Excess Commitment in R&D*

Marius Guenzel

Tong Liu

February 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 termination propensity. This excess project commitment intensifies when the CEO has higher stock-price–compensation sensitivity and is personally responsible for the project's initiation. Welfare implications are nuanced: delay-driven commitment induces investment crowd-out, while not predicting increased adverse effects in marginally-launched drugs and predicting continuation of drugs for diseases lacking alternative treatments.

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

*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; Sertkaya et al., 2016). 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 R&D and consumer 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 trial projects 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 in 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 a 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, combined with full information about managerial ability, yields that a delay in completing the first project phase will not affect

the manager's propensity to subsequently suspend the project. This result motivates our terminology of referring to a positive relation between project delays and continuation probability as *excess* commitment, particularly considering 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, the extent of drug dosage experimentation (i.e., learning-by-doing) in the preceding phase, as well as expected drug sales.

To further hone in on identifying plausibly-exogenous trial delays, 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. We conduct a series of further tests regarding the IV analysis, including using an alternative instrument that isolates trial site congestion solely from drugs in unrelated disease fields. Our IV results remain robust.

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 drug projects (Guedj and Scharfstein, 2004) or by financial constraints preventing them from pivoting. Moreover, the results are difficult to reconcile with reporting incentives or (misspecified) beliefs of firms about anticipated trial phase completion dates, considering our IV design, and because our results remain identical when directly controlling for firms' reported anticipated trial duration.

Instead, the effect of unexpected trial delays on subsequent trial commitment is amplified in the presence of management-related frictions and when the chief executive officer's (CEO's) personal stakes in the project are heightened. Delay is substantially more predictive of continuation among CEOs with greater wealth exposure to stock price changes, 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 predicts project continuation twice as strongly when the CEO heading the firm remains the same between trial phase start and end, consistent with personal responsibility for initiating the project and the project phase playing an important mediating role.

In the final part, we evaluate the implications of our findings for firm investment and consumer outcomes. While unanticipated trial phase completion delay increases the firm's subsequent commitment

[2]Drug development delays are inherently linked to increases in project costs, as longer trial timelines generate, for example, higher labor costs (investigators and staff must be compensated for longer) and increased patient monitoring and care costs (Wong et al., 2014; Sertkaya et al., 2016; Shadbolt et al., 2023).

to that specific project, we find that it crowds out the initiation of new drug projects, particularly those targeting other diseases. In the long run, this may negatively affect both firm profits and drug choices available to consumers. Regarding consumer outcomes observable in the more near-term, we find that 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 of adequate innovation activity in the economy. Seminal prior work has, for instance, uncovered various reasons for potential aggregate under-investment in innovation, including low spillover effects, lack of 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 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 market failures in the innovation and intellectual property markets. In the context of pharmaceutical innovations, Acemoglu and Linn (2004), Budish et al. (2015), and Azoulay et al. (2019) investigate factors influencing and potentially distorting pharmaceutical innovations, including market size potential, short-termism, and scientific grant funding. Additionally, Higgins and Rodriguez (2006), Krieger et al. (2022a), Thakor and Lo (2022), Bonaime and Wang (2023), and Li et al. (2023) delve into how risk aversion, product market competition, mergers and acquisitions, and common ownership among venture capital firms, respectively, shape innovation dynamics and R&D investment.[4] Departing

[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] In the realm of 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 biotechnology firms, and Krieger et al. (2022b) on profit erosion of existing products.

from the previous work, this paper studies how firms' (excess) commitment to the R&D process, an organizational response within firms to existing R&D strategies, affects innovation efforts and outcomes. Furthermore, we examine the potential welfare consequences, documenting the opportunity costs and investment crowd-out effects associated with continuing existing R&D projects, and challenging the conventional wisdom that friction-induced (R&D) decisions by firms are unequivocally detrimental to consumers.

Second, we contribute to the literature in economics and finance on the effects of top management on firm investment. Guenzel (2024) documents distortions in firm investment due to managerial sunk cost effects—which can be viewed as one form of "path-dependent decision-making," similar to excess commitment—within the realm of mergers and acquisitions, with exogenously more expensive acquired businesses being less likely to be abandoned through divestiture.[5] In relation to aggregate efficiency and welfare, Barrero (2022) and Ma et al. (2024) study the macro and equilibrium effects of managerial overconfidence and diagnostic expectations. However, they do not consider consumer-relevant effects on product quality or variety resulting from managerial influences on firm investment. In contrast to the focus in this prior work, ours lies on a distinct investment context—R&D. We expand the scope of management-influencing factors by studying the effect of project delays and quantify various welfare-relevant consumer outcomes. Despite its importance, the exploration of consumer and welfare implications arising from managerial influences on firm decisions remains largely 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 the literature in economics and related fields on the effects of commitment and escalation of commitment. As alluded to above, in an influential laboratory study, Staw (1976) finds that subjects demonstrate stronger commitment to a chosen (hypothetical) R&D project if they were personally responsible for the initial project selection and, consistent with an escalation mechanism, in response to negative signals about the investment. Subsequent research has presented empirical evidence supporting escalating commitment in various contexts including "problem loan" write-offs (Staw et al., 1997) and

[5]Using standard firm-wide R&D data from Compustat, Guenzel (2024) also presents suggestive correlational patterns consistent with a sunk cost mechanism in the realm of R&D investment, suggesting overinvestment in R&D after prior R&D investments.

NBA draft picks (Staw and Hoang, 1995; Camerer and Weber, 1999). A common issue in these empirical studies is the lack of plausibly-exogenous variation in initial investments made, which leaves room for alternative explanations related to information, beliefs, or selection to explain subsequent commitment. We advance this literature by deliberately focusing on how *unanticipated* factors, specifically project delays, influence subsequent commitment.[6] We further advance this literature by examining not only the 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

We first outline a simple theoretical framework to establish a benchmark for the relationship between project delays and the decision to continue versus suspend the project. We sketch the framework here and present the full version in Section A of the Online Appendix.

There are two types of managers: high ability ("good") and low ability ("bad") managers. Each firm is randomly matched with a manager to implement a project. 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 of completing each project phase. After completing Phase I, both types of managers can choose to continue the project or suspend it 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.[7]

[6] We note two papers that use price variation induced by auctions to study commitment in consumer settings, rather than firm settings (Augenblick, 2016 on consumers' penny auction behavior; Ho et al., 2018 on consumers' car usage behavior). Guenzel (2024) exploits post-acquisition-agreement aggregate stock market fluctuations to obtain quasi-random variation in sunk acquisition costs. Additionally, we note the literature on R&D that focuses on how commitment by 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.

[7] The simplified framework abstracts from firms being able to dismiss a bad manager. Simple extensions of the framework could microfound this abstraction, including search frictions and costs of CEO dismissal (see, e.g., Taylor, 2010).

This simple setup is sufficient to characterize our notion of excess commitment. Specifically, in the base case of full information about managerial ability, the framework predicts that manager's probability to subsequently suspend the project will *not* depend on the experienced delay in completing Phase I.

Prediction 1: *With full information about managerial ability, unexpected delays in the already-completed trial phase do not affect a manager's propensity to subsequently suspend the project.*

Proof. See Section A.1 of the Online Appendix. □

This result motivates us to term the positive relation between project delay and continuation probability observed empirically as *excess* commitment, especially 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 A 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 framework.

Our framework abstracts from other underlying frictions, such as more psychological frictions (e.g., those grounded in cognitive dissonance 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. Institutional Background and Data

3.1 Institutional Background

The pharmaceutical industry is highly R&D intensive. For example, in 2019, the whole industry spent $83 billion dollars on R&D, and R&D intensity, defined by R&D spending as a share of net revenues, reached 25 percent, higher than that of the software and semiconductor industries (Austin and Hayford, 2021). A significant portion of the R&D spending in the pharmaceutical industry is dedicated to conducting clinical trials, representing a substantial component of the overall cost of bringing a new drug to market, which frequently exceeds $1 billion (DiMasi et al., 2003; DiMasi et al., 2016; Wouters et al., 2020).

Regulated by the FDA, clinical trials involve three stages: Phase I, II, and III. Phase I trials assess the safety of a drug candidate on a small group of humans and determines how and where the drug distributes within the body. Phase II trials, involving a larger patient pool, seek to determine optimal drug dosage and effectiveness. Phase III trials primarily focus on safety and efficacy across diverse populations, typically involving a large-scale patient group. Trials in all phases necessitate the design of a strict protocol, a scientific plan of action, outlining patient recruitment, the inclusion/exclusion criteria for eligible patients, testing procedures, data collection methods, endpoints (outcomes of interest), and criteria for discontinuation. Successful trial outcomes in each phase are required for progression to the subsequent stages. Upon successful completion of all three phases, drug developers can submit a new drug application to the FDA for market launch approval.

Clinical trials in all phases are frequently conducted at many different sites and typically require several years to complete. In our sample, consisting of Phase I and II trials, and as detailed in Section 3.2, the median trial phase duration is 2.4 years (Table 1; see Section 3.3 for further summary statistics). Wong et al. (2019) report similar numbers in their sample, with a median duration of Phase I, II, and III of 1.6, 2.9, and 3.8 years, respectively.

Reflecting the time-intenstive nature and complexity of clinical trials, their costs are significant. Factors that significantly contribute to the overall costs 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 encompass costs such as patient retention costs, payments to nurses and physicians, laboratory costs, and clinical procedure costs. Per-site costs are approximately the sum of site recruitment costs, site retention costs, administrative staff costs, and site monitoring costs (Wong et al., 2014; Sertkaya et al., 2016).

Notably, these cost factors are heavily dependent on the duration of a trial, with longer trial timelines leading to increased drug development costs, stemming from both heightened per-patient and per-site expenses (Wong et al., 2014; Sertkaya et al., 2016). For example, longer timelines lead to higher labor costs as investigators and staff must be compensated for longer hours. Extended timelines also necessitate more dedicated hours to care for trial participants, further increasing costs for drug developers.

3.2 Data and Sample Construction

Clinical Trials Data. We focus on clinical trials initiated by U.S.-based companies. The trials data come from Cortellis Clinical Trial Intelligence and allow us to observe detailed trial information, including the trial title, trial 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, along with information on the drug developer ID and name, its role in the trial (e.g., sponsor), and details on drug candidates, drug indications,[8] drug dosages tested during the trial, and associated technology. All our analyses are implemented at the trial-phase-by-drug-indication level, i.e., a drug project in our data is defined as a drug-by-indication combination. To obtain the outcome of trials and the corresponding drug projects, we merge the trials data with drug development data from Cortellis Competitive Intelligence. We classify a project as suspended at a particular trial phase if it is explicitly coded as suspended, discontinued, withdrawn, or coded as "no development" for over three years, or if it remains 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 main analysis sample. First, we focus on completed Phase I or II trials and non-missing records on outcomes after these trials, i.e., suspension or continuation. We do not consider Phase III trials as our focus is on whether firms suspend or advance trials through the subsequent phase, necessitating the possibility of a follow-up phase. Second, we drop any trials completed after 2020, because these projects would not have had sufficient time to yield any outcomes. Third, we remove a Phase I or II trials that did not achieve their endpoints. This ensures that if a project is suspended in our data, this is not solely due to unfavorable readouts in the previous trial phase. Finally, we require a project to have both a realized trial start and completion date, thereby excluding projects with estimated completion dates. These filters result in 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 manually match our sample with additional trial data scraped

[8] In all analyses, we will use the most granular, three-level categories of ICD classification codes. We will sometimes refer to them simply as ICD categories or ICD codes for conciseness.

from ClinicalTrials.gov, maintained by U.S. National Library of Medicine. For each clinical trial, drug developing firms must submit initial and follow-up reports to ClinicalTrials.gov. Crucially for our purposes, firms are required to indicate the anticipated trial completion date in their initial submission. This information enables us to construct the key variable 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 actual trial completion date.[9] By matching the dataset with the ClinicalTrials.gov reports, we also obtain trial-specific information on clinical holds imposed by the FDA, as well as obtain detailed location information for all sites involved in a given trial, including street addresses, cities, states, countries, and zip codes. Online Appendix Section B.1 provides further details on the matching procedure of the dataset with the ClinicalTrials.gov reports. Out of the 11,228 projects, we are able to match 9,448 with the ClinicalTrials.gov reports.

Adverse Events Data. To shed light on welfare-relevant consumer outcomes, we also manually link 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 B.2 describes the corresponding matching procedure in more detail.

Other Data. Additionally, we obtain financial accounting and management information for a subset of the drug developers that are public companies traded on U.S. exchanges from Compustat, BoardEx, and Execucomp. We obtain data on expected revenues of drugs upon market launch from Cortellis Competitive Intelligence. We acquire drug different quality proxies, including the drug-specific likelihood of approval (LOA) and the drug's phase transition success rate (PTSR) from the GlobalData Pharma database. Lastly, 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.

[9] Throughout the paper, we use the terms "trial phase completion delay," "trial completion delay," and "trial delay" interchangeably.

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 subsequent phase (Phase II or III), respectively. The average trial takes approximately three years to complete. (We discuss further variables related to trial duration below in Sections 4.1 and 4.2 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 under development, i.e., our sample firms are *not* single-product firms as the firms analyzed in Guedj and Scharfstein (2004), and the average project faces about 400 competing projects within the same ICD code.

Figure 1 shows the geographic distribution of clinical trial sites in our sample. A significant portion of clinical trials is conducted in the U.S, which accounts for 63% of the total trial sites. At the same time, trial activity is substantial in many other countries. Apart from the U.S., other major countries include Canada, Germany, France and Japan, contributing an additional 12% of the total trial sites. These patterns are 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. Canada, Germany, France, and Japan jointly 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 completing the preceding clinical trial phase induce firms to be more likely to "stay committed" to the project and advance the clinical trial through the next trial phase. Figure 2 visualizes the key elements of the project timeline and empirical design.

4.1 Delays 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 first estimate:

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

using 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 (in months) between the actual and the anticipated trial phase completion date as indicated by the firm at the start of the trial phase. As control variables, we include the number of sites in the clinical trial (in logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (in logs), and number of competing projects in the same ICD category (in logs). We also include various levels of fixed effects (FEs) to control for unobserved heterogeneity. These include FEs for the drug company, for the drug's ICD category, and the previously completed trial phase (Phase I or II). We further include FEs for the year-by-quarter of the trial phase start date.[10] As a result, our estimation isolates differences in the suspension-versus-continuation probability of projects that were initiated at the same time (in the same year-quarter) but experienced differential trial delays. We use two-way clustered standard errors, clustered at the ICD category and year-by-quarter levels.

Discussion of the Delay Measure. Inspecting the key variable of interest, the delay in the completion of the preceding trial phase, the average trial phase completion in our sample is nearly one year later relative to the initially anticipated end date (Table 1). Additionally, there is substantial heterogeneity in the extent of delay experienced, which proves useful from an identification standpoint. The 25th-percentile trial phase is not completed with delay, whereas the 75th-percentile trial phase experiences a delay of

[10] Table OA.2 of the Online Appendix examines the influence of the included control variables and fixed effects on a trial's anticipated duration (i.e., the drug developing firm's expectation regarding trial duration at trial start). The analysis suggests that the disease field targeted by a drug project, the year of trial initiation, as well as the identity of drug developing company are three important factors affecting the expected trial duration. As detailed above, our analysis accounts for these differerencs in anticipated durations by controlling for the effects of the various controls and fixed effects.

approximately one and a half years.

Building on the discussion in Section 3.1, trial delays do not solely increase the temporal project investments, but also increase the associated project costs. As previous research discusses, extended trial timelines result in, for instance, escalated labor expenses, as investigators and staff require compensation for the extended duration, alongside heightened costs associated with patient monitoring and care due to the prolonged trial period (Wong et al., 2014; Sertkaya et al., 2016; Shadbolt et al., 2023). Increased project costs associated with delays can serve to reinforce the key prediction that as project delays increase, firms are increasingly inclined to continue rather than abandon ongoing R&D projects.

In terms of measurement of trial completion delays, as the difference between the actual and the anticipated trial end dates, 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 were to misreport expected trial end dates at the outset. Additionally, reporting guidelines for ClinicalTrials.gov explicitly stress that "it is important these dates are accurate."[11] These reporting incentives reduce the possibility of measurement error (which would plausibly bias the estimated coefficient towards zero) or distortions in the delay variable. Additionally, our IV strategy (discussed next in Section 4.2) based on changes in trial site congestion is further suited to address possible remaining 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.

4.2 Increase in Trial Site Congestion as an Instrument for Trial Phase Completion Delays

The *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, we introduce an instrumental variables (IV) strategy that uses *trial 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. Focusing on the change

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

in congestion since trial start allows us to account for the fact that trial sites (and therefore initial trial busyness) are not chosen randomly by firms.

Intuitively, the instrument is based on the notion that trial sites have limited capacity to accommodate clinical trials. When a trial site becomes overburdened with trials, it becomes more challenging to, for example, recruit a sufficient number of patients on time, ensure proper staffing to monitor all trials simultaneously, and have timely access to specialized medical devices, diagnostic tools, laboratory equipment, or other technology required for data collection and analysis. Such challenges increase the likelihood of bottlenecks and trial slowdowns.

Construction of the Congestion Measure. To construct our instrument, we download the universe of trial records across all phases from ClinicalTrials.gov, i.e., all available Phase I, II, and III trials, which is a superset of our main sample. We then standardize the zip codes of the 20 most frequent countries where these trials are conducted, covering about 90% of total trials at the ClinicialTrials.gov. We define each unique zip code in a country as one trial location, z, and drop any trials with missing addresses, number of participants, and start and completion dates. In a first step, we 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 enrollment per year and 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 for location z in year t, G_{zt}, is the sum of E_i for all trials that have one site located in zipcode z and year t:

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

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 a second step, we calculate the *change* in congestion at location z between the start year τ_1 and the completion year τ_2 for trial i. We scale this change by the mean level of congestion in location z across years, $\overline{G}_z$, i.e, one can interpret it as the relative change in congestion at location z. Our instrument for trial completion delay is then

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

i.e., the instrument uses the maximum congestion change across trial *i*'s locations, since a clinical trial cannot be completed until the final set of trial results is recorded. For the average trial in our sample, the maximum increase in trial congestion between trial start and end across trial sites, normalized by the mean congestion, is slightly more than five percent (Table 1).

Discussion of the Exclusion Restriction. Besides the instrument relevance condition—i.e., a change in trial site congestion strongly affects trial completion delays (which we will establish in Section 5.2)—the key identifying IV assumption is that the change in trial site congestion affects project suspension only through its effect on trial delays. One argument in favor of this assumption is that the change in trial site crowdedness of a zip code is an aggregate measure of trial site activity across drug developers and drug fields, and thus unlikely to be correlated with unobserved shocks to firms' decisions to advance a specific clinical trial (a firm-level condition). At the same time, there are plausible concerns regarding the exclusion restriction. One concern is that more crowded trial locations with greater participant numbers may suggest higher service quality, including enhanced site monitoring, data collection, and implementation of clinical protocols. Clinical trials conducted at these sites, while potentially experiencing delays in schedule, may yield results of higher quality. Consequently, firms would be more inclined to advance trials to the next phase. To ameliorate this concern, we collect detailed hospital quality data as a proxy for trial site quality (see Section 3.3) and examine their correlation with our congestion instrument. As Figure 3 shows, we find no evidence indicating a meaningful correlation, either positive or negative, between the congestion instrument and trial site quality, as proxied by the average clinical score/evaluation across hospitals in a zip code, encompassing a total of 26 hospital quality measures.

A second concern is that while the instrument picks up aggregate rather than project-specific variation, it could be that there is an aggregate, correlated shock (e.g., an innovation or drug demand shock in a given disease field) that increases trial site activity and prospects of drugs under development. We relegate the discussion of this concern to Section 5.2 when we present the IV results. To preview, the results are robust to controlling, linearly or non-parametrically, for the number of competing drug projects in the same ICD category. Additionally, we construct an alternative instrument that excludes patient flows from drugs in the same ICD category, thus purging the instrument of the possibility of correlated disease field shocks, and isolating trial site busyness and congestion induced by unrelated drug projects. Our IV results

are very similar when using this alternative measure.

Finally, we also find that the congestion 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 more detail in Section 5.3.

5. Results

5.1 Delays in Trial Phase 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 within 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 phase 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 include fixed effects for the drug's ICD category and for the trial phase (completed Phase I versus II trials). Finally, the effect remains unchanged in Column (4) when we include control variables for the number of sites where the clinical trial is conducted (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). The coefficients on the included control variables (unreported) are intuitive. More projects in the pipeline and more competing projects are positively (but insignificantly) associated with suspension. Having more trial sites increases the probability of suspension, whereas having more trial participants reduces it.

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* firms' initial expectations, increases subsequent project commitment by firms as measured by the decision to subsequently suspend versus continue the trial. The next two sections will expand on this evidence, addressing the key question of whether unanticipated project delays might contain or produce information that increases the value of projects and therefore should lead firms to increase their commitment and propensity to continue delayed projects. In particular, Section 5.2 will present the delay-from-congestion IV results, while Section 5.3 will offer further corroborating evidence from drug quality and experimentation proxies.

5.2 Increase in Trial Site Congestion as an Instrument for Trial Phase Completion Delays

Table 3 presents the main IV results, estimating 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 control variables from the final column of Table 2. Both columns reveal a strongly positive 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 well above the common threshold for weak instruments of 10 (Stock et al., 2002). Figure OA.2 of the Online Appendix confirms the strongly positive, and nearly monotonic, relationship between trial site congestion increase and the resulting unanticipated clinical trial delays through a binned scatter plot.

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 project subsequently. 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.) These 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.

IV Robustness. As alluded to in Section 4.2, in spite of the non-correlations of the instrument with

trial site quality proxies (Figure 3), another plausible concern regarding the congestion-based IV approach is the possibility of disease-category-wide shock affecting both trial activity and project viability. Three pieces of evidence alleviate this concern. First, like in the final column of the OLS results in Table 2, the final IV column in Table 3 directly controls for the number of competing drug projects in the same ICD category, thus accounting for broader, disease-field-wide effects. Second, Online Appendix Table OA.3 shows that the IV results are virtually unchanged when controlling for the number of competing drug projects in the same ICD category nonparametrically through number-of-competing-drug decile fixed effects rather than parametrically, and further, that the estimates are also quantitatively similar in the subsamples with below-median and above-median drug project competition, respectively. Third, in Online Appendix Table OA.4, we go one step further and re-estimate the IV model using an alternative congestion instrument that excludes patient flows from drugs in the same ICD category. This alternative instrument, whose construction we describe in detail in Online Appendix Section B.3, identifies trial site busyness and congestion solely from unrelated drug projects, eliminating any effects from drugs within the same ICD disease field. The IV results remain unchanged, economically and statistically, under this alternative specification.[12]

5.3 Association Between Trial Delays and Project Experimentation and Quality

A key question regarding the main finding that the experience of unanticipated trial delays increases firms' subsequent project continuation propensity is whether delays might produce information that increases the value of projects and as a result should prompt firms to increase their project commitment. The IV results from the preceding section serve as a first way to alleviate this concern, as long as change in trial site congestion is uninformative of and uncorrelated with project quality and information. This section presents additional corroborating evidence using various measures of drug project experimentation and quality.

Drug Experimentation. Related to drug experimentation, it could be the case that drug developers

[12] As shown in Online Appendix Table OA.4, both the first-stage and second-stage results remain unchanged when using the alternative, other-ICD-category-based congestion instrument. In terms of the IV relevance condition, the Kleibergen and Paap (2006) $F-$statistic remains well above 10, confirming that delays continue to be predicted from congestion induced by projects in other ICD codes, e.g., due to fewer remaining resources available for patient recruitment efforts, staffing challenges, or physical space constraints.

who experiment with multiple dosages in a clinical trial are more likely to encounter (unanticipated) delays while at the same time generating additional scientific data points on optimal dosages. This, in turn, could enhance the project's value and a firm's inclination to continue the project, constituting an alternative explanation of the observed patterns.

While this alternative explanation is more relevant in the OLS approach, and less so in the IV approach that uses the instrumented delay measure, we examine it further by exploiting project-specific information on drug dosages tested. Specifically, in Table 4, we investigate the association of our delay and instrument measures with the number of dosages tested in each clinical trial, *Num of Doses in a Trial*. Column (1) shows that the number of dosages experimented with in a trial is not significantly correlated with the experienced trial phase completion delay. Similarly, Column (2) shows that the extent of dose experimentation is also insignificantly correlated with our instrument, the change in trial site congestion. Overall, this evidence is inconsistent with drug experimentation being a primary driver of our findings.

Drug Quality. A natural concern is that there might be some remaining omitted variable that simultaneously affects project delays, quality, and continuation probability. Even though, similar to the argument above, our IV strategy can alleviate this concern as long as the congestion measure is uncorrelated with project quality, we collect additional project-specific information to further allay the concern. Specifically, we collect and utilize two objective, third-party measures of drug quality from the GlobalData Pharma database: the drug-specific likelihood of final FDA approval (LOA) and 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).[13] In Table 5, we then examine the association of our delay and instrument measures with the drug approval and phase transition likelihood measures. In Columns (1) and (2), the correlation between the LOA and PTSR measures with the experienced trial phase completion delay is insignificant, both economically and statistically. Furthermore, in Columns (3) and (4), the correlation between the LOA and PTSR measures with our instrument, the change in trial site congestion, is also insignificant.

[13]For example, the Phase-II drug candidate, *ceftolozane sulfate + tazobactam sodium*, developed by Merck & Co Inc for Febrile Neutropenia, is estimated to have a 33% drug-specific LOA and a 65% drug-specific PTSR. To obtain the drug quality measures, we match each drug in our sample with the GlobalData Pharma database by drug name and compute the average LOA and PTSR across companies associated with the drug and indications as the proxies for drug quality.

In the Online Appendix, we also perform two further tests relating experienced trial delays to additional trial quality and profitability proxies. First, in Panel A of Online Appendix OA.5, we show that trial delays are uncorrelated with expected drug sales conditional on drug launch, data obtained from Cortellis Competitive Intelligence. This holds for various expected sales measures, including the sum of the expected first five-year sales and the expected average annual sales after FDA drug approval (over all available years). Second, in Panel B of Table OA.5, we show that trial delays are also uncorrelated with the frequency of severe adverse reactions during the already completed trial phase.[14]

Overall, the evidence in both Tables 5 and OA.5 speaks against an omitted variable related to project quality as an underlying explanation.

Sample Selection. The drug experimentation and quality evidence discussed in the preceding paragraphs is also helpful in addressing a related alternative explanation revolving around sample selection. Our sample is constructed to only include completed trials that achieved their endpoints (cf. Section 3.2). As discussed, this requirement is necessary to ensure that if a project is suspended in our data, this decision is not solely governed by unfavorable readouts in the previous trial phase. Simultaneously, the requirement may introduce the possibility that certain drug developers are inherently more committed to their drug projects, and thus more likely to sustain project delays and successfully complete the trial phase achieving all endpoints. However, this potential underlying mechanism is not easily consistent with three aspects.

First, all analyses include firm fixed effects, thereby absorbing any potential differences in firms' ex-ante drug project commitment across drug developers. Second, our results are also robust when adopting a more stringent specification that includes firm-by-ICD-category fixed effects. That is, this specification further accounts for any potential differences in ex-ante commitment within the same firm but across disease fields and project types (e.g., it accounts for a firm possibly having a different ex-ante commitment for cancer drugs compared to pneumonia drugs). As shown in Online Appendix Tables OA.6 and OA.7, both the OLS and IV results remain similar with the inclusion of the tighter firm-by-ICD-category fixed

[14] We examine both the the total and maximum percentage of life-threatening reactions relative to the total number of enrolled trial participants. Specifically, for a life-threatening reaction $r \in R$, such as acute myelogenous leukemia, basal cell carcinoma, and brain hemorrhage, we first calculate the ratio δ_r as the number of participants experiencing reaction r over the total number of trial participants. The total percentage of life-threatening reactions is the sum of δ_r over $r \in R$. The maximum percentage of life-threatening reactions is the largest δ_r among $r \in R$.

effects.[15] Third, we would expect—particularly after the inclusion of the various fixed effects controlling for unobserved heterogeneity—any possible remaining differential ex-ante commitment across drug projects to be correlated with *some* project observable. In contrast to this, the previous tables show that our delay measures are insignificantly correlated with various observables, including drug likelihood of approval, phase transition success rate, expected sales, and severe adverse reactions during clinical trials. Consequently, the hurdle for a sample-selection-based explanation of our findings becomes that there exist unobservable differences in ex-ante commitment across projects within the same firm (and disease field) that are both unrelated to all considered project observables and simultaneously affect firms' likelihood of achieving successful trial completion. While impossible to rule out, we view this as unlikely compared to our proposed more parsimonious explanation based on delay-*induced* commitment differences.

5.4 Underlying Channels

This section examines various potential underlying channels that might contribute to (excess) commitment and project continuation in response to trial phase completion delays. In Table 6, we examine firm- and project characteristics-related channels, including agency conflicts inside the firm, initial firm expectations about trial length, financial constraints, and heterogeneities by trial phase. In Table 7, we examine mechanisms pertaining to the ultimate decision-maker responsible for R&D investment, the CEO of the drug developing firm.

5.4.1 Project-Characteristics- and Firm-Related Channels

Lack of Other Viable Drug Candidates. In 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

[15] We note the statistical significance drops in in Column (3) of Table OA.7, but all results remain significant at 5% or 1% in the most stringent specification with both firm-by-ICD-category fixed effects (as well as other fixed effects) and control variables.

(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 6 shows that unexpected delays continue to predict trial continuation decisions with nearly identical 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).

Trial Duration Expectations. 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 (realized end date minus anticipated end date). Consequently, it could be the case that when firms are more optimistic about a particular project (and therefore more inclined to advance it regardless of circumstances), they underestimate or strategically underreport the trial's anticipated duration. As discussed in Section 4.1, two arguments speaking against such a channel are that firms are required to truthfully report anticipated trial end dates and that our IV design accounts for this, as long as remaining reporting incentives or project expectations are uncorrelated with the congestion instrument. Furthermore, when we when we directly control for firms' reported anticipated duration in Column (2) of Table 6, our results remain unaffected.[16] Considering this evidence, a mechanism based on trial duration expectations is unlikely to be an underlying mechanism driving our findings.

Financial Constraints. Next, we test whether firms' sustained commitment to drug projects in response to trial delays may be driven by financial constraints of 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 various financial constraints measures (thereby restricting the sample to public firms only). Specifically, the *Constrained* variable in the third column of Table 6 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.[17] Thus, the path dependence in

[16] In unreported tests, we include second- and third-order polynomials for anticipated duration and find nearly identical results compared to those reported in Table 6.

[17] We show robustness to other measures of financial constraints, including 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.8. Irrespective of the measure we use, we continue to find a strong, positive effect of trial delay on subsequent continuation probabilities.

decision-making is not explained by financially constrained firms not being able to pivot.

Phase I Versus II Trials. Finally, the last two columns in Table 6 separate the main result by whether the preceding trial phase was a Phase I or Phase II project. 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 explained by any trial-phase-specific channels or confounders, and instead represent a robust phenomenon in the realm of new drug development.

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' personal stakes in the project and firm, specifically, how it 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. The analysis of how CEO pay-performance sensitivity (delta) mediates the relationship between trial delays and continuation decisions is motivated by our theoretical framework from Section 2.[18] Specifically, the extended framework with asymmetric 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 (2024) shows that managers engage in sunk-cost thinking with respect to acquired targets, distorting their subsequent decisions to abandon acquired businesses through divestiture. Guenzel (2024) finds that the sunk cost effects are driven by the CEO who made the initial acquisition.

[18] 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).

To test for a similar CEO effect in our setting, we separately estimate the effect of delay on trial continuation in two subsamples: the subsample of CEOs who initiated the project phase, i.e., the subsample of firms where CEO remained the same between trial phase start to completion (Column (3) of Table 7), and the subsample of firms that appointed a new CEO between trial phase start and completion (Column (2)). Comparing the two columns, the commitment effects we uncover are substantially 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 trial commitment is about twice as large when there is no CEO change between trial start and end, with the difference being statistically significant as well ($p = 0.057$).

Additionally. we have also explored CEO age as a possible further mediating factor in the project-initiating CEO subsample. We find no statistically significant difference in the effect of delay on project continuation between younger CEOs, who have longer careers ahead of them, and older CEOs closer to retirement. This suggests that the managerial responsibility channel likely arises not only from externally-oriented career and reputation concerns (which are expected to be 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 a significant underlying mechanism for the documented delay-induced (excess) commitment to R&D projects within firms.

6. Further Identification Challenges and Robustness Tests

Throughout the preceding sections, we have discussed a range of robustness tests for the key finding that trial completion delays increase firms' subsequent project commitment, including tests based on drug experimentation and quality proxies (Tables 4, 5, and OA.5), tests based on alternative specifications with even more granular, firm-by-ICD-code fixed effects (Tables OA.6 and OA.7), and robustness tests related to the congestion instrument based on trial site quality proxies, nonparametric competition controls, and congestion induced by unrelated drugs (Figure 3, Tables OA.3 and OA.4). This section extends the discussion of robustness tests to address potential additional identification concerns, as well as presents

additional corroborating evidence from exchange-rate-induced variation in trial costs that supports our main hypothesis.

Trial Delays Reflecting Real Options. First, directly related to the prior discussion on the association between trial completion delays and drug experimentation and quality measures, one may be worried that trial delays might reflect firms exercising a real option in light of favorable new information (Pindyck, 1991; Dixit et al., 1994; McAfee et al., 2010). The evidence discussed speaks against this channel as the driver of our findings. In particular, the instrumented, congestion-induced delay is plausibly unrelated to a "learning-by-doing" channel, and neither project delay nor congestion is significantly associated with any observable experimentation and quality measures (drug dosages, project success probabilities, expected sales, adverse events during trials), which is difficult to reconcile with a real options mechanism.

FDA Clinical Trial Holds. Second, we address possible concerns that the link between trial delays and continuation may reflect the effect of FDA clinical trial holds, which are orders issued by the FDA to the sponsor of the drug development to delay a proposed clinical investigation. Common reasons for clinical holds include the presence of an 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). During the hold, drug developers would work with the FDA to improve the clinical protocol and address all deficiencies. Consequently, a trial that experienced a hold may have a prolonged duration, but the FDA may potentially help improve the trial quality throughout the hold period, which, in turn, may increase the probability of project continuation. The concern mostly applies to the OLS-based delay results in Table 2, rather than the IV-based delay results in Table 3 and, to address it further, Online Appendix Table OA.9 shows that there is no statistically significant relationship between the presence of clinical holds and trial duration, trial delays, or the trial site congestion instrument. We conclude that our findings are not driven by FDA clinical holds.

Human Capital. Third, one remaining concern may be that firms' commitment to ongoing R&D activities reflects their efforts to retain valuable human capital. For example, abandoning one project and transitioning to an unrelated one might pose challenges in retaining scientists specialized in a specific disease field. Such constraints, while plausible, are unlikely to affect our estimates since human capital incentives, if present, will be inherent within the R&D process and should thus not vary by the extent of

experienced project delays.

Further Miscellaneous Tests. Fourth, we implement additional miscellaneous robustness tests. In Table OA.10, 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.[19] In Table OA.11, 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 presented in Table 6, these results highlight the general applicability of our findings.

Further Validation From Exchange-Rate-Induced Variation in Trial Costs. Finally, before delving into the possible implications of our key finding that trial delays strongly predict subsequent trial continuation for firm investment and consumer outcomes, we conduct one further validation test, based on exchange-rate-induced variation in trial costs. We summarize this exercise here, relegating the details to Online Appendix Section B.4.

The key idea behind the validation test is that if firms respond to trial delays (i.e., temporal project-specific investments made) by increasing their commitment to the project, we would expect to see similar patterns for unexpected increases in trial-specific project financial investments. To test this, we leverage the institutional feature that a large portion of clinical trials are conducted outside of the U.S. (see Figure 1). For trials that involve international sites, the contracts and actual payments are typically 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.[20]

Appendix Table OA.12 shows how the exchange rate-induced project cost fluctuations (see Online Appendix Section B.4 for details on the construction of the measure) affect firms' decision to continue

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

[20] 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."

versus suspend trial projects. Columns (1) and (2) are estimated on the subsample of foreign-based trials, with Column (2) adding control variables. 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. Across columns, the coefficient on exchange rate changes is negative and statistically significant, implying that exchange-rate-driven cost increases make trial suspension less likely. 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, while slightly lower compared to the effects of unexpected delays on suspension decisions, are economically meaningful.

To further probe the robustness of the exchange rate results, we also conduct a placebo test that randomly assigns trial countries to trials in the sample and re-runs the regressions from Appendix Table OA.12 using these placebo cost fluctuations. Figure OA.3 of the Online Appendix plots the distribution of the estimated coefficients of interest on the exchange rate variables, repeating the estimation 1,000 times. In contrast to the results in Appendix Table OA.12, 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.

7. Implications for Firm Investment and Consumer Welfare

In the final part of the paper, we explore the effects of delay-induced commitment to drug projects on various firm and consumer outcomes.

7.1 Outcomes Variables Related to Firm Investment

To explore the implications for firm investment, we focus on spillover effects on other investments firms make. Specifically, we examine the take-up of new drug projects, and test whether delay-induced commitment leads to crowd-out of other projects. Any potential crowd-out effects in firm investment

may worsen firms' future drug pipeline and revenue prospects and, ultimately, may also have negative implications for consumers in terms drug choices available to them. Concretely, we compute three outcomes variables: 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), the total number of new projects that share the same ICD indication as the focal project, and the total number of new projects that do not share the same ICD indication. Throughout our analysis, projects are defined at the drug-by-indication level, and we include all newly initiated projects with development status "Discovery" or "Preclinical."

7.2 Outcome Variables Related to Patient Welfare

To explore the implications of delay-induced drug project commitment for direct consumer outcomes observable in the near term, we consider two distinct outcome variables. At the extensive margin, an approved new drug can save lives if there is no alternative drug available on the market to target the disease. To measure this effect, we examine the effect of delay-induced commitment on drug approval outcomes for so-called orphan drugs targeting infrequent diseases with typically few available treatment options. At the intensive margin, drugs might yield adverse reactions among users once approved, which is indirectly related to the drug's safety and immediately relevant for consumer welfare. To assess the impact along this dimension, we study the occurence and frequency of adverse events occurring after the drug is launched in the marketplace.

7.2.1 Econometric Framework

In contrast to the OLS and 2SLS approaches discussed in Section 4, one complication in examining consumer outcomes, particularly adverse reactions to drugs after they are launched in the market, arises from the fact that adverse events are observed only for the drugs that are ultimately approved and launched. To account for this endogenous sample selection, we apply a parametric framework for the tests related to adverse events. Specifically, this framework addresses sample selection with endogenous explanatory variables by modifying the classic Heckman selection model (e.g., Heckman, 1979; Wooldridge, 2010). An alternative approach is 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 the econometric 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 previously incurred delays, or temporal costs, influence 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

$$\tilde{Y}_i^{app} = \beta_1 c_i + \beta_2' \mathbf{X}_i + u_i \tag{2}$$

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 $\tilde{Y}_i^{out}$. We assume that $\tilde{Y}_i^{out}$ is a linear function of observables $\mathbf{X}_i$ and c_i, such that

$$\tilde{Y}_i^{out} = \alpha_1 c_i + \alpha_2' \mathbf{X}_i + v_i \tag{3}$$

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, we assume c_i is generated according to the model

$$c_i = \gamma_1 z_i + \gamma_2' \mathbf{X}_i + \xi_i \tag{4}$$

where z_i is the instrumental variable.

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

$$\begin{pmatrix} u_i \\ \xi_i \\ v_i \end{pmatrix} \sim Normal \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix} \right] \tag{5}$$

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.

The above setting and assumptions are similar to the sample selection model with endogenous explanatory variables in Chapter 19 of Wooldridge (2010). One difference is that we allow the endogenous variable to enter the sample selection Equation (2). To discipline the model, we therefore impose extra structure on the residuals through Equation (5). Note that Equation (5) 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 $E[v_i \mid \beta_1 \xi_i + u_i] = \alpha_3(\beta_1 \xi_i + u_i)$.

Our estimation procedure, which is similar to that proposed in Wooldridge (2010). then proceeds as follows (with a more detailed derivation available in Section B.5 of the Online Appendix):

Step 1: 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 (4). Compute the estimated inverse Mills ratio, $\hat{\lambda}_i = \lambda(\frac{\hat{\beta}_1 \cdot \hat{\gamma}_1 z_i + (\hat{\beta}_1 \cdot \hat{\gamma}_2' + \hat{\beta}_2')\mathbf{X}_i}{\sqrt{\hat{\beta}_1^2 \hat{\sigma_2^2} + 2\hat{\beta}_1 \hat{\sigma_{12}} + 1}})$, where $\lambda(\cdot) = \frac{\phi(\cdot)}{\Phi(\cdot)}$ with $\phi(\cdot)$ being the probability density function of the standard normal distribution and $\Phi(\cdot)$ being the cumulative density function of the standard normal distribution.

Step 2: Use the selected sample in which Y^{out} is observed and estimate the equation

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

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

Standard errors and test statistics are corrected for the generated regressor by bootstrapping.

7.3 Effects on Firm Investment

We first assess the effects of delay-induced drug continuation on other firm investments, and specifically potential crowd-out effects of other drug projects. For this part, we can continue to use the IV

approach from Section 4.2. That is, we continue to use the change in trial site congestion as an instrument for trial completion delay, studying the number of newly initiated drug projects in the year after the completion of the focal project (i.e., the trial phase in our main sample) as the outcome variable of interest.

Table 8 presents the second-stage estimation results of this IV approach, 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-year delay is associated with approximately one fewer new project being started, compared 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 investment in alternative fields. 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 as they pertain to the development and launch of orphan drugs—drugs for rare diseases for which there are commonly no other or only few other treatment alternatives available. For this part, we can also continue to rely on the IV strategy from Section 4.2. Specifically, *conditioning* on clinical trials involving orphan drugs, we examine how the likelihood of eventual FDA approval for orphan drugs (i.e., the launch of the drug in the market) varies with congestion-induced delay in the completion of Phase I or Phase II clinical trials.[21]

Table 9 presents the results, structured similarly to Table 3. Despite the reduction in sample size (due to the restriction to orphan drug clinical trials), the first stage remains strongly significant (Column (1)). In the second stage, using the indicator variable for FDA orphan drug approval as the dependent variable of interest, we estimate a significant, positive effect of instrumented delay on the probability that orphan

[21] Conditioning on orphan drugs in this analysis, and providing a within-orphan-drug test rather than comparing across orphan and non-orphan drug estimates, is important to account for the regulatory and institutional differences between orphan and non-orphans (e.g., Ridley et al., 2006).

drugs are ultimately launched in the marketplace (Column (2)). In terms of magnitudes, 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. Overall, these results point to possible positive implications for (a subset of) consumers associated with commitment effects in new drug development.

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, using the econometric framework detailed in Section 7.2.1. Table 10 shows the results, with standard errors bootstrapped using 100 iterations. As in Table 9, in Column (1) we continue to observe a strong relation between trial site congestion and trial delay in the first stage (i.e., the first part of the second step outlined in Section 7.2.1), despite a reduction in sample size, now due to the fact that we focus on outcomes of approved drugs. In the second stage, we estimate modest increases in adverse events after congestion-induced delay that are, however, notably insignificant. This is the case both when examining adverse events over a one-year horizon since drug launch in Column (2), and when examining a three-year horizon in Column (3). Economically, 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 we find evidence that the commitment effects expand the range of treatments available for rare diseases with limited existing medications, they may also result in the launch of drugs with more adverse effects, and lead to a reduction in firms' take-up of other drug projects. More broadly, the results highlight that for non-firm parties, firms' friction-induced R&D decisions need not exclusively entail welfare costs; instead, they can, and may frequently, generate positive externalities for (subsets of) other parties.

8. Conclusion

In this paper, we study how firms' commitment to ongoing R&D projects is influenced by the amount of temporal (and financial) resources the firm has already invested in the project, and examine resulting welfare implications. We study these questions using project-level R&D data from the pharmaceutical industry, specifically clinical trials data with granular information on key dimensions such as project timelines and outcomes.

Clinical trials that experienced unanticipated and plausibly exogenous delays are significantly more likely to be continued by firms through the next trial phase, even though delays are empirically uncorrelated with various project quality measures. We estimate a 0.6 percentage point increase in the probability of drug project continuation (2% relative to the baseline continuation probability) associated with a median delay in our sample of three months. These economic effects are sizable, especially considering that project delays reduce drugs' patent protection time. In terms of implications for firm investment, delay-induced commitment to drug projects crowds out the initiation of new drug projects. The 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 several avenues for future research. One promising area for investigation would be to examine heterogeneity along various dimensions. For example, it would be interesting to explore how the extent of excess R&D commitment varies with organizational structure (e.g., top-down versus decentralized decision-making) and organizational culture (e.g., the degree of tolerance for failure). Moreover, investigating whether there are heterogeneous effects on consumer welfare outcomes (e.g., based on the severity of adverse effects or different consumer demographics) could yield additional valuable insights. Finally, another important open question relates to how competitors respond to firms' excess commitment to R&D projects and the resulting general equilibrium effects. We leave these topics for future research.

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. (2024): "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.

Higgins, M. J. and D. Rodriguez (2006): "The outsourcing of R&D through acquisitions in the pharmaceutical industry," *Journal of Financial Economics*, 80, 351–383.

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 (2024): "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.

RIDLEY, D. B., H. G. GRABOWSKI, AND J. L. MOE (2006): "Developing drugs for developing countries," *Health Affairs*, 25, 313–324.

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

SERTKAYA, A., H.-H. WONG, A. JESSUP, AND T. BELECHE (2016): "Key cost drivers of pharmaceutical clinical trials in the United States," *Clinical Trials*, 13, 117–126.

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.

TAYLOR, L. A. (2010): "Why are CEOs rarely fired? Evidence from structural estimation," *The Journal of Finance*, 65, 2051–2087.

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.

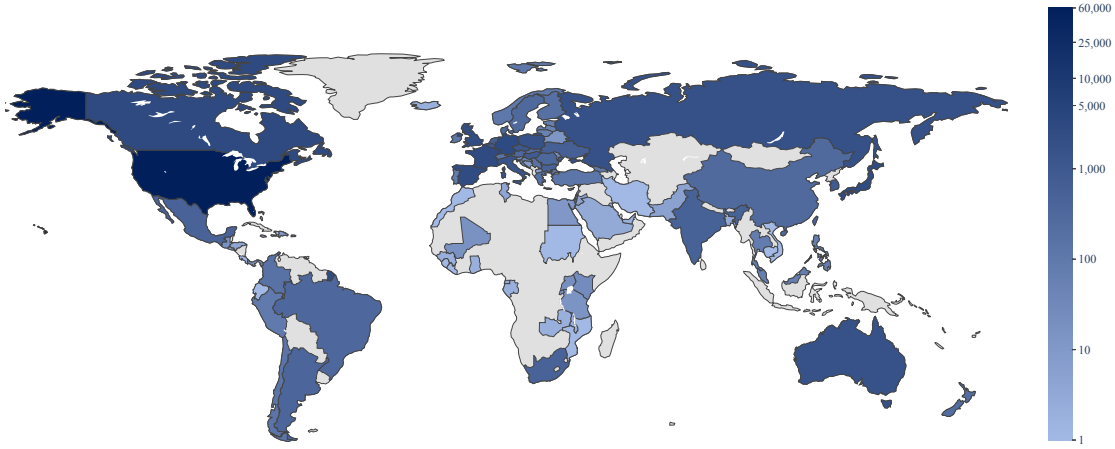
WONG, H.-H., A. JESSUP, A. SERTKAYA, A. BIRKENBACH, A. BERLIND, AND J. EYRAUD (2014): "Examination of clinical trial costs and barriers for drug development final," *Reports Submitted to Office of the Assistant Secretary for Planning and Evaluation*.

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.

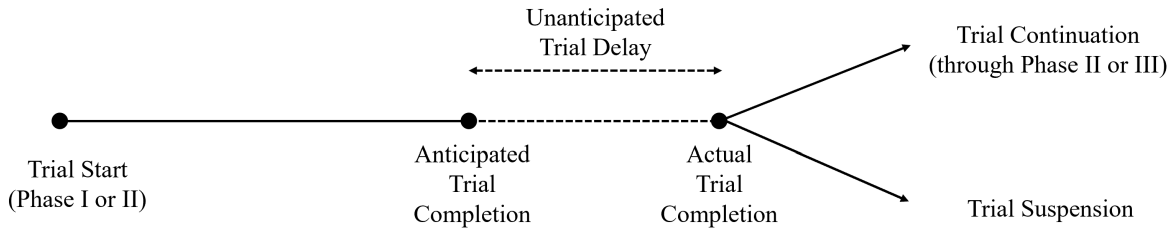
Figures

Figure 1: Geographic Distribution of Trial Sites



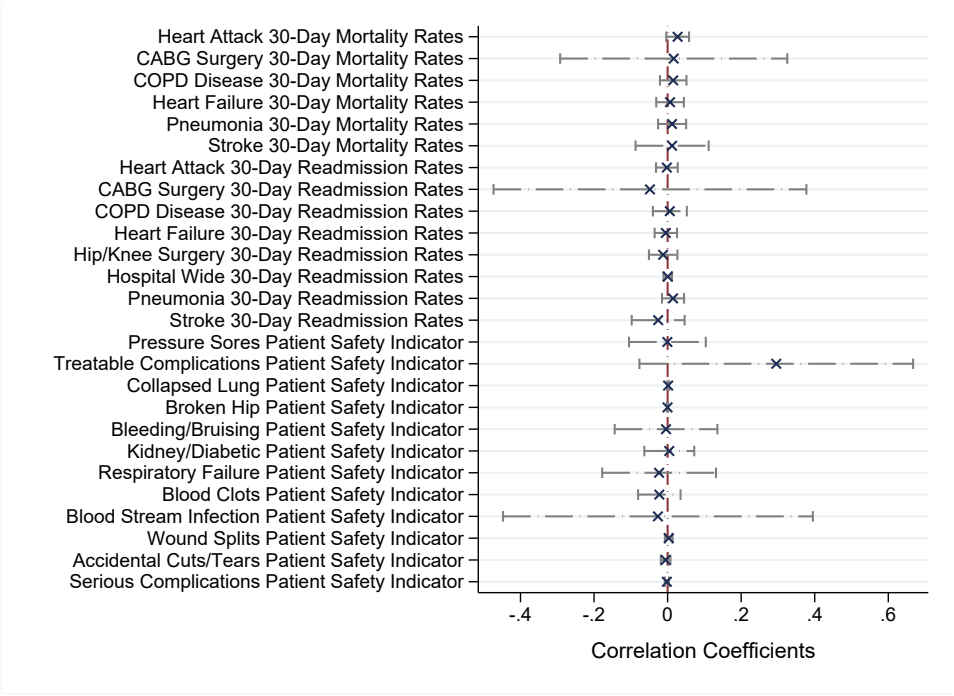
This figure shows the trial-site geographic distribution of clinical trials included in the analysis sample. For each country, we calculate the natural logarithm of the number of trial sites.

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: Association Between Change in Trial Site Congestion and Hospital Care Quality

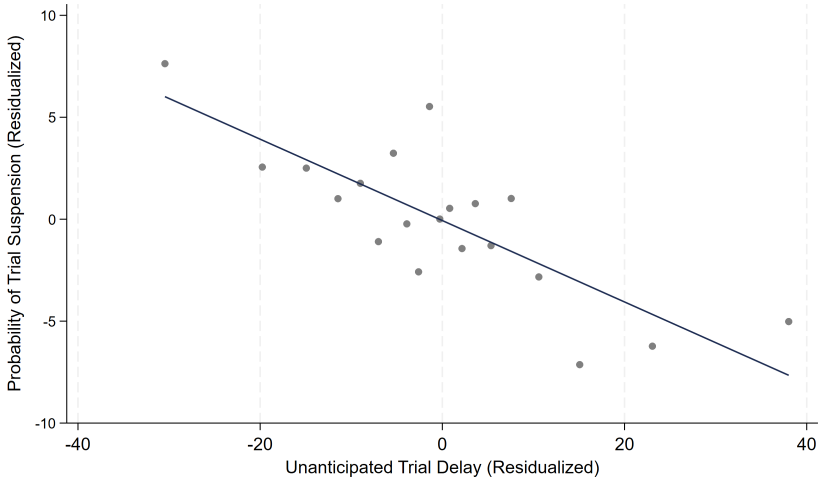


This figure depicts the associations between the trial site congestion measure and the various hospital care quality measures in a zip code. The plotted coefficients, β, are estimated by the regression

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

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, and γ_z and τ_t are zip code and year fixed effects, respectively. The y-axis shows the names of various quality measures. All standard errors are two-way clustered at the zip code and year levels. Capped lines represent 95% confidence intervals.

Figure 4: Delays in Trial Phase 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	11,182	402.07	812.80	81.00	206.00	406.00
Change in Trial Site Congestion	9,948	5.28	10.04	0.00	0.93	4.19

Table 2: Delays in Trial Phase 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 the 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 year-by-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 Phase Completion and Project Continuation Versus Suspension: Increase in Trial Site Congestion as an Instrument for Delay

This table reports the IV estimation results of Equation (1), using the change in trial site congestion as an instrument for delay in 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, zip-code-level 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 the 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 year-by-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: Association Between Trial Delays and Drug Experimentation

This table examines the association between trial delays and drug experimentation (number of doses tested in the clinical trial). The dependent variable is *Delay in Trial Completion* and *Trial Site Congestion*, respectively. The independent variable, *Num of Doses in a Trial*, counts the number of drug doses tested in the trial. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in the 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 year-by-quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	DV: Delay in Trial Completion	**DV: 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 5: Association Between Trial Delays and Drug Quality

This table examines the association between trial delays and and drug quality measures. Columns (1) and (2) use *Delay in Trial Completion* and columns (3) and (4) use *Trial Site Congestion* as the dependent variable. The two drug quality measures, *Drug Likelihood of Approval* and *Phase Transition Success Rate (PTSR)*, measure the predicted probabilities to obtain final FDA approval and to successfully complete the current phase of trial and move forward to the next one for a drug candidate, respectively. These measures 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 the 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 year-by-quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	DV: Delay in Trial Completion		**DV: Trial Site Congestion**	
	(1)	(2)	(3)	(4)
Drug Likelihood of Approval (LoA)	0.083		0.010	
	(0.79)		(0.20)	
Phase Transition Success Rate (PTSR)		0.017		0.003
		(0.38)		(0.10)
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	2,884	2,925	2,884	2,925
Adj. R-squared	0.3581	0.4877	0.3580	0.4877

Table 6: Firm- and Project Characteristics-Related Channels

This table examines 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 the 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). See Section 5.4 for further details. Standard errors are two-way clustered at the ICD category and year-by-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 7: CEO-Related Channels

This table examines 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 the 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). See Section 5.4 for further details. Standard errors are two-way clustered at the ICD category and year-by-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	New CEO Subsample	Project-Initiating CEO Subsample
	(1)	(2)	(3)
Delay in Trial Completion	−0.163***	−0.188***	−0.359***
	(−2.97)	(−3.36)	(−3.12)
CEO Delta	0.373		
	(0.27)		
Delay × *CEO Delta*	−0.081**		
	(−2.36)		
p-value for difference of coefficients on *Delay*, Columns (2) vs. (3): 0.057			
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	1,207	2,143
Adj. R-squared	0.2587	0.2219	0.3075

Table 8: Effects on Other Firm Investments

This table examines effects on new drug project initiation, using the change in trial site congestion as an instrument for delay in 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 is the total number of new projects (in logs) initiated by the company within one year of completing the focal project (i.e., the project included in the main sample). New projects are defined as a drug by indication, and include newly initiated drug projects with development status "Discovery" or "Preclinical." Columns (1) and (2) consider all new projects, while 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, zip-code-level 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 the 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 year-by-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 examines the effect of delay-induced R&D commitment on the probability of new drug launches, among drug projects with no or few existing medications (orphan drugs), using the change trial site congestion as an instrument for delay in 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, zip-code-level 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 the 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 year-by-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 examines the effects of delay-induced R&D commitment on the occurrence of adverse events in approved drugs, using the change in trial site congestion as an instrument for delay in 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, zip-code-level 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 the 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). 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(**1-Yr Adverse Events**) Second Stage	**DV:** log(**3-Yr Adverse Events**) 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

A. Theoretical Framework

This section extends the discussion of the simplified theoretical framework from Section 2. Our framework extends Li et al. (2023) by including unexpected delays in project development and differential ability of CEOs running the project. Below, we discuss two different scenarios: a frictionless scenario as a benchmark and a scenario with asymmetric information about managers' ability.

The framework is in continuous time. A pharmaceutical firm develops a drug project i which, in order to be launched, has to complete two phases (Phase I and Phase II clinical trials). Favorable trial results in each phase are required for the project to continue to the next phase or be launched. 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 each time t, and with instantaneous probability λ_j the project completes the current trial phase. Since the time to complete a stage of clinical trials follow a Poisson process, the anticipated duration for completing Phase $j \in \{I, II\}$ is $\frac{1}{\lambda_j}$. Consequently, the unexpected delay in the completion of each phase is $D_j = t - \frac{1}{\lambda_j}$, where t is the actual time to complete phase j. 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 cumulative density function $F(W)$. W is revealed after the completion of Phase I and is observed publicly. All agents are risk neutral with a discount rate r.

The firm is randomly matched with a manager who oversees the development of the drug project. Managers can be of two types: good (g) or bad (b), with the prior for a good manager in the market being $p \in (0,1)$. The manager's ability governs the instantaneous probability of the project completing each trial phase. Specifically, for a project in Phase $j \in \{I, II\}$ that 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 each trial phase and the overall project faster.

Upon the completion of Phase I trials, the manager can choose to continue or suspend the project. If the manager suspends the project, she will leave the firm 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 when Phase I trial is completed at time t. In the compensation package, a represents her base wage and b represents the sensitivity of her compensation to the project's value.

If the manager chooses to suspend the project and re-enters the labor market, her compensation is $M' = a + b\,[V - N(p_b)] + \varepsilon$, where $N(p_b)$ represents the manager's labor market search costs. $N(p_b)$ is an increasing function of the market's belief that the manager is of the bad type, motivated by the empirical evidence on managerial labor markets.[22] ε is the idiosyncratic component in the change of the manager's base wage, with CDF $K(\cdot)$. The manager observes ε before deciding to continue or suspend the project.

A.1 Frictionless Benchmark

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

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

where rV_2^s represents the flow value in the project, consisting of two components. First, c_2 are the R&D costs for Phase II. Second, with probability λ_2^s, the project completes Phase II and generates the final payoff W; thus, $W - V_2^s$ is the expected payoff jump. Rearranging this equation yields

$$V_2^s = \frac{\lambda_2^s W - c_2}{\lambda_2^s + r}. \tag{6}$$

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

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

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 forms an expected value of W.

[22] 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 signals 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, 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.

Therefore,

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

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

How does delay affect the manager's continuation decision?

Suppose the project that has completed Phase I trials and enters Phase II at time t, is run by manager s and generates final project payoff W, completed Phase I trials with an unexpected delay $D = t - \frac{1}{\lambda_1^s}$. In the frictionless benchmark case, the ability of the manager is perfectly observed in the market. According to Equation (6), the expected value of the project at that moment in time 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'.

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

$$Prob(Suspend \mid D, W, s = g) = 1 - K(bN(0)) \tag{7}$$

where $K(\cdot)$ is the cumulative density function of ε, b is the sensitivity of the manager's compensation to the project's 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

$$Prob(Suspend \mid D, W, s = b) = 1 - K(bN(1)) \tag{8}$$

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

Note that in Equations (7) and (8), the project suspension probabilities are independent of D or t for $s \in \{g,b\}$. This proves Prediction 1 of Section 2.[23]

[23] It is also interesting to consider the problem from the social planner's perspective. At time t, for a project that has already completed Phase I trials, it is optimal to suspend it if and only if it has a negative NPV, i.e., $V_2^s < 0$. This implies that the social planner will 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 (7) and (8), the project suspension probability with a social planner is also independent of the unexpected delays in the previous trial phase. We also note that the difference between the manager's and the social planner's decision rule is mechanical under the maintained assumption in our stylized framework that managers' compensation structure is linear.

A.2 An Extended Framework With Asymmetric Information

Now suppose the market cannot observe managers' ability. The only possible equilibrium is a pooling equilibrium since the bad-type manager has an incentive to mimic the good-type manager and there is no cost for them to do so. At time t, conditional on the project having completed Phase I, the market's belief of a bad-type manager is

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

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

The expected market value of a project with payoff W that has just completed Phase I and enter 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 $[1-A(t)]\lambda_2^g$ represents the probability of completing Phase II conditional the probability of a good-type manager. This implies

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

Analyzing managers' decision of whether to suspend the project, we compare their compensation packages in the asymmetric information scenario. Given a project W having completed Phase I with an unexpected 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 is

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

If the manager suspends 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 $\varepsilon > bN(A(t))$. Consequently, the probability of suspending the project is

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

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

Prediction 2: *With asymmetric information about managerial ability, the suspension probability following the completion of the first project phase is (1) decreasing in the unexpected delay of the already-completed trial stage D; and (2) decreasing in the sensitivity of the CEO's compensation with respect to the firm's stock price given a level of unexpected delay in the already-completed trial phase.*

B. Data and Estimation Appendix

B.1 Additional Detail on Data Collection 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 include additional information from ClinicalTrials.gov (number of participants, trial start and end dates), and compare the added information to that from our main dataset. This allows us to determine the set of correctly matched trials. Before further assessing the accuracy of each match, as described 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.[24]

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 (we determine a match as correct when, e.g., 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 (we determine a match as incorrect when, e.g., the anticipated completion date from ClinicalTrials.gov precedes the start date of the trial in our main dataset), we manually assess the match, that is, we assess whether, e.g., the sponsors and collaborators match, taking into account acquisitions, joint ventures, and name changes.

Table OA.1 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 candidate matches (adding information

[24] 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 is adjusted in subsequent records.

from ClinicalTrials.gov, scraping NCT identifier, and assessing correctness of each match).

Table OA.1: 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

B.2 Additional Detail on Matching With FDA Drug Adverse Events Database

For the clinical trials in our main analysis sample, we identify a subset of trials where the experimental drugs ultimately receive FDA approval and are launched in the market. We manually gather the brand names of these drugs by searching their experimental names used during the trials (typically containing the active ingredients or chemical compounds of the drugs) through online searches. To ensure accuracy in identifying the correct brand names, we also cross-reference the drug developers and disease indications by searching relevant information from the internet, including prominent websites such as Drug Bank and Adis Insight.

We then use both the brand names and experimental names of drugs from the clinical trials to match with the drug names listed in the FDA Drug Adverse Events Database. We include the drugs' experimental names (such as active ingredients) in the matching process, as they sometimes appear in the adverse event reports. However, due to the presence of biosimilars and generic drugs, for which the adverse events database also lists the corresponding active ingredients, we take additional steps to ensure the quality of the matching process. Specifically, we determine 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 with the drug sponsor names in the trial data.

After obtaining the matched adverse events sample, we restrict to the incidents where drugs are listed as the primary suspect. We exclude incidents with any missing information on incident dates or incident primary ID. Ultimately, we are able to match 3,890,964 adverse events with the drug projects that received FDA approval in our clinical trials data.

B.3 Additional Detail on Construction of Alternative Instrument (Excluding Same-ICD-Category Patient Flows)

To exclude patient flows from trials in the same ICD category in the construction of our alternative instrument, we proceed in five steps as described below. These steps are necessary because, although ICD designations are readily available for our main sample, they are not provided by ClinicalTrials.gov for the universe of trials forming the basis of our instrument construction. For the universe of trials from ClinicalTrials.gov, we instead have a written description summarizing each project.

First, we begin with a list of key words describing each ICD code, identifying those key words that are sufficiently rare to distinguish a small number of ICD codes. Specifically, we focus on key words associated with at most ten unique ICD codes. (For concrete examples, we omit common keywords such as "infection" or "syndrome" as they are too ubiquitous to be effective for ICD code mapping.) Second, we further eliminate key words that are commonly used in general language and might appear in a context unrelated to disease in the trial descriptions from ClinicalTrials.gov (e.g., "complex," "stem," "status"). Third, for each drug project in the universe of trials on ClinicalTrials.gov used in our main instrument construction, we verify whether any of the main text items on ClinicalTrials.gov (official title, brief title, brief summary, detailed description) contain any of the final key words associated with a given ICD code. If such key words are present, we assign the project the corresponding ICD code. (Following this procedure, a project may be assigned multiple ICD codes. This is a more conservative approach.) Fourth, for each ICD code, we construct an ICD-code-specific instrument using the subset of ClinicalTrials.gov projects that are *not* assigned the respective ICD code in the previous step. Finally, for each project in our main sample, we use the available ICD designation to identify the corresponding ICD-code-specific instrument for analysis.

B.4 Additional Detail on Validation Exercise Using Exchange-Rate-Induced Variation in Trial Costs

Here, we provide additional detail on the construction of the measure capturing exchange-rate-induced trial cost fluctuations, as well as on the corresponding estimation. We measure the extent to which foreign currency becomes more expensive (or cheaper) across the duration of a clinical trial by taking 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. We then construct an average exchange rate change for clinical trial i as

$$\Delta FX_i = \sum_c \frac{n_{ic}}{N_i} \sum_{\tau=1}^{T} \Delta FX^{c}_{[\tau]} \tag{9}$$

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-sites-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 $\Delta FX_i = 0$. Among foreign trials, the interquartile range in exchange variation is slightly above six percentage points in our data.

Similar to Equation (1), we then estimate:

$$\text{Suspension}_i = \beta \Delta FX_i + \gamma \text{Controls}_i + \text{FEs} + \eta_i \tag{10}$$

to study how firms' trial continuation decisions are affected by prior exchange-rate-induced investments made in the clinical trial project, with Online Appendix Table OA.12 reporting the results. Regarding the results shown in the table, 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 in the table solely for visual purposes. Further, all columns in Appendix Table OA.12 control for a trial's total duration, to account for the fact that total trial duration determines the period over which exchange rate fluctuations unfold and affect financial obligations.

B.5 Additional Detail on Patient Outcomes Estimation Procedure

Our derivation of the estimation equation in Section 7.2.1 follows Wooldridge (2010). Let us rewrite Equation (3) such 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 can estimate the above equation by 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,

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

Given the assumption in Equation (5), 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 to Equation (11),

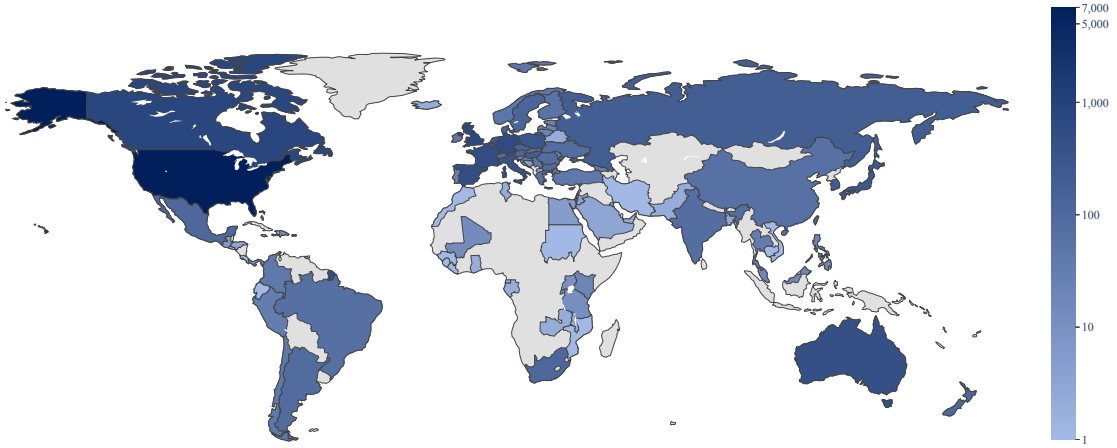
$$
\begin{aligned}
g(z_i, \mathbf{X}_i, Y_i^{app} = 1) &= E[E[v_i | z_i, \mathbf{X}_i; \beta_1 \xi_i + u_i] \mid z_i, \mathbf{X}_i; \beta_1 \xi_i + u_i > -\beta_1 \gamma_1 z_i - (\beta_1 \gamma_2' + \beta_2')\mathbf{X}_i] \\
&= \alpha_3 E[\beta_1 \xi_i + u_i \mid z_i, \mathbf{X}_i; \beta_1 \xi_i + u_i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

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.

C. Supplementary Results

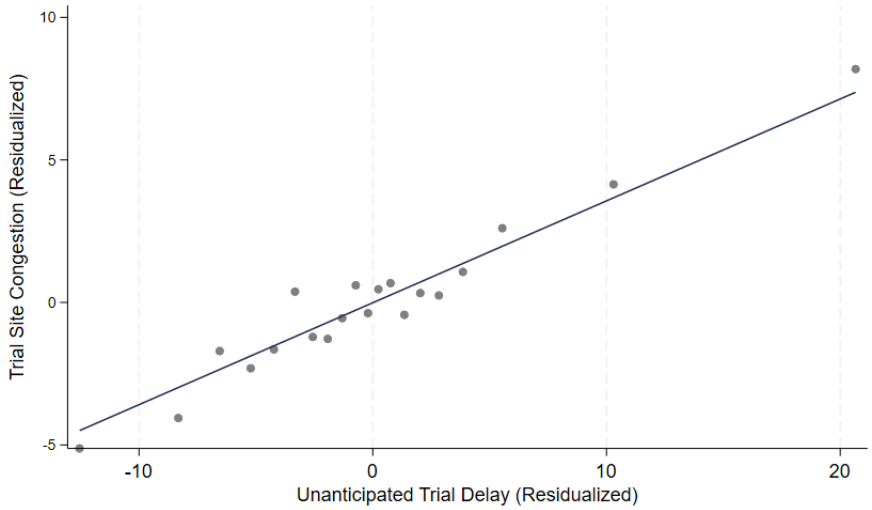
C.1 Figures

Figure OA.1: Geographic Distribution of Clinical Trials



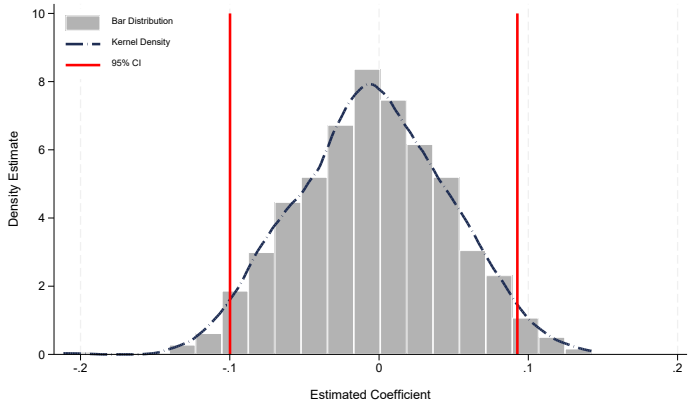
This figure shows the geographic distribution of clinical trials in our analysis sample. For each country, we calculate the natural logarithm of the number of trials.

Figure OA.2: First Stage

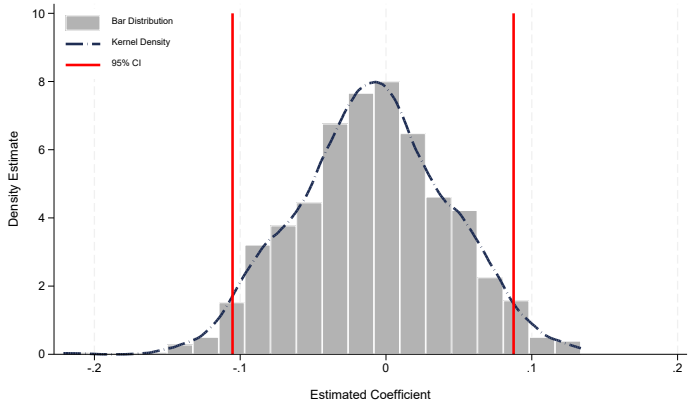


This figure shows a binscatter plot visualizing the relationship between unanticipated clinical trial delays and trial site congestion (IVs). Both the delay and trial site congestion are residualized with respect to the control variables and fixed effects in column (2) of Table 3.

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

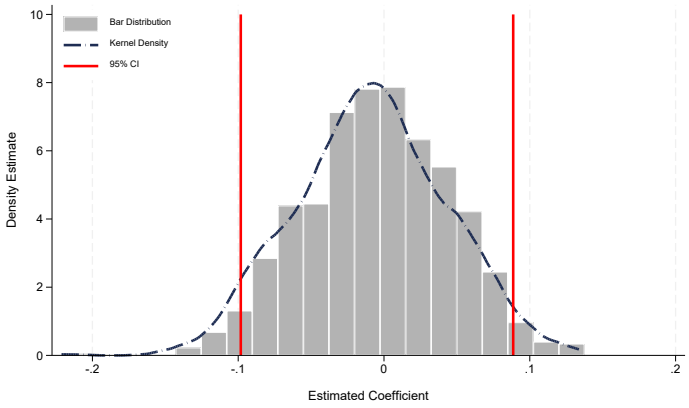


(a) Distribution of Estimated Coefficient of ΔFX without Controls

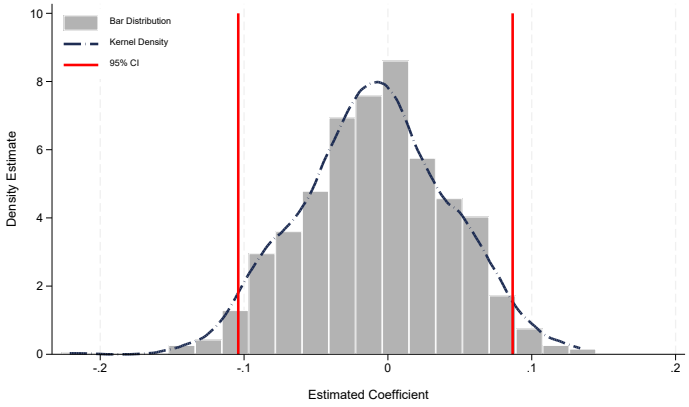


(b) Distribution of Estimated Coefficient of ΔFX with Controls

Figure OA.3: ***continued***



(c) Distribution of Estimated Coefficient of $\Delta FX \times ForeignTrial$ without Controls



(d) Distribution of Estimated Coefficient of $\Delta FX \times ForeignTrial$ with Controls

This figure shows the distribution of estimated placebo coefficients for ΔFX and $\Delta FX \times ForeignTrial$ corresponding to Table OA.12, where 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. Panels A and C show the estimation without controls. Panels B and D show the estimation with controls.

C.2 Tables

Table OA.2: 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 the 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 year-by-quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: Anticipated Duration (Months)							
	(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.3: IV Robustness: Controlling for Number of Competing Drug Projects Nonparametrically

This table replicates the specification in Columns (2) and (4) of Table 3 except that we control for the number of competing drug projects in the same ICD category nonparametrically through number-of-competing-drug decile fixed effects. Columns (2) and (3) repeat the estimation separately for the subsamples with below-median and above-median drug project competition, respectively. All other details are the same as in Table 3.

	DV: Suspension		
	Full Sample	No. Competing Projects $\leq p50$	No. Competing Projects $> p50$
	(1)	(2)	(3)
Delay in Trial Completion	-0.530^{***}	-0.539^{**}	-0.542^{**}
	(-2.93)	(-2.40)	(-2.05)
Controls	Y	Y	Y
No. Competing Projects Decile FE	Y	Y	Y
Year $\times$ Quarter FE	Y	Y	Y
Firm FE	Y	Y	Y
ICD FE	Y	Y	Y
Trial Phase FE	Y	Y	Y
First-stage F-stat	18.95	16.10	15.78
Observations	9,161	4,437	4,438

Table OA.4: IV Robustness: Excluding Same-ICD-Category Patient Flows from Congestion Measure

This table replicates Table 3 except that we modify the construction of the instrument, measuring the change in trial site congestion excluding patient flows from drugs in the same ICD category (see Section 5.2 and Online Appendix Section B.3 for details). All other details are the same as in Table 3.

	DV: Delay First Stage		**DV: Suspension** Second Stage	
	(1)	(2)	(3)	(4)
Delay in Trial Completion			-0.432^{**}	-0.672^{***}
			(-2.55)	(-3.21)
Other-ICD Trial Site Congestion	0.377^{***}	0.374^{***}		
	(4.09)	(4.09)		
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
F-stat	16.75	16.72	–	–
Observations	9,397	9,161	9,397	9,161

Table OA.5: Trial Delays, Expected Drug Sales, and Patient Adverse Events During Clinical Trials

This table examines the relationship between trial delays, expected drug sales, and patient adverse events during clinical trials. Panel A examines the relationship between the log of expected drug sales (in million U.S. dollars adjusted to 2017 by GDP deflators) 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. Sales are the sum of expected sales during the initial five years after drug launch and the expected average annual sales after drug launch (averaging over all years with expected sales data), respectively. Panel B examines the relationship between the occurrence of adverse reactions among trial participants for project i that already completed Phase I or II clinical trials and $Delay in Trial Completion_i$. Occurrence of adverse events is the total percentage of severe (life-threatening) adverse reactions (i.e., summing over the ratios of the number of participants experiencing life-threatening reaction $r \in R$ to the total enrolled participants across all life-threatening reactions R) and the maximum percentage of severe adverse reactions, respectively. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in the 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 year-by-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		
	DV: Five-Year Sales	**DV: Average Annual Sales**
	(1)	(2)
Delay in Trial Completion	-0.910×10^{-4}	4.756×10^{-4}
	(-0.10)	(0.41)
Controls	Y	Y
Year × Quarter FE	Y	Y
Firm FE	Y	Y
ICD FE	Y	Y
Trial Phase FE	Y	Y
Observations	3,611	3,951
Adj. R-squared	0.5576	0.5722
Panel B: Adverse Reactions in Trials		
	DV: Total Severe Reaction	**DV: Max Severe Reaction**
	(1)	(2)
Delay in Trial Completion	0.002	0.003
	(0.49)	(0.91)
Controls	Y	Y
Year × Quarter FE	Y	Y
Firm FE	Y	Y
ICD FE	Y	Y
Trial Phase FE	Y	Y
Observations	9,161	9,161
Adj. R-squared	−0.0547	−0.0557

Table OA.6: Controlling for Firm-by-ICD Fixed Effects

This table replicates Columns (3) and (4) in Table 2 except that we include firm-by-ICD-code fixed effects. All other details are the same as in Table 2.

	Dependent Variable: Suspension	
	(1)	(2)
Delay in Trial Completion	-0.186^{***}	-0.174^{***}
	(-6.21)	(-5.62)
Controls	N	Y
Year $\times$ Quarter FE	Y	Y
Firm $\times$ ICD FE	Y	Y
Trial Phase FE	Y	Y
Observations	7,546	7,337
Adj. R-squared	0.5010	0.5005

Table OA.7: Controlling for Firm-by-ICD Fixed Effects: IV

This table replicates Table 3 except that we include firm-by-ICD-code fixed effects. All other details are the same as in Table 3.

	DV: Delay First Stage		**DV: Suspension** Second Stage	
	(1)	(2)	(3)	(4)
Trial Site Congestion	0.361^{***}	0.372^{***}		
	(4.50)	(4.61)		
Delay in Trial Completion			-0.214	-0.373^{**}
			(-1.47)	(-2.14)
Controls	N	Y	N	Y
Year × Quarter FE	Y	Y	Y	Y
Firm × ICD FE	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y
F-stat	20.28	21.27	–	–
Observations	7,546	7,337	7,546	7,337

Table OA.8: Controlling for 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 third column of Table 6. 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 the 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). 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.9: Trial Duration and FDA Clinical Holds

This table examines the association between trial delays and FDA 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, with the main independent variable of interest being an indicator variable for whether project *i* experienced a clinical hold during the already completed Phase I or II trials. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in the 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). 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.10: Positive Vs. Negative Project Delays

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 the 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). 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 × 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.11: 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 the 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). 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

Table OA.12: Exchange-Rate-Induced Trial Cost Changes and Project Continuation Versus Suspension

This table reports the OLS estimation results of Equation (10). 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 (9). $Trial\,Duration_i$, 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 the 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 year-by-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