate the total number of new projects that share (do not share) the same ICD indication as the focal project. The instrument, $TrialSiteCongestion_i$, is the normalized change in the number of participants in the universe of trials on ClinicalTrials.gov between trial start and end date, taking the maximum across all trial sites of a given trial in the sample. The second-stage independent variable, $Delay in Trial Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD category (taking logs). Standard errors are two-way clustered at the ICD category and quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: Suspension					
	Second Stage					
	All Projects		Same-ICD Projects		Different-ICD Projects	
	(1)	(2)	(3)	(4)	(5)	(6)
Delay in Trial Completion	-0.011^{***}	-0.011^{***}	0.002	0.002	-0.011^{***}	-0.011^{***}
	(-3.33)	(-3.17)	(1.01)	(0.86)	(-3.30)	(-3.18)
Controls	N	Y	N	Y	N	Y
Year × Quarter FE	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y	Y	Y
First-stage F-stat	26.56	25.46	26.56	25.46	26.56	25.46
Observations	7,838	7,636	7,838	7,636	7,838	7,636

Table 9: Effects on New Drug Launches With No or Few Existing Medications (Orphan Drugs)

This table investigates the effects of delay-induced R&D investment distortions on the probability of new drug launches with no or few existing medications (orphan drugs), using trial site congestion as an instrument for delay in clinical trial completion. The unit of observation is a drug project. The sample contains all orphan-drug-designated projects that completed Phase I or II trials with detailed clinical trials records. The second-stage dependent variable, $FDAApproval_i$, is an indicator variable (multiplied by 100 for ease of exposition) for whether a drug is eventually approved by the FDA. The instrument, $TrialSiteCongestion_i$, is the normalized change in the number of participants in the universe of trials on ClinicalTrials.gov between trial start and end date, taking the maximum across all trial sites of a given trial in the sample. The second-stage independent variable, $Delay in Trial Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD category (taking logs). Standard errors are two-way clustered at the ICD category and quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	DV: Delay First Stage	**DV: FDA Approval** Second Stage
	(1)	(2)
Trial Site Congestion	0.551***	
	(5.17)	
Delay in Trial Completion		0.249***
		(2.65)
Controls	Y	Y
Year × Quarter FE	Y	Y
Firm FE	Y	Y
ICD FE	Y	Y
Trial Phase FE	Y	Y
F-stat	26.75	–
Observations	1,859	1,859

Table 10: Effects on Adverse Events Related to Approved Drugs

This table investigates the effects of delay-induced R&D investment distortions on adverse events related to approved drugs, using trial site congestion as an instrument for delay in clinical trial completion. The unit of observation is a drug project. The sample contains all approved projects that completed Phase I or II trials with detailed clinical trials records. The second-stage dependent variables are the count of adverse events reported by the FDA (taking logs) within the first and the first three years since drug launch, respectively. The instrument, $TrialSiteCongestion_i$, is the normalized change in the number of participants in the universe of trials on ClinicalTrials.gov between trial start and end date, taking the maximum across all trial sites of a given trial in the sample. The second-stage independent variable, $Delay in Trial Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Please see Section 7.2.1 for further details on the econometric specification and estimation. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). We report t−statistics, based on standard errors accounting for the generated regressors problem by bootstrapping, in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	DV: Delay First Stage	**DV:** log(**Adverse Events in 1 Year**) Second Stage	**DV:** log(**Adverse Events in 3 Years**) Second Stage
	(1)	(2)	(3)
Trial Site Congestion	0.903***		
	(5.25)		
Delay in Trial Completion		0.014	0.014
		(0.57)	(0.58)
Inverse Mills Ratio	−1.572	1.101	1.467
	(−0.12)	(0.54)	(0.66)
Controls	Y	Y	Y
Year × Quarter FE	Y	Y	Y
Firm FE	Y	Y	Y
ICD FE	Y	Y	Y
Trial Phase FE	Y	Y	Y
Observations	438	438	438

Online Appendix

Excess Commitment in R&D

Marius Guenzel
Wharton

Tong Liu
MIT Sloan

Not For Publication

I Theoretical Framework

This section discusses a simplified theoretical model to motivate the empirical findings. Our model extends Li et al. (2023) by including unexpected delay in project development and differential ability of the CEO in running the project. We discuss two different scenarios: a frictionless scenario as a benchmark and a scenario with asymmetric information about managers' ability.

The model is in continuous time. There is a pharmaceutical firm developing the drug project i. In order to launch the drug project in the market, the project has to go through Phase I and then Phase II clinical trials. Only favorable trial results allow for the project to continue to the next phase. At $t = 0$, the firm begins Phase I clinical trials. For a project in Phase $j \in \{I, II\}$, the firm pays a flow R&D cost c_j at any time t so that with instantaneous probability λ_j the project completes the current phase of trials. Since the time to complete a stage of clinical trials follow the Poisson process, we know the anticipated duration to complete Phase $j \in \{I, II\}$ would be $\frac{1}{\lambda_j}$. The unexpected delay in a stage of trial can be defined as $D_j = t - \frac{1}{\lambda_j}$ where t is the actual time to complete Phase j trials. The cumulative density function of t follows $G(x) = 1 - e^{-\lambda_1 t}$. After the project completes both phases, it receives a payoff $W \in [W_{\min}, W_{\max}]$ with mean $\overline{W}$ and the cumulative density function $F(W)$. Suppose W will be revealed after Phase I is completed and can be observed publicly. All agents in the model are risk neutral with a discount rate r.

The firm hires a manager to oversee the development of the drug project. Assume managers can be of two types: good (g) or bad (b). The project, denoted by its final payoff W, is randomly matched with managers. The prior for a good-type manager in the market is $p \in (0,1)$. The ability of a manager governs the instantaneous probability of the project completes the current phase of trials. Specifically, if a project in Phase $j \in \{I, II\}$ is run by a manager of type $s \in \{g, b\}$, the instantaneous probability of completing the phase is λ_j^s with $\lambda_j^g > \lambda_j^b$. Therefore, a good manager is expected to complete the trial faster.

Upon the completion of Phase I trial, the manager has a choice to continue or suspend the project. To simplify the analysis, we assume that the manager would leave the firm after suspending the project and re-enter the labor market. If the manager chooses to continue the project, she receives a compensation package $M = a + bV$ which is linear in the market value of the project V at time t when Phase I trial is completed. a represents her base wage and b represents the sensitivity of her compensation to the project value. If the manager chooses to suspend the project, she has to re-enter the labor market and her

potential compensation becomes $M' = a + b[V - N(p_b)] + \varepsilon$, where $N(p_b)$ represents the search costs in the labor market. $N(p_b)$ is an increasing function of the market belief of the bad type CEO, motivated by the empirical evidence on managerial labor markets.[18] ε is the idiosyncratic component in the change of her base wage. The manager observes ε before deciding to continue or suspend the project and the CDF of ε is $K(\cdot)$.

I.A Frictionless Benchmark

Consider a project run by a manager of type $s \in \{g, b\}$ that completes Phase I trials and is entering Phase II at time t. The final payoff of the project W is observed. Therefore, the Bellman equation for the project's expected value follows

$$rV_2^s = -c_2 + \lambda_2^s(W - V_2^s)$$

in which rV_2^s represents the flow value in the project, and it has two components. First, it spends the R&D costs c_2 for Phase II trials. Second, with probability λ_2^s, the project completes Phase II and receives its final payoff W. So $W - V_2^s$ is the expected jump in payoff. Rearranging this equation yields

$$V_2^s = \frac{\lambda_2^s W - c_2}{\lambda_2^s + r}. \tag{8}$$

Moving backward, the Bellman equation of the project's value in Phase I follows

$$rV_1^s = -c_1 + \lambda_1^s(\overline{V}_2^s - V_1^s)$$

[18] Prior work documents a significant cost to managers of losing the CEO position. Most recently, Cziraki and Jenter (2022) report that "unattached [outside] managers are more frequently chosen by firms with low stock returns, low ROA, low sales growth, low market-to-book, and high leverage." Fee et al. (2018) find that "new positions [of ex-CEOs] tend to be substantially inferior." Furthermore, beyond CEO job loss constituting a negative career event in general, Fee et al. (2018) find evidence that individuals who have sent more negative about their managerial ability "fare significantly worse in the retread labor market." Examining labor market outcomes of ex-CEOs *and* ex-non-CEOs jointly provides further insight. Similar to above, also for executives in general, "on average their new positions are significantly inferior" (Fee and Hadlock, 2004). Additionally, the statistics in Fee and Hadlock (2004) indicate that following job loss and in comparison to CEOs, non-CEOs—who Kaplan and Sorensen (2021) find to have lower general ability than CEOs—secure new employment opportunities that are even less favorable.

where $\overline{V}_2^s = \frac{\lambda_2^s \overline{W} - c_2}{r + \lambda_2^s}$ since W is unknown in Phase I and the market takes an expected value of W. Therefore,

$$V_1^s = \frac{\lambda_1^s \overline{V}_2^s - c_1}{\lambda_1^s + r}.$$

To have the incentive constraint be satisfied, we impose the following parametric assumption $\overline{W} \geq \frac{c_1}{\lambda_1^b} \frac{r+\lambda_2^b}{\lambda_2^b} + \frac{c_2}{\lambda_2^b}$ so that $V_1^g > V_1^b \geq 0$.

How does delay affect the CEO's continuation decision?

Suppose at time t a project W run by a manager s completes its Phase I trials with an unexpected delay $D = t - \frac{1}{\lambda_1^s}$. The ability of the manager is perfectly observed in the market. According to Equation (8), the expected value of the project at the moment is V_2^s with $s \in \{g, b\}$. The manager weighs her compensation of continuing the project, M, against the compensation if suspended, M'. We have the following conclusion:

The suspension probability of the project run by a good-type manager is

$$Prob(Suspended \mid D, W, s = g) = 1 - K(bN(0)) \tag{9}$$

where $K(\cdot)$ is the cumulative density function of ε, b is the sensitivity of the manager's compensation to the project value, and $N(0)$ indicates that the market knows she is a good-type manager. Similarly, the suspension probability of the project run by a bad-type manager is

$$Prob(Suspended \mid D, W, s = b) = 1 - K(bN(1)) \tag{10}$$

where $N(1)$ indicates that the market is aware that the project is run by a bad-type manager.

Notice in Equations (9) and (10), the project suspension probabilities are independent of D or t for $s \in \{g, b\}$. So the benchmark suspension probability can be summarized in the following prediction:

Prediction: *In the benchmark model with full information about managers' ability, unexpected delays in the already-completed trial stage do not affect the manager's propensity to suspend the project subsequently.*

In addition, we analyze the optimal suspension decision from the social planner's perspective. At time t for a project that has already completed Phase I trials, it is optimal to suspend any negative NPV project, i.e., $V_2^s < 0$. This implies that the social planner would suspend a project as long as $W < \frac{c_2}{\lambda_2^s}$. Though this suspension decision rule differs from the manager's in Equations (9) and (10),[19] the project suspension probability under the social planner is also independent of the unexpected delay of the trial. So the unexpected delay in the already-completed trials does not affect the suspension probability in the follow-up stage of clinical trials.

I.B A model with asymmetric information

Suppose the market cannot observe managers' ability. There would only exist a pooling equilibrium because the bad-type manager has an incentive to mimic the good one and there is no cost for them to do so. At time t, conditional on the project completed Phase I, the market belief of a bad-type manager would be

$$Pr(b \mid \text{complete Phase I at time } t) = \frac{1-p}{(1-p)+p\frac{\lambda_1^g}{\lambda_1^b}e^{-(\lambda_1^g-\lambda_1^b)t}} = A(t).$$

It is easy to verify that $A(t)$ is an increasing function in t.

The expected market value of a project W that just completed Phase I and is in Phase II at time t satisfies the Bellman equation

$$rV_2 = -c_2 + A(t)\lambda_2^b(W-V_2) + [1-A(t)]\lambda_2^g(W-V_2)$$

where $A(t)\lambda_2^b$ represents the probability of completing Phase II conditional on a bad-type manager and $[1-A(t)]\lambda_2^g$ is the probability for a good type. It implies

$$V_2 = \frac{\{A(t)\lambda_2^b + [1-A(t)]\lambda_2^g\}W - c_2}{r + A(t)\lambda_2^b + [1-A(t)]\lambda_2^g}.$$

Analyzing managers' decision to whether suspend the project, we compare their compensation under the asymmetric information scenario. Given the project W completing its Phase I with an unexpected

[19] The difference between the manager's and the social planner's decision rule is mechanical under the assumption of managers' linear compensation structure in our stylized model. We would rather not call it distortion.

delay $D = t - \left[p\frac{1}{\lambda_1^g} + (1-p)\frac{1}{\lambda_1^b} \right]$ relative to the initial anticipated duration at $t = 0$, the compensation for a manager to continue the project would be

$$M(D,W) = a + bV_2.$$

If the manager decides to suspend the project, her compensation becomes

$$M'(D,W) = a + b(V_2 - N(A(t))) + \varepsilon.$$

The condition for the manager to suspend the project is $M'(D,W) > M(D,W)$, implying $\varepsilon > bN(A(t))$. The probability of suspending the project is

$$Pr(Suspended \mid D,W) = 1 - K(b \times N(A(\mathrm{t}))) = 1 - K(b \times N(A\left(D + p\frac{1}{\lambda_1^g} + (1-p)\frac{1}{\lambda_1^b}\right))).$$

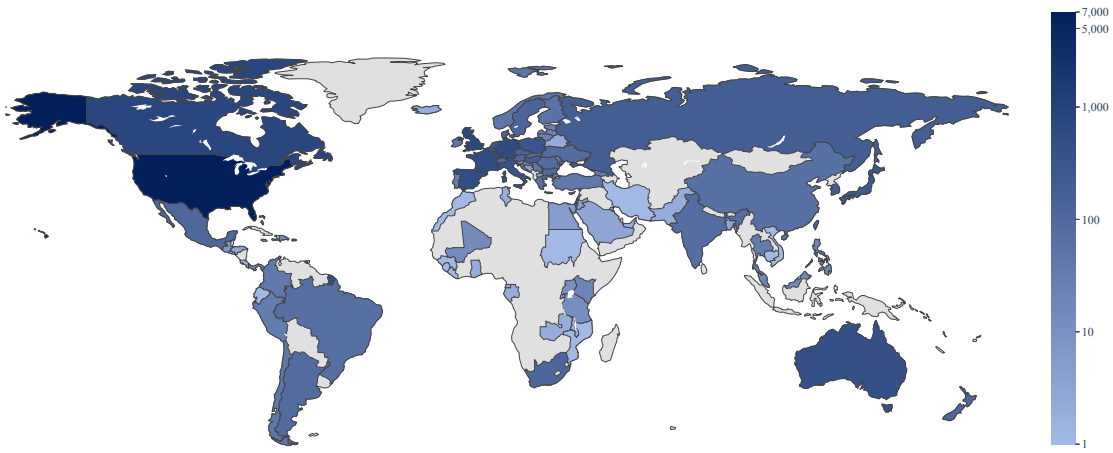
Since $K(\cdot)$ and $N(\cdot)$ are increasing functions in their respective arguments and $A(\cdot)$ is an increasing function in D, we have the following statements:

Prediction: *In a model with asymmetric information about managers' ability, the suspension probability in the follow-up stage of clinical trials is (1) decreasing in the unexpected delay of the already-completed trial stage D; (2) decreasing in the sensitivity of the CEO's compensation w.r.t. the firm's stock price given a level of unexpected delay in the already-completed trial stage.*

II Supplementary Results

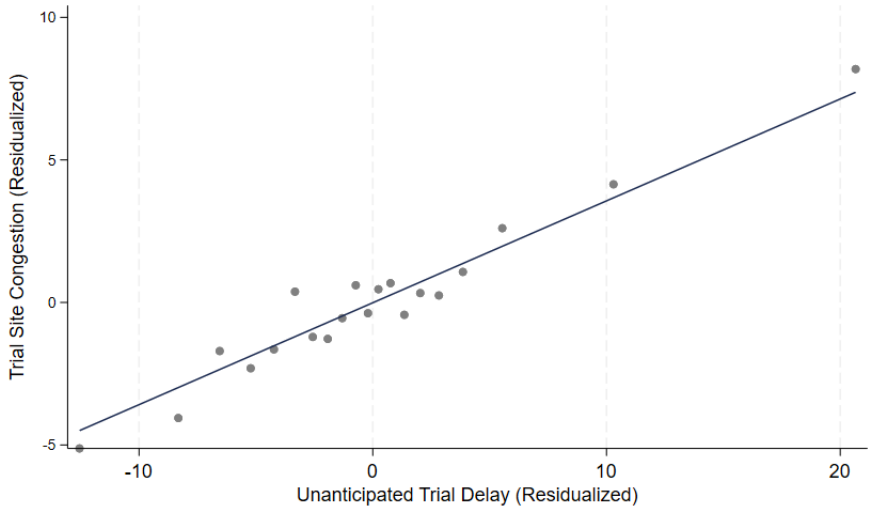
II.A Figures

Figure OA.1: Geographic Distribution of Clinical Trials



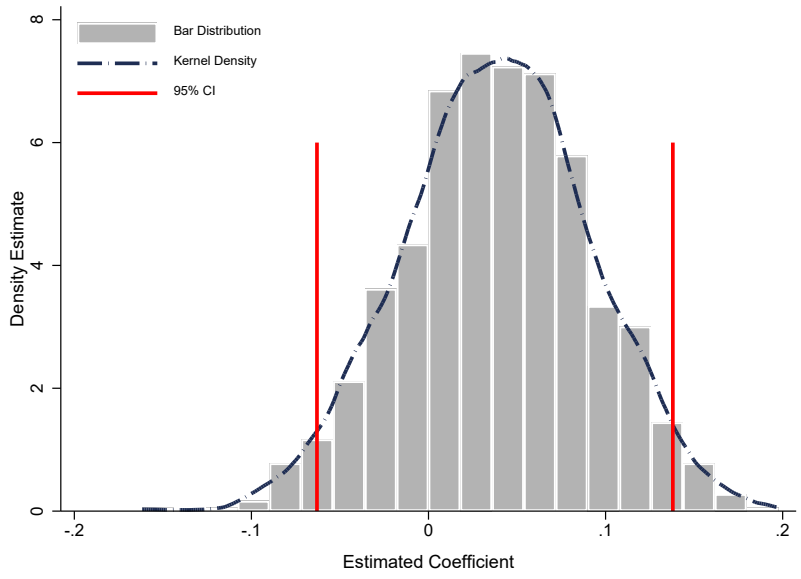
This figure exhibits the geographic distribution of clinical trials in our analysis sample. For each country, we calculate the natural logarithm of the number of trials.

Figure OA.2: First Stage

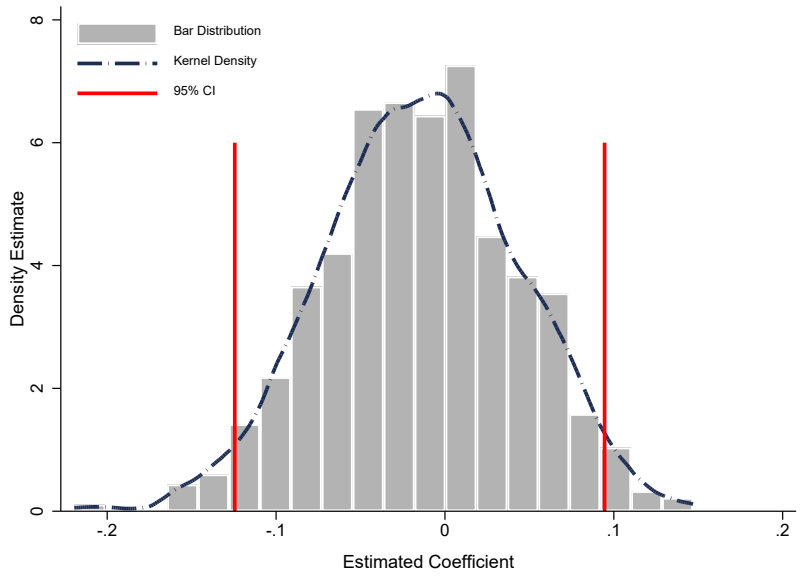


This figure shows a binscatter plot visualizing the relationship between unanticipated clinical trial delays and trial site congestion (IVs). Both the delay and trial site congestion are residualized with respect to the control variables and fixed effects in column (2) of Table 3.

Figure OA.3: Placebo Test for Variation in Exchange Rates



(a) Distribution of Estimated Coefficient of ΔFX without Controls



(b) Distribution of Estimated Coefficient of ΔFX with Controls

This figure exhibits the distribution of the estimated coefficient for ΔFX in Equation (2). In each estimation, we randomly assign countries to the clinical trials in the sample. The distribution of the estimated coefficients is derived by repeating the estimation 1,000 times. Panel A exhibits the estimation without controls. Panel B exhibits the estimation with controls.

II.B Tables

Table OA.1: Determinants of Anticipated Trial Duration

This table examines the determinants of a clinical trial's anticipated trial duration by gradually adding control variables and fixed effects in Table 2. The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records. The dependent variable, $Anticipated\,in\,Duration_i$, is the difference between trial start month and expected trial completion month of project i. The independent variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD category (taking logs). Fixed effects are indicated in the bottom rows. Standard errors are two-way clustered at the ICD category and quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: Anticipated Duration (Month)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Num of Trial Sites)					1.222***	0.949***	0.950***	0.968***
					(3.80)	(2.85)	(2.89)	(2.94)
Num of Patient Enrolled (×100)						0.504**	0.501**	0.489**
						(2.16)	(2.16)	(2.16)
log(Num of Company's Projects)							−0.173	−0.100
							(−0.29)	(−0.17)
log(Num of Competing Projects)								−3.939***
								(−3.32)
Year × Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	N	Y	Y	Y	Y	Y	Y	Y
ICD FE	N	N	Y	Y	Y	Y	Y	Y
Trial Phase FE	N	N	N	Y	Y	Y	Y	Y
Observations	9,654	9,193	9,161	9,161	9,161	9,161	9,161	9,161
Adj. R-squared	0.130	0.362	0.456	0.486	0.492	0.494	0.493	0.495

Table OA.2: Delays in Trial Completion and Project Continuation Versus Suspension: Further Evidence on Financial Constraints

This table further explores the role of financial constraints for explaining the relationship between trial suspension and delay in completion of the preceding trial phase, adding to the evidence in the final column in Panel A of Table 5. The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records. The dependent variable, $Suspension_i$, is an indicator variable for whether project i was suspended after completing Phase I or II clinical trials. The independent variable, $Delay\,in\,Trial\,Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). See Section 5.4 for further details. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: Suspension					
	(1)	(2)	(3)	(4)	(5)	(6)
Delay in Trial Completion	−0.196***	−0.179***	−0.197***	−0.197***	−0.195***	−0.193***
	(−4.08)	(−3.91)	(−4.10)	(−4.84)	(−4.69)	(−4.65)
Constrained (WW)	0.238					
	(0.06)					
Constrained (HP)		13.722***				
		(3.20)				
Constrained (KZ)			−0.766			
			(−0.27)			
Constrained–Trial Start (WW)				−0.905		
				(−0.25)		
Constrained–Trial Start (HP)					4.572	
					(1.00)	
Constrained–Trial Start (KZ)						3.429
						(1.18)
Controls	Y	Y	Y	Y	Y	Y
Year × Quarter FE	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y	Y	Y
Observations	4,746	4,855	4,759	4,987	5,174	4,996
Adj. R-squared	0.2701	0.2736	0.2714	0.2603	0.2646	0.2616

Table OA.3: FDA Clinical Holds and Trial Duration

This table examines if the trial delays are mechanically caused by FDA's clinical holds. We adopt a specification similar to Column (4) of Table 2 except that we use *Trial Duration*, *Delay in Trial Completion*, and *Trial Site Conjestion* as the dependent variables. The independent variable is an indicator variable for whether project *i* experienced any clinical hold in Phase I or II clinical trials that already completed. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

DV:	**Trial Duration**	**Delay in Completion**	**Instrument for Delay**
	(1)	(2)	(3)
Had Clinical Hold	−6.032	1.122	5.465
	(−1.65)	(0.32)	(1.30)
Controls	Y	Y	Y
Year × Quarter FE	Y	Y	Y
Firm FE	Y	Y	Y
ICD FE	Y	Y	Y
Trial Phase FE	Y	Y	Y
Observations	9,257	9,161	9,257
Adj. R-squared	0.5745	0.3187	0.4517

Table OA.4: Correlation between Delay Measures and Number of Doses in a Trial

This table examines the correlation between trial delay measures and number of doses tested in a trial. We use *Delay in Trial Completion* and *Trial Site Conjestion* as the dependent variables. The independent variable,*Num of Doses in a Trial*, counts the number of drug doses tested in a trial. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Delay in Trial Completion	**Trial Site Congestion**
	(1)	(2)
Num of Doses in a Trial	0.389	0.009
	(1.42)	(0.09)
Controls	Y	Y
Year × Quarter FE	Y	Y
Firm FE	Y	Y
ICD FE	Y	Y
Trial Phase FE	Y	Y
Observations	3,853	3,905
Adj. R-squared	0.3561	0.5241

Table OA.5: Robustness: Control for the Number of Doses in a Trial

This table replicates Table 3 except that we include *Num of Doses in a Trial* as an additional dependent variables. Other details are the same as in Table 3.

	DV: Delay First Stage		**DV: Suspension** Second Stage	
	(1)	(2)	(3)	(4)
Trial Site Congestion	0.425***	0.434***		
	(4.43)	(4.76)		
Delay in Trial Completion			−0.557***	−0.619***
			(−3.19)	(−3.07)
Num of Doses in a Trial	0.411	0.386	−0.189	−0.137
	(1.66)	(1.54)	(−0.80)	(−0.55)
Controls	N	Y	N	Y
Year × Quarter FE	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y
F-stat	19.60	22.64	–	–
Observations	3,947	3,853	3,947	3,853

Table OA.6: Correlation between Delay Measures and Drug Quality Proxies

This table examines the correlation between trial delay measures and drug quality proxies. We use *Delay in Trial Completion* in columns (1) to (4) and *Trial Site Conjestion* in columns (5) to (8) as the dependent variables. Two drug quality proxies, *Drug Likelihood of Approval* and *Phase Transition Success Rate (PTSR)*, measure the predicted probabilities to obtain the final FDA approval and to successfully complete the current phase of trial and move forward to the next one for a drug candidate. The drug quality proxies are sourced from the GlobalData's Likelihood of Approval (LoA) database. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Delay in Trial Completion				**Trial Site Congestion**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Drug Likelihood of Approval (LoA)	0.009	0.053			−0.009	0.013		
	(0.07)	(0.40)			(−0.15)	(0.32)		
Phase Transition Success Rate (PTSR)			0.050	0.023			0.056	0.034
			(1.05)	(0.52)			(1.40)	(1.31)
Controls	N	Y	N	Y	N	Y	N	Y
Year × Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,436	2,388	2,436	2,388	2,467	2,417	2,467	2,417
Adj. R-squared	0.3712	0.3777	0.3712	0.3777	0.3443	0.5036	0.3443	0.5035

Table OA.7: Positive Vs. Negative Delay Robustness

This table implements robustness based on whether the gap between realized and anticipated trial phase end date is positive or negative. The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records. The dependent variable, $Suspension_i$, is an indicator variable for whether project i was suspended after completing Phase I or II clinical trials. The independent variable, $Delay in Trial Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials, split by whether this gap takes a positive or negative value. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). See Section 5.4 for further details. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: Suspension			
	Full Sample		Delay ≥ 0 Sample	
	(1)	(2)	(3)	(4)
Delay in Trial Completion$^+$	-0.229^{***}	-0.223^{***}	-0.219^{***}	-0.211^{***}
	(-6.31)	(-6.00)	(-5.45)	(-5.10)
Delay in Trial Completion$^-$	0.079	0.086		
	(0.43)	(0.48)		
Controls	N	Y	N	Y
Year $\times$ Quarter FE	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y
Observations	9,397	9,161	7,954	7,757
Adj. R-squared	0.3087	0.3100	0.3178	0.3198

Table OA.8: Additional Miscellaneous Robustness Tests

This table implements further miscellaneous robustness tests. The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records. The dependent variable, $Suspension_i$, is an indicator variable for whether project i was suspended after completing Phase I or II clinical trials. The independent variable, $Delay in Trial Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). See Section 5.4 for further details. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: Suspension		
	Interacted FE	US Trials	Non-US Trials
	(1)	(2)	(3)
Delay in Trial Completion	-0.211^{***}	-0.226^{***}	-0.169^{***}
	(-4.81)	(-5.12)	(-3.13)
Controls	Y	Y	Y
Year × Quarter FE	N	Y	Y
Firm FE	Y	Y	Y
ICD FE	N	Y	Y
Year × Quarter × ICD FE	Y	N	N
Trial Phase FE	Y	Y	Y
Observations	6,693	5,695	3,156
Adj. R-squared	0.3568	0.3337	0.2890

III Data and Estimation Appendix

III.A Details on Information from ClinicalTrials.gov

In a first step, we use the official and short titles of observations in our main dataset to find the corresponding National Clinical Trial (NCT) identifier. For this, we use a combination of searching the ClinicalTrials.gov database for the official and short titles, and fuzzy-string matching between the titles in our dataset and those on ClinicalTrials.gov (using the jellyfish package in Python). After matching, we add additional information from ClinicalTrials.gov (number of participants, trial start and end dates), in order to compare the added information to that from our main dataset. This allows us to determine whether we matched the correct trial. Before determining the correctness of the match (described further below), we also scrape the anticipated completion date as well as primary anticipated completion date from ClinicalTrials.gov's "History of Changes" records associated with each trial (i.e., each NCT identifier). We use the information from the earliest available historical record, and use the anticipated completion date (rather than the primary anticipated completion date) when available.[20]

In a second step, we further assess the correctness of the matched trial from ClinicalTrials.gov. Whenever it is not clear whether a match is correct (e.g., we conclude the match is correct when the information on number of participants as well as trial start and end dates in both datasets coincides, and the Levenshtein distance between titles is at most 10, or the jellyfish similarity score is above 0.99) or incorrect (e.g., we conclude the match is incorrect when the anticipated completion date from ClinicalTrials.gov precedes the start date of the trial in our main dataset), we manually assess the match (e.g., we assess whether the sponsors and collaborators match, taking into account acquisitions, joint ventures, and name changes). Table OA.9 below contains several examples of correctly or incorrectly matched trials between our main dataset and ClinicalTrials.gov. The first (third) example contains a clear correct (incorrect) match. The second example contains a slightly more subtle incorrect match. Finally, for incorrect matches that were based on searching the ClinicalTrials.gov database for the official and short titles, we also implement the fuzzy-string matching, and repeat the subsequent steps for the new matches (adding information from ClinicalTrials.gov, scraping NCT identifier, assessing correctness of each match).

[20] For one observation in the dataset, we use the primary anticipated end date, as the anticipated completion date is implausibly high in the first record (year 2087) and corrected in subsequent records.

Table OA.9: Examples

Main Dataset					ClinicalTrials.gov				
Company	Start Date	Completion Date	No. of Partici-pants	Title	Sponsor/ Collabo-rator	Start Date	Completion Date	No. of Partici-pants	Title
Celgene Corp	8/1/2008	5/1/2014	18	Lenalidomide and Low Dose Dexamethasone Induction Therapy Followed by Low Dose Melphalan, Prednisone, Lenalidomide and Bortezomib Sequential Maintenance Therapy for Newly Diagnosed High-Risk Multiple Myeloma	Celgene Corpora-tion	8/1/2008	5/1/2014	18	Lenalidomide and Low Dose Dexamethasone Induction Therapy Followed by Low Dose Melphalan, Prednisone, Lenalidomide and Bortezomib Sequential Maintenance Therapy for Newly Diagnosed High-risk Multiple Myeloma
Pfizer Inc	3/1/2007	9/1/2009	37	A Phase I, Randomized, Placebo-Controlled, Double-Blind, Dose-Escalation Study Of The Safety, Tolerability, Pharmacokinetics, Pharmacodynamics And Immunogenicity Of a Single Intravenous Dose Of PF-04360365 In Adults With Mild To Moderate Alzheimer's Disease	Pfizer	2/1/2008	10/1/2010	20	A Phase 1, Randomized, Placebo-Controlled, Double Blind, Dose-Escalation, Multicenter Study Of The Safety, Tolerability, Pharmacokinetics, Pharmacodynamics, And Immunogenicity Of A Single Intravenous Dose Of PF-04360365 In Japanese Patients With Mild To Moderate Alzheimer's Disease
TESARO Inc	1/1/2019	6/1/2020	40	Phase II, Multi-cohort Study to Evaluate the Safety and Efficacy of Novel Treatment Combinations in Patients With Recurrent Ovarian Cancer	Cubist Pharma-ceuticals LLC	3/1/2007	9/1/2012	30	Phase II, Open-Label Study to Evaluate the Safety and Efficacy of Daptomycin in the Treatment of Catheter-Related Gram Positive Bloodstream Infections

III.B Matching with FDA Drug Adverse Events Database

For the clinical trials in our main analysis sample, we locate a subsample of trials in which their experimental drugs eventually obtained the FDA approval and are launched in the market. We manually collect drugs' brand names by Googling drugs' experimental names under trials (typically contain the names of drugs' active ingredients or chemical compounds). To ensure that we find the correct brand names, we also match the drug developers and disease indications by searching relevant information from the Internet (e.g., prominent websites including Drug Bank and Adis Insight).

We then use both brand names and experimental names under trials of drugs to match with drug names appearing in the FDA Drug Adverse Events Database. We include drugs' experimental names (e.g., active ingredients) in the matching since sometimes they also show up in the adverse adverse report. But due to the presence of biosimilars and generic drugs, for which the adverse events database also reports the corresponding active ingredients contained in drugs, we conduct extra steps to guarantee the matching quality. Specifically, we identify if a drug (with experimental names reported in the adverse events database) is uniquely developed and marketed by the company. If so, we compare the manufacturers' names in the adverse events database to the drug sponsor names in the trial data.

After obtaining the matched adverse events sample, we restrict our analysis to the incidents with drugs listed as the primary suspect. We drop incidents with any missing information on incident dates and incident primary ID. In the end, we are able to match 3,890,964 adverse events to the drug projects that obtained the FDA approval in our clinical trials data.

III.C Estimation Procedure

Our derivation of the estimation equation follows Wooldridge (2010). Let us rewrite Equation (5) so that

$$\tilde{Y}_i^{out} = \alpha_1 c_i + \boldsymbol{\alpha_2'} \mathbf{X}_i + g(z_i, \mathbf{X}_i, Y_i^{app}) + e_i$$

where $g(z_i, \mathbf{X}_i, Y_i^{app}) = E[v_i \mid z_i, \mathbf{X}_i, Y_i^{app}]$ and $e_i = v_i - E[v_i \mid z_i, \mathbf{X}_i, Y_i^{app}]$. Since $E[e_i \mid z_i, \mathbf{X}_i, Y_i^{app}] = 0$ by construction, we could estimate the above equation y 2SLS on the selected sample using $(z_i, \mathbf{X}_i, g(z_i, \mathbf{X}_i, Y_i^{app} = 1))$ if we know $g(z_i, \mathbf{X}_i, Y_i^{app})$. In the following step, we focus on the derivation of $g(\cdot)$.

By definition,

$$\begin{aligned} g(z_i, \mathbf{X}_i, Y_i^{app} = 1) &= E[v_i \mid z_i, \mathbf{X}_i; Y_i^{app} = 1] \\ &= E[v_i \mid z_i, \mathbf{X}_i; \beta_1 c_i + \boldsymbol{\beta_2'} \mathbf{X}_i + u_i > 0] \\ &= E[v_i \mid z_i, \mathbf{X}_i; \beta_1 (\gamma_1 z_i + \boldsymbol{\gamma_2'} \mathbf{X}_i + \xi_i) + \boldsymbol{\beta_2'} \mathbf{X}_i + u_i > 0] \\ &= E[v_i \mid z_i, \mathbf{X}_i; \beta_1 \xi_i + u_i > -\beta_1 \gamma_1 z_i - (\beta_1 \boldsymbol{\gamma_2'} + \boldsymbol{\beta_2'}) \mathbf{X}_i]. \end{aligned} \tag{11}$$

Given the assumption in Equation (7), we have

$$\begin{pmatrix} v_i \\ \beta_1 \xi_i + u_i \end{pmatrix} \sim Normal \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_3^2 & \beta_1 \sigma_{23} + \sigma_{13} \\ \beta_1 \sigma_{23} + \sigma_{13} & \beta_1^2 \sigma_2^2 + 2\beta_1 \sigma_{12} + 1 \end{pmatrix} \right].$$

Therefore, the conditional expectation of v_i is

$$E[v_i \mid \beta_1 \xi_i + u_i] = \frac{\beta_1 \sigma_{23} + \sigma_{13}}{\beta_1^2 \sigma_2^2 + 2\beta_1 \sigma_{12} + 1} (\beta_1 \xi_i + u_i) = \alpha_3 (\beta_1 \xi_i + u_i)$$

where $\alpha_3 = \frac{\beta_1 \sigma_{23} + \sigma_{13}}{\beta_1^2 \sigma_2^2 + 2\beta_1 \sigma_{12} + 1}$.

Applying iterated expectations on Equation (11),

$$
\begin{aligned}
g(z_i, \mathbf{X}_i, Y_i^{app} = 1) &= E[E[v_i | z_i, \mathbf{X}_i; \beta_1 \xi_i + u_i] \mid z_i, \mathbf{X}_i; \beta_1 \xi_i + u_i > -\beta_1 \gamma_1 z_i - (\beta_1 \gamma_2' + \beta_2') \mathbf{X}_i] \\
&= \alpha_3 E[\beta_1 \xi_i + u_i \mid z_i, \mathbf{X}_i; \beta_1 \xi_i + u_i > -\beta_1 \gamma_1 z_i - (\beta_1 \gamma_2' + \beta_2') \mathbf{X}_i] \\
&= \alpha_3 \lambda \left(\frac{\beta_1 \gamma_1 z_i + (\beta_1 \gamma_2' + \beta_2') \mathbf{X}_i}{\sqrt{\beta_1^2 \sigma_2^2 + 2\beta_1 \sigma_{12} + 1}} \right)
\end{aligned}
$$

where $\lambda(\cdot) = \frac{\phi(\cdot)}{\Phi(\cdot)}$ with $\phi(\cdot)$ as the probability density function of the standard normal distribution and $\Phi(\cdot)$ as the cumulative density function of the standard normal distribution.