

Excess Commitment in R&D*

Marius Guenzel

Tong Liu

May 2023

Please do not distribute without permission.

Abstract

We investigate how excess commitment to R&D activities impacts innovation outcomes and consumer welfare. Using project-level data on clinical trials by pharmaceutical firms, we document that trials which faced *unanticipated* delays in the previous trial stage, relative to initial firm expectations or induced by plausibly-exogenous trial site congestion, are significantly less likely to be subsequently suspended. These effects are amplified when the firm's CEO has a greater wealth exposure to stock price changes and is personally responsible for the project initiation. Consumers may, in some ways, *benefit* from firms' excess commitment in new drug development. Marginally-launched drugs because of commitment distortions are not significantly associated with more adverse events, but are significantly more likely to target diseases with no or only few existing medications in the marketplace (orphan drugs).

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

*Guenzel is at the Wharton School, University of Pennsylvania. Liu is at the MIT Sloan School of Management. Emails: mguenzel@wharton.upenn.edu, tongl@mit.edu. We are grateful for comments from Amber Batata, Ashley Litwin, Andrew Lo, Deborah Lucas, Antoinette Schoar, Jerome Taillard, and Richard Thakor (discussant), as well as seminar and conference participants at MIT Sloan, the Virtual Corporate Finance Seminar, and the 2023 Kentucky Finance Conference. 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 and central to economic growth. One defining feature of innovation and R&D is the frequent need for a long-term commitment to a given course of action in order to achieve breakthroughs. While this suggests a positive influence of commitment on innovation, seminal prior work, in particular [Staw \(1976\)](#)'s evidence from the lab, points to a potentially large role of *excess* commitment to previously chosen R&D activities. Such excess commitment may arise from misaligned incentives as well as behavioral biases of decision-makers.

Despite the critical role of commitment in R&D, we still lack a comprehensive understanding of how decision-makers determine their level of commitment to ongoing projects. It is even less clear how firms' commitment levels to their R&D endeavors impact aggregate welfare. For example, does elevated commitment lead to inferior or even unsafe innovation, or does it drive the creation of otherwise nonexisting innovation that benefits consumers?

There are at least two reasons for why studying the level of commitment and path-dependent decision-making with respect to innovation, as well as resulting welfare consequences, is challenging. First, it requires granular data on R&D activities at the project level, including observable information on R&D milestone achievements and final R&D outcomes. Second, it requires not only reliable measures capturing the level of commitment to a certain R&D activity over time, but also a setting in which there is plausibly exogenous variation in the intensity of initial R&D commitments made.

In this paper, we overcome these twin challenges related to data and identification in the important setting of new drug development by pharmaceutical firms. Drug projects in clinical trials offer a near-ideal, project-level setting of R&D investment with detailed observable and quantifiable outcomes. Additionally, in this setting, we can isolate plausibly-exogenous variation in R&D commitment to drug projects. Specifically, we study whether *unanticipated* delays in trial completion (of the previous trial stage) distort firms' decision to suspend versus advance the project through the next trial phase. We document substantial excess commitment by firms in response to trial delays, and quantify the impact on ultimate R&D outcomes, in particular patient outcomes of approved drugs. One broad-level takeaway from our results is that firms' excess commitment can have positive externalities on consumers.

To test for and study the effects of firms' (excess) commitment to R&D activities, we assemble a new

dataset of clinical trials initiated by U.S.-based companies, with detailed information on trial timelines (start date, end date, and, importantly, *anticipated* end date as filed by the drug sponsor on ClinicalTrials.gov), trial sites (both domestic and international sites), trial participants (patient eligibility, number of participants, trial sponsors), drug details (drug indications, ICD category, etc.), post-trial-completion outcomes (project continuation versus suspension), and ultimate drug outcomes (drug launches, sales, regulatory designations, and adverse events). The data come from a variety of sources, including, among others, Cortellis Clinical Trial Intelligence, web-scraped and hand-matched data from ClinicalTrials.gov, and the FDA Adverse Event Reporting System database. Our final sample is comprised of more than 10,000 clinical trial projects initiated between 1985 and 2019 and completed between 1991 and 2020, with trial sites spanning 91 countries including the U.S.

We organize our analysis and findings into four main parts. In the first part of the paper, we document a strong link between firms' initial and follow-on commitment to R&D activities. We leverage the information on drug firms' self-reported anticipated trial end dates—data which, to our knowledge, have not been explored in prior work—and study how unexpected delays in trial completion affect subsequent decision-making. The average trial completion in our data is delayed by nearly one year, with the 25th (75th) percentile trial being not delayed (delayed by 1.6 years). Unanticipated trial delays have substantial predictive power for firms' post-trial-completion decision to suspend versus continue the drug development in the next-stage trial. Consistent with path dependence in R&D decision-making, already-completed trials that faced an unexpected delay are significantly less likely to be subsequently suspended, with a one standard deviation increase in delay reducing the suspension probability by 4 percentage points, or 14% relative to the baseline suspension probability of 31%.

In the second part of the paper, we find that this path dependence in R&D decision-making is most consistent with firms exhibiting excess commitment to R&D projects, as supported by two sets of additional analyses. Both analyses assess firms' excess commitment behavior by examining the effect of exogenously-induced variation in initial trial commitments on subsequent decision-making.

First, we implement an instrument variables (IV) approach that exploits the fact that trial sites have limited capacity to accommodate clinical trials. When more trials take place simultaneously in a given location, bottlenecks are more likely to occur. Leveraging this, we instrument for unanticipated trial

completion delay with a measure capturing *trial site congestion*. This measure is the normalized change in the average patient enrollment speed of a zip code where the trial takes place between trial start and end date. The identifying assumption is that changes in trial site congestion affect firms' decisions to advance new drug candidates through the next clinical trial phase only through the effect of trial completion delay. We argue this exclusion restriction plausibly holds since other potential factors, such as trial site quality proxied by hospital care quality provided in a region, are empirically uncorrelated with the congestion measure. Trial site congestion and delay are significantly positively related, confirming the existence of a strong first stage. In the second stage, delay in trial completion when instrumented with trial site congestion continues to significantly affect the decision to continue versus suspend the trial, consistent with excess commitment. The IV-based economic magnitude of the effect is similar to the OLS estimates.

Second, we validate the findings based on the delay and time-based commitment measures with an additional cost-based commitment measure, exploiting fluctuations in exchange rates in foreign-based (i.e., outside the U.S.) trials between trial start and end date. Exchange rate movements induce plausibly-exogenous variation in firms' monetary commitments to a given trial, as the contracts and actual payments are typically conducted in local currency.¹ Again, we find a significant influence of commitments made by a firm to a trial, now arising from exchange rate movements, on subsequent commitment to advance the trial through the next phase. In the exchange rate context, we estimate an excess commitment effect, in response to a plausibly-exogenous increase in trial costs by one standard deviation, amounting to a 6% reduction in trial suspension probability.

In the third part of the paper, we investigate underlying channels for the positive link between trial delays and subsequent trial suspension decisions. The results further reinforce the interpretation that the observed R&D path dependence reflects excess commitment. Our findings are not driven by early-stage firms with no or few alternative drugs as in [Guedj and Scharfstein \(2004\)](#). They are also not driven by firms being unable to pivot due to financial constraints as gauged by a series of common financial constraint measures. Differential sorting by firms into trials with different lengths or delays based on initial expectations is equally unlikely to be an underlying mechanism, since we can control for trial

¹See, for example, <https://www.appliedclinicaltrials.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."

length expectations. Moreover, the path dependence in decision-making in our setting is unlikely to reflect a real options channel, since delayed drugs are not associated with favorable outcomes for firms such as drug sales upon drug approval and launch.

Instead, the effect of unexpected trial delays on subsequent trial commitment is amplified in the presence of management-related frictions. For one, delay is substantially more predictive of continuation among chief executive officers (CEOs) with greater wealth exposure to stock price changes (Δ), 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, R&D path dependence is significantly greater when the CEO in charge remains the same between trial start and end, consistent with personal responsibility for initiating the project playing an important mediating role.

In the final part of the paper, we evaluate the welfare implications of our findings, particularly for consumers, shedding light on various welfare-relevant outcomes. These include the occurrence of adverse events in patients, as well as the impact of the excess commitment in new drug development on the availability of drugs in the marketplace. With respect to adverse effects, we find that marginally-approved drugs due to delay-induced firm behavior is associated with economically modest but statistically insignificant increases in harmful events. At the same time, with respect to drug availability, we find that delay-induced increased commitment to trials increases the probability that firms ultimately bring the drug to market among drugs with no or few existing treatment options, as gauged by a drug's designation as an orphan drug.² Taken together, these findings suggest a nuanced relation between firms' excess commitment and welfare implications for consumers.

Overall, we document significant path dependence and excess commitment in firms' R&D decision-making. From a welfare perspective, this excess commitment in new drug development plausibly benefits consumers in important ways, for instance, through greater availability of orphan drugs. Our findings complement a large literature and active debate about the extent to which there is adequate innovation activity in the aggregate economy. For instance, seminal prior work has uncovered many reasons for why there might be aggregate under-investment in innovation, including low spillover effects, low competition, patent protection, infringement lawsuits, and risk aversion (Hall and Lerner, 2010). Our contribution to

²See, e.g., <https://www.fda.gov/news-events/fda-voices/rare-disease-day-2021-fda-shows-sustained-support-rare-disease-product-development-during-public>.

this debate is unique, in that we examine the level of 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 the market failures in the markets for innovation and intellectual property. In the context of pharmaceutical innovations, Acemoglu and Linn (2004) investigate the effect of potential market size on drug innovations. Budish et al. (2015) show that short-termism and the fixed patent term distort private research investments. Azoulay et al. (2019) study the impact of scientific grant funding at the NIH on pharmaceutical innovations. Krieger et al. (2022a) study how risk aversion leads pharmaceutical firms to underinvest in radical innovation. Thakor and Lo (2022) investigate how product market competition affects R&D investment and financing. Li et al. (2023) show that common ownership among venture capital firms could be beneficial to innovation efficiency.³ Departing from the previous work, this paper studies how firms' (excess) commitments to the R&D process, an organizational response within firms, distort their innovation efforts. In particular, we discuss the potential welfare consequences of these distortions, and challenge the conventional wisdom that R&D distortions are typically and unequivocally bad for consumers.

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

³With respect to biomedical innovation, Lo and Thakor (2022) and Lo and Thakor (2023) review the recent literature on how external financing frictions affect drug development, including Lerner et al. (2003) on equity financing cycles, Robinson and Stuart (2007) on financial contracting in strategic alliances, Lerner and Malmendier (2010) on contractibility in research agreements, Cunningham et al. (2021) on the M&A market, Aghamolla and Thakor (2022) and Liu (2021) on the IPO decisions of biotech, and Krieger et al. (2022b) on profit erosion of existing products.

its importance, studying welfare implications of firm investment distortions is still mostly absent in the literature on nonstandard firm decision-making (see [Malmendier \(2018\)](#) and [Guenzel and Malmendier \(2020\)](#) for recent literature surveys).

Third, we contribute to a literature in economics and related fields on the effects of commitment and escalation of commitment. As alluded to above, in an influential laboratory study on commitment and path dependence, [Staw \(1976\)](#) finds that subjects are more committed to a chosen (hypothetical) R&D project if they were personally responsible for the initial action and, consistent with an escalation mechanism, in response to negative signals about the investment. Subsequent work has provided empirical evidence consistent with escalating commitment in various contexts including “problem loan” write-offs ([Staw et al., 1997](#)) and NBA draft picks ([Staw and Hoang, 1995](#); [Camerer and Weber, 1999](#)). A common problem in empirical studies is the lack of plausibly-exogenous variation in initial commitment decisions, which opens the door for alternative explanations related to information, beliefs, or selection to explain subsequent commitment. We advance this literature by deliberately focusing on how *unanticipated* factors, delays and cost variation, affect initial and distort subsequent commitment.⁴

The remainder of the paper is structured as follows. Section 2 provides details on institutional background and data. Section 3 introduces the empirical approach. Sections 4 and 5 discuss the results and welfare implications. Section 6 concludes.

2 Data and Institutional Background

2.1 Institutional Background

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

⁴We note two papers that use price variation induced by auctions to study commitment in *consumer*, rather than firm, settings ([Augenblick, 2016](#) in consumers’ penny auction behavior; [Ho et al., 2018](#) in consumers’ car usage behavior). [Guenzel \(2023\)](#) exploits post-acquisition-agreement aggregate stock market fluctuations to obtain quasi-random variation in sunk acquisition costs. We further note the existence of a literature on R&D focusing on how commitments by some firms can lead to “path dependence” in a *different* sense, namely creating barriers to operate for *other* firms (e.g., [Manez et al., 2009](#)). In contrast to such studies, we focus on *within-firm* commitment effects.

devoted to conducting clinical trials.

Clinical trials are regulated by the FDA and typically involve three stages: Phase I, II, and III trials. Phase I tests the safety of a drug candidate on a small group of humans and determines how and where it distributes within the body. Phase II involves a larger number of patients, aiming to determine the dosage and effectiveness of the drug. Phase III primarily focuses on the safety and efficacy for a wide variety of population by testing the drug candidate on a much larger group of patients. After completing all three phases, a drug developer (sponsor) can submit the new drug application to the FDA to seek the market launch permission. For each stage of trials, a sponsor needs to design a strict protocol—a scientific plan of action. The protocol explains how many patients will be recruited, the inclusion/exclusion criteria for eligible patients, what tests will be performed, how results will be measured and collected, and why the trial may be discontinued. Clinical trials are often conducted in different settings and locations. These locations are across cities, states, and countries.

Clinical trials are costly. Recent studies reveal that it costs over \$1 billion to bring a new drug to the market (DiMasi et al., 2003; DiMasi et al., 2016; Wouters et al., 2020). A variety of factors significantly affect the final cost of a clinical trial, including but not limited to clinical procedure costs, administrative staff costs, site monitoring costs, site retention costs, and central laboratory costs (Wong et al., 2014). Noticeably, these cost factors heavily hinge on the length of time to complete a trial. For example, longer timelines mean higher labor costs as investigators and staff must be compensated for longer hours. Longer timelines also imply more devoted hours to take care of trial participants, which leads to higher costs to drug developers. As documented by Wong et al. (2019), the median duration of Phase I, II, and III are 1.6 years, 2.9 years, and 3.8 years,⁵ In our sample, whose construction we describe next, the median trial (comprised of Phase I and II trials) similarly takes 2.4 years to complete (Table 1; see Section 2.3 for further summary statistics).

2.2 Data and Sample Construction

Clinical Trials Data. We construct a sample of clinical trials initiated by U.S.-based companies. The trials data come from Cortellis Clinical Trial Intelligence, in which we can observe detailed information

⁵However, the time to complete a trial varies widely across indications. For example, the median duration of Phase III oncology trials is about six years.

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

Data Filters. We apply several filters to construct our final analysis sample. First, we focus on projects with already-completed Phase I or II trials and non-missing records on their outcomes after these trials, e.g., suspension or continuation. Second, we drop any projects whose Phase I or II trials were completed after 2020, because these projects would have insufficient time to reach any outcomes. Third, we drop any projects in Phase I or II trials where their endpoints were not achieved, so that the suspension in the next phase of trials is not due to the unfavorable readouts in the previous trials. Next, we require a project to have a realized trial start and completion dates and hence drop those with estimated completion dates. These filters yield a sample of 11,228 projects, among which trials were initiated between 1985 and 2019 (completed years between 1991 and 2020) and trial sites were spanning across 91 countries including the U.S.

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

provided in Section II.A in the Online Appendix.

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

Other Data. Lastly, we obtain financial accounting and management information for a subset of the drug developers that are public pharmaceutical companies traded on U.S. exchanges from Compustat, BoardEx, and Execucomp. We get the expected drug revenues once launching in the market from the Cortellis Competitive Intelligence. We collect the quality information of U.S. hospitals from various CMS-managed quality programs, including 30-day mortality rates, 30-day readmission rates, and patient safety indicators. Clinical hold records are collected from the ClinicalTrials.gov. Exchange rates data are from Bloomberg.

2.3 Summary Statistics

Table 1 presents summary statistics for our main sample comprised of completed clinical trials. Just under 60% of trials in our sample are Phase II trials, the remainder are Phase I trials. About one in three trials are suspended, i.e., not advanced through Phase II or III, respectively. 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. The average firm in our sample has 136 drug projects (drug by ICD indications) under development, i.e., our sample firms are *not* single-product firms as analyzed in Guedj and Scharfstein (2004). We discuss variables related to trial duration and exchange rates below in Sections 3.1 to 3.3, when we introduce the empirical design. We also note that the number of observations in the regression results will sometimes differ from the summary statistics numbers in Table 1, depending on the availability of the duration and exchange rate variables.

Figure 1 shows the geographic distribution of clinical trial sites in our sample. Clinical trials have gone global in the recent decades, with trial sites being widely spread across countries. Though a large

share of clinical trials are clustered in the U.S, accounting for 63% of the total trial sites (59,617 out of 94,261 clinical trial sites), other countries important including Canada, Germany, France and Japan contribute another 12% of the total trial sites. This pattern is robust to tabulating the number of trials, rather than trial sites, across countries (Figure OA.1 of the Online Appendix). In terms of number of trials, the U.S. accounts for 45% of observations (6,847 out of 15,222 clinical trials). Canada, Germany, France, and Japan contribute 16%.

3 Empirical Approach

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

As a validation exercise, we also analyze the effect of trial cost increases of foreign-based trials due to exchange rate fluctuations on subsequent trial commitment. That is, in a nutshell, we ask whether (exogenously) higher levels of prior commitment of a firm to a clinical trial creates path dependence in decision-making and increased commitment going forward.

3.1 Delay in Trial Completion and Project Continuation Versus Suspension

We first examine how firms factor unexpected delays in clinical trial completion into their R&D decision-making regarding continuation of the trial through the next trial phase, by estimating:

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

Suspension_i is an indicator variable for whether project i was suspended in the follow-on trial phase (Phase II or III clinical trial) after completing Phase I (II) clinical trials. The sample includes all projects i that already completed Phases I or II with detailed clinical trials records. The independent variable, $\text{DelayInTrialCompletion}_i$, is the difference between actual trial completion date and the *anticipated* completion date as indicated by the firm at the start of the trial. In our sample, the average trial completion

is nearly one year later relative to the initially anticipated end date (Table 1).

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

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

As mentioned above, our *Delay in Trial Completion* measure captures the unanticipated component of the length of a given clinical trial, as the difference between realized end date and anticipated end date filed by the firm with ClinicalTrials.gov at trial start. To further isolate the effects of plausibly-exogenous variation in trial length on firm continuation-versus-suspension decision-making, and assess the degree of excess commitment in R&D decision-making, we introduce an instrumental variables (IV) strategy that uses *trial site congestion* as an instrument for trial completion delay. This instrument is based on the notion that trial sites have limited capacity to accommodate clinical trials. If a trial site hosts too many trials, it is challenging to recruit patients on time and monitor multiple trials simultaneously, which increases the likelihood of a slowdown of trials.

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

trial completion year τ_2 :

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

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

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

The instrument takes the maximum across trial i 's sites because a clinical trial is not completed unless the last trial site is closed. For the average trial in the sample, the increase in trial congestion between trial start and end, normalized by the mean congestion, is slightly more than five percent (Table 1).

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

3.3 Exchange Rate Variation in Foreign Trials

As a validation exercise, we also analyze a second measure of plausibly-exogenous prior commitments to a clinical trial that is independent of the time-based delay and congestion measures. Specifically, we leverage the institutional feature that a large portion of clinical trials are conducted outside of the U.S. For a trial involving multiple sites internationally, the contracts and actual payments are conducted in other currencies rather than the U.S. dollar. 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.

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

$$\Delta FX_i = \sum_c \frac{n_{ic}}{N_i} \sum_{\tau=1}^T \Delta FX_{[\tau]}^c \quad (2)$$

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

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

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

to study how firms' trial continuation decisions are affected by prior exchange-rate-induced commitments to clinical trials.

4 Results

4.1 Delay in Trial Completion and Project Continuation Versus Suspension

Table 2 presents the results for how firms' decision to continue versus suspend a clinical trial after a given completed phase depends on the experienced delay in trial completion. Column (1), comparing trials initiated in the same quarter, finds a strong negative effect of delay on the likelihood of suspension, both economically and statistically speaking. A one standard deviation increase in delay reduces the suspension probability by 4.4 percentage points, or 14% relative to the baseline suspension probability of 31% (Table 1).

The effect of trial completion delay on trial continuation remains unchanged in Column (2) when we add firm fixed effects, as well as in Column (3) when we further add fixed effects for the drug's ICD category and for the trial phase (Phase I versus II trials). Finally, the effect remains unchanged in Column (4) when we add control variables for the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs).⁶

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 commitment by firms to a given clinical trial, in the form of time it takes to complete the trial *beyond* what firms anticipated at trial start, induces path dependence and increases subsequent commitment by firms when it comes to the decision to suspend versus continue the trial.

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

Table 3 presents the IV results, studying the effect of trial completion delay due to 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 trial site congestion as detailed in Section 3.2, with and without

⁶The coefficients on controls (unreported) are intuitive. More projects in the pipeline and more competing projects are positively (but insignificantly) associated with suspension. More trial sites increases the likelihood of suspension, whereas more trial participants reduces it.

the controls included in the last column of Table 2. The columns reveal a strong association between congestion and delay. Economically, a one standard deviation increase in the congestion instrument is associated with an increase in trial completion delay of close to four months. Statistically, the Kleibergen and Paap (2006) F -statistic is clearly above the common threshold for weak instruments of 10 (Stock et al., 2002).

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

4.3 Exchange Rate Variation in Foreign Trials

As a validation exercise for the delay-based measure and the interpretation that the previous results reflect excess commitment to R&D projects, Table 4 shows the results for how firms' decision to continue versus suspend a clinical trial after a given completed phase depends on plausibly-exogenous exchange rate fluctuations experienced in foreign-based trials between trial start and completion. Columns (1) and (2) are estimated on the subsample of foreign-based trials, with Column (2) adding controls as in Column (5) of Table 2. Columns (3) and (4) are estimated on the full sample including U.S.-based trials, and interact the exchange rate fluctuation variable, as well as controls and the drug indication and year-quarter fixed effects with an indicator for foreign-based trials.⁷ All columns include a control for a trial's total duration, given that total trial duration determines the period over which exchange rate fluctuations can affect financial obligations.

Across columns, the coefficient on exchange rate changes is negative, implying exchange-rate-driven cost increases make trial continuation more likely, and statistically significant. In terms of magnitudes, a

⁷We note that since $\Delta FX = 0$ for domestic trials, it is in fact the case that $\Delta FX = \Delta FX \times \text{Foreign Trial}$, i.e., we adopt the interaction term notation solely for visual purposes.

one standard deviation increase in trial costs through exchange rates is estimated to reduce the likelihood of subsequent trial suspension by 1.75 percentage points, or 6% relative to the baseline. These estimated magnitudes are slightly lower compared to the effects of unexpected delays on suspension decisions, but still economically meaningful.

To further probe the robustness of the exchange rate results, in the Internet Appendix we conduct a placebo test by randomly assigning trial countries to trials in the sample and re-running the full-sample regressions in Table 4. Figure OA.2 of the Online Appendix plots the distribution of the estimated coefficients for ΔFX_i by repeating the estimation 1,000 times. In contrast to the results in Table 4, the placebo test coefficients are not significantly different from zero, and not as large in economic magnitudes.

Overall, the results exploiting exchange rate variation provide additional evidence in favor of the idea that plausibly-exogenous increases in prior commitment to a chosen action induce distortions and excess commitment in firms' R&D decision-making.

4.4 Robustness: FDA Clinical Holds

Despite the corroborating trial-cost-based evidence in the preceding section, one potential concern regarding the main, delay-based results is that the delay in trials might be caused by FDA-imposed clinical holds. An FDA clinical hold is an order issued by FDA to the sponsor of the drug development to delay a proposed clinical investigation. Common reasons for clinical holds include the presence of impurity profile indicative of a potential health hazard, missing supporting information on the anticipated dose or exposure, or insufficient information on the clinical protocol (Lapteva and Pariser, 2016). After the hold, drug developers would work with the FDA to improve the clinical protocol and address all deficiencies until the hold is lifted. A trial that experienced a hold may have a prolonged trial duration, but the FDA may potentially help improve the trial quality throughout the hold period. This could also explain a relation between trial completion delays and continuation decision, but would be due to external, and potentially project-improving factors. We note that this concern applies mostly to the results in Table 2, rather than the IV-based results in Table 3). To allay this concern further, Online Appendix Table OA.2 shows that there is no statistically significant relationship between the presence of clinical holds and trial duration, trial delays, or trial site congestion measures (our instrument for delays). We conclude that it is

unlikely that clinical holds are a driver of any of our results.

4.5 Underlying Channels

The empirical findings above show that pharmaceutical firms' R&D decisions exhibit excess commitment with respect to R&D projects, with unanticipated, and plausibly-exogenous delays in trial completions strongly affecting trial continuation decisions. In this section, we examine various potential underlying channels that might lead to path dependence and (excess) commitment in response to drug delays. In Table 5 we examine agency conflicts inside the firm, initial firm expectations about trial length, financial constraints, and heterogeneities by trial phase. In Table 6 we discuss real options as a potential underlying mechanism. In addition to these firm- and project characteristics-related channels, in Table 7 we examine mechanisms pertaining to the ultimate decision-maker responsible for the R&D investment, the drug developing firm's CEO.

4.5.1 Firm- and Project Characteristics-Related Channels

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

Doing so, we conclude that a mechanism related to an aversion to "start over or liquidate" does not explain our results. For one, the many-projects-in-the-pipeline restriction only drops relatively few observations from the sample (less than 25%). That is, the findings in the previous sections predominantly come from large firms with many projects to begin with, leaving little room for an early-stage-firm channel as in [Guedj and Scharfstein \(2004\)](#). Additionally, and consistent with the previous point, Column (1) of Table 5 shows that unexpected delays continue to have distortive effects on trial continuation decisions that are of very similar magnitudes among firms with more than ten drug projects in the pipeline. Thus, the effects we find are not driven by early-stage firms with no or few alternative drugs as in [Guedj and Scharfstein \(2004\)](#).

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

However, the empirical correlation between anticipated duration and delay, while positive, is relatively small ($\rho = 0.07$). Furthermore, when we directly control for anticipated duration in Column (2) of Table 5, we obtain very similar results.⁸ In light of this evidence, differential sorting by firms into trials with different lengths or delays based on initial expectations is unlikely to be an underlying mechanism of our findings.

Financial Constraints. Next, we test whether the path dependence and commitment in drug-related decision-making is driven by financial constraints of the drug-developing firms. Intuitively, if firms are financially constrained, they may be more likely to stick to a chosen course of action. To investigate this, we augment our specifications with a series of financial constraints measures.⁹ Specifically, the *Constrained* variable in the third column of Table 5 is an indicator variable that equal one if the firm’s [Whited and Wu \(2006\)](#) (WW) index at trial end is in the top quartile of the index’s sample distribution. Column (3) shows that the distortive effect of delay on trial suspension remains unchanged with the added constraints measure. Thus, the path dependence in decision-making is not merely the result of constrained firms not being able to pivot (also in combination with the fact that most firms in our sample have many other viable drug candidate projects).¹⁰

Phase I Versus II Trials. The final two columns in Table 5 separate the effect by whether a trial is a Phase I or Phase II trial. As discussed in Section 2.1, the primary endpoint of Phase I trials is safety, whereas that of Phase II trials is drug efficacy, and Phase II trials tend to be longer and involve more participants. The effect of trial delay on subsequent continuation versus suspension is robust and

⁸In unreported tests, we include second- or third-order polynomials for anticipated duration and find nearly identical results compared to those reported in Table 5.

⁹We note that adding the financial constraints measures restricts the sample to include public firms only.

¹⁰We show robustness to other measures of financial constraints, in particular the [Hadlock and Pierce \(2010\)](#) (HP) and [Kaplan and Zingales \(1997\)](#) (KZ) indices, as well as to constraint measures defined as of the time of trial start in Appendix Table OA.1. Irrespective of the measure we use, we continue to find a strong effect of trial delay on subsequent suspension probabilities.

large for both Phase I and II trials. Thus, our findings cannot be not explained by any trial-specific channels or confounds, and instead unveil a robust and wide-ranging phenomenon in the realm of new drug development.

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

To test for a potential real options channel, we examine if delays in trial completion predict higher expected drug sales conditional on drug launch or fewer adverse reactions among subsequent trial participants. We adopt a specification following Equation (1). Panel A of Table 6 examines the outcome variables pertaining to expected drug sales. We use four different measures, including the expected first-year sales, the sum of first three-year sales, the sum of first five-year sales, and the average annual sales after the drug is approved by the FDA. All sales are million U.S. dollars in 2017 adjusted by GDP deflators and logged. Contrary to the real option hypothesis, we find that delays in trial completion do not predict the expected drug sales, neither statistically nor economically. Panel B examines the outcome variables pertaining to adverse reactions in clinical trials. Columns (1) and (2) use the total and maximum percentage of raw adverse reaction counts (an enrolled participant might have multiple reports) relative to the total number of enrolled patients across various adverse symptoms. Instead of a negative correlation, we find a significantly positive relation between delays in trial completion and the number of adverse reaction reports. We note that the two measures include many light adverse reactions, such as headache, nausea, and fatigue, which are of less importance for drug evaluation by the FDA. In columns (3) and

(4), we only include life-threatening adverse reactions, such as acute myelogenous leukemia, basal cell carcinoma, and brain hemorrhage. Again, the results are inconsistent with the real option hypothesis as the coefficients on delay in trial completion are statistically and economically insignificant. These results suggest that a real option channel is unlikely to explain the path-dependence in pharmaceutical R&D in our setting.

4.5.2 CEO-Related Channels

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

CEO Pay Sensitivity to Firm Value Changes. We first study how CEO pay-performance sensitivity (δ) mediates the relationship between trial delays and continuation versus suspension decisions.¹¹ Conditional on having started a drug project, CEOs with high δ may be particularly reluctant to suspend a delayed drug project, if they (correctly or falsely) believe that this would be interpreted as a negative signal (e.g., about the CEO's ability) by shareholders and depress the stock price. Consistent with this conjecture, Column (1) of Table 7 finds that the effect of trial delay on project continuation is substantially more pronounced among CEOs with greater exposure to stock price changes.

CEO Changes. Second, we study how firms' project continuation decisions after trial delays vary with CEOs' personal responsibility for having initiated the project. In the context of mergers and acquisitions, Guenzel (2023) shows that managers engage in sunk-cost thinking with respect to acquired targets, distorting their subsequent decisions to abandon acquired businesses through divestiture. Guenzel (2023) finds that the sunk-cost distortions are driven by the CEO who made the initial acquisition. To test for a similar CEO effect in our setting, we create a *Same CEO* indicator that equals one if the firm's CEO is the same at trial start and completion, and then interact the CEO variable with our delay measure. We emphasize the test of interest is *not* whether an incumbent is more or less likely to suspend a trial, but

¹¹We obtain the CEO δ 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 δ 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.

how the effect of delay varies depending on the CEO regime.

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

Overall, the results in Table 7 strongly support the existence of economically and statistically significant management-induced frictions in new drug development within firms.

5 Welfare Implications

5.1 Outcome Variables Related to Patient Welfare

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

5.2 Econometric Framework

Unlike the OLS and 2SLS approaches adopted in Section 3, one complication in our welfare analysis is endogenous sample selection. Specifically, while we can study the effect of investment distortions

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

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

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

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

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

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

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

In our setting, the endogenous sample selection is embodied in the possible correlation between c_i and u_i while the endogenous explanatory variable (or endogenous treatment) is captured by the fact that c_i

might be correlated with v_i . Following a similar strategy as in Section 3.2, we use an exogenous variable to instrument for c_i . Specifically, assume c_i is generated according to the model

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

where z_i is the instrumental variable.

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

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

where (u_i, ξ_i, v_i) is jointly normally distributed and independent of all observables z_i and \mathbf{X}_i . Without loss of generality, we normalize $var(u_i) = 1$ as Y_i^{app} is a binary variable. Our setting and assumptions are similar to the sample selection model with endogenous explanatory variables in Chapter 19 of Wooldridge (2010). The key difference is that we allow the endogenous variable to enter the sample selection Equation (4). To discipline the model, we hence impose extra structure on the residuals through Equation (7). Note that Equation (7) is a stronger version of the restriction imposed on the residuals. Our estimation can also be obtained under a weaker set of assumptions where (u_i, ξ_i) is jointly normally distributed and $E[v_i | \beta_1 \xi_i + u_i] = \alpha_3(\beta_1 \xi_i + u_i)$.

We detail our estimation procedure as follows:

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

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

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

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

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

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

We first assess welfare implications as they pertain to the development and launch of orphan drugs—drugs for infrequent diseases for which there are commonly no other or only few other treatment alternatives. As alluded to above, for this part we can in fact continue to use the IV strategy from Section [3.2](#), and do not have to rely on the more advanced framework from Section [5.2](#). Specifically, *conditioning* on clinical trials involving orphan drugs, we examine how the likelihood of eventual orphan drug approval by the FDA (i.e., launch of the drug in the market) varies with congestion-induced delay in the completion of Phase I or Phase II clinical trials.

We present the corresponding results in Table [8](#). Similar to Table [3](#), Column (1) presents the first-stage results. Despite the reduction in sample size (due to the restriction to orphan drug clinical trials), the first stage remains strongly significant. Column (2) then presents the second-stage results, with the indicator variable for FDA orphan drug approval as the dependent variable of interest. We find a significant positive effect of instrumented delay on the probability that orphan drugs are ultimately launched in the marketplace. For example, a six-month congestion-induced delay in trial completion is associated with a 1.5 percentage point increase in the probability of orphan drug market launch. This is an economically large effect, given the relatively low baseline probability of eventual FDA approval of orphan-drug-related clinical trials of just above 10% in the sample.

5.4 Effects on Adverse Events Related to Approved Drugs

We next assess welfare implications as they pertain to adverse effects in patients of approved drugs. For this part, we make use of the econometric sample-selection framework detailed in Section [5.2](#). Table [9](#) shows the results, with standard errors bootstrapped using 100 iterations. Column (1) shows the first stage results, i.e., the first part of the second step outlined in Section [5.2](#). Similar to Table [8](#), we continue to observe a strong relation between trial site congestion and delay despite the reduction in sample size.

(In Table 9, the sample size reduction is due to the fact that we focus on outcomes of approved drugs.)

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

Taken together, the results in Sections 5.3 and 5.4 clarify the intricacies and complexities of how distortions in firm decision-making with respect to innovation can affect (consumer) welfare. While the distortions we uncover increase the breadth of treatments for infrequent diseases for which there are often no other existing medications, they might also lead to the launch of drugs with more adverse effects. More broadly, the results in this paper highlight that for other parties, distortions in firm investment behavior do not need to exclusively entail welfare *costs*; instead, investment distortions can, and in fact may frequently, induce positive externalities on third parties, e.g., in the form of increased product variety.

6 Conclusion

In this paper, we study the extent to which existing commitments made to R&D endeavors induce path dependence in subsequent R&D decision-making, and examine resulting welfare implications. We study these questions using project-level R&D data from the pharmaceutical industry, namely clinical trials data with granular information on many important dimensions including project timelines and project outcomes. Clinical trials that firms complete but that experienced unanticipated increases in commitment, due to being unexpectedly delayed or unexpectedly costlier as a result of exchange rate fluctuations, are significantly more likely to be advanced by firms through the next trial phase. The economic magnitudes we uncover suggest substantial path-dependence in firm decision-making. The implications for consumers as a result of distorted R&D undertakings by firms are nuanced. On the one hand, delay-induced elevated commitment to clinical trials 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) may be associated with modestly higher adverse event counts.

Our findings suggest a variety of potential directions for future research. One promising area for future work would be to investigate heterogeneities along several dimensions. For example, it would be interesting to explore how the magnitude of R&D distortions varies with organizational structure (e.g., top-down versus decentralized decision-making) and by target disease and drug type (e.g., depressants versus stimulants), and whether there are heterogeneous effects on consumer welfare outcomes (e.g., by severity of adverse effects, or by various demographic characteristics). Alternatively, it would also be interesting to bring the analysis of “firm distortions and consumer welfare effects” to other contexts. The notion that distorted firm decision-making can come with both positive and negative consequences for consumers applies broadly, whenever firms and consumers interact in goods or financial markets.

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*.
- 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.
- 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.

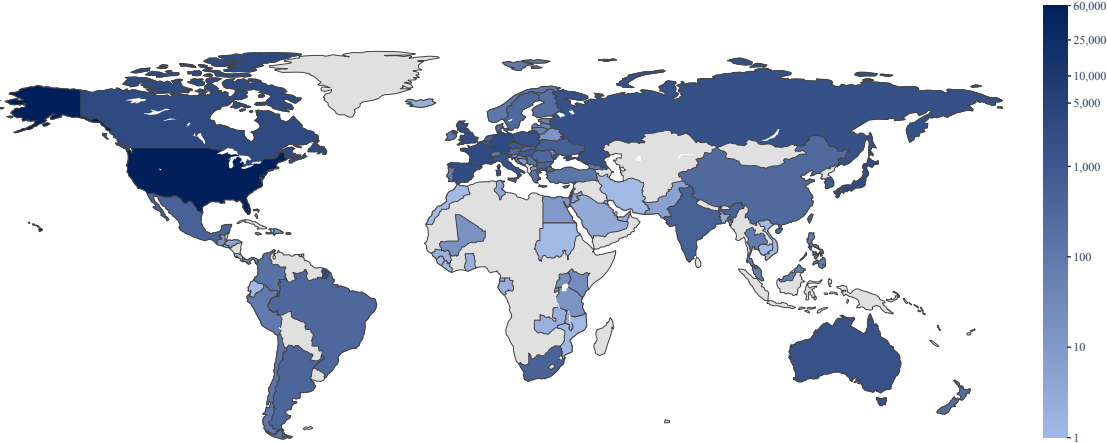
- DIXIT, R. K., A. K. DIXIT, AND R. S. PINDYCK (1994): *Investment under uncertainty*, Princeton university press.
- FAMA, E. F. (1980): “Agency problems and the theory of the firm,” *Journal of Political Economy*, 88, 288–307.
- FESTINGER, L. (1962): *A theory of cognitive dissonance*, vol. 2, Stanford University Press.
- GIBBONS, R. AND K. J. MURPHY (1992): “Optimal incentive contracts in the presence of career concerns: Theory and evidence,” *Journal of Political Economy*, 100, 468–505.
- GUEDJ, I. AND D. S. SCHARFSTEIN (2004): “Organizational scope and investment: Evidence from the drug development strategies and performance of biopharmaceutical firms,” .
- GUENZEL, M. (2023): “In Too Deep: The Effect of Sunk Costs on Corporate Investment,” *The 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.
- 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: Svenska Handelshögskolan.
- 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 L. ZINGALES (1997): “Do investment-cash flow sensitivities provide useful measures of financing constraints?” *Quarterly Journal of Economics*, 112, 169–215.
- KLEIBERGEN, F. AND R. PAAP (2006): “Generalized reduced rank tests using the singular value decomposition,” *Journal of Econometrics*, 133, 97–126.
- KRIEGER, J., D. LI, AND D. PAPANIKOLAOU (2022a): “Missing novelty in drug development,” *Review of Financial Studies*, 35, 636–679.
- KRIEGER, J. L., X. LI, AND R. T. THAKOR (2022b): “Find and replace: R&D investment following the erosion of existing products,” *Management Science*.

- LAPTEVA, L. AND A. R. PARISER (2016): “Investigational New Drug applications: a 1-year pilot study on rates and reasons for clinical hold,” *Journal of Investigative Medicine*, 64, 376–382.
- LEE, D. S. (2009): “Training, wages, and sample selection: Estimating sharp bounds on treatment effects,” *Review of Economic Studies*, 76, 1071–1102.
- LERNER, J. AND U. MALMENDIER (2010): “Contractibility and the design of research agreements,” *American Economic Review*, 100, 214–46.
- LERNER, J., H. SHANE, AND A. TSAI (2003): “Do equity financing cycles matter? Evidence from biotechnology alliances,” *Journal of Financial Economics*, 67, 411–446.
- LI, X., T. LIU, AND L. A. TAYLOR (2023): “Common ownership and innovation efficiency,” *Journal of Financial Economics*, 147, 475–497.
- LIU, T. (2021): “Financing Experimentation,” Working paper, MIT Sloan School of Management.
- LO, A. W. AND R. T. THAKOR (2022): “Financing biomedical innovation,” *Annual Review of Financial Economics*, 14, 231–270.
- (2023): “Financial intermediation and the funding of biomedical innovation: a review,” *Journal of Financial Intermediation*, 101028.
- MA, Y., T. ROPELE, D. SRAER, AND D. THESMAR (2020): “A quantitative analysis of distortions in managerial forecasts,” Working paper, National Bureau of Economic Research.
- MALMENDIER, U. (2018): “Behavioral Corporate Finance,” in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 1, 277–379.
- 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.
- ROBINSON, D. T. AND T. E. STUART (2007): “Financial contracting in biotech strategic alliances,” *Journal of Law and Economics*, 50, 559–596.
- STAW, B. M. (1976): “Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action,” *Organizational behavior and human performance*, 16, 27–44.
- STAW, B. M., S. G. BARSADE, AND K. W. KOPUT (1997): “Escalation at the credit window: A longitudinal study of bank executives’ recognition and write-off of problem loans,” *Journal of Applied Psychology*, 82, 130.

- STAW, B. M. AND H. HOANG (1995): "Sunk costs in the NBA: Why draft order affects playing time and survival in professional basketball," *Administrative Science Quarterly*, 474–494.
- STOCK, J. H., J. H. WRIGHT, AND M. YOGO (2002): "A survey of weak instruments and weak identification in generalized method of moments," *Journal of Business & Economic Statistics*, 20, 518–529.
- THAKOR, R. T. AND A. W. LO (2022): "Competition and R&D financing: evidence from the biopharmaceutical industry," *Journal of Financial and Quantitative Analysis*, 57, 1885–1928.
- VYTLACIL, E. (2002): "Independence, monotonicity, and latent index models: An equivalence result," *Econometrica*, 70, 331–341.
- WHITED, T. M. AND G. WU (2006): "Financial constraints risk," *Review of Financial Studies*, 19, 531–559.
- WONG, C. H., K. W. SIAH, AND A. W. LO (2019): "Estimation of clinical trial success rates and related parameters," *Biostatistics*, 20, 273–286.
- WONG, H.-H., A. JESSUP, A. SERTKAYA, A. BIRKENBACH, A. BERLIND, AND J. EYRAUD (2014): "Examination of clinical trial costs and barriers for drug development final," *Reports Submitted to Office of the Assistant Secretary for Planning and Evaluation*.
- WOOLDRIDGE, J. M. (2010): *Econometric analysis of cross section and panel data*, MIT press.
- WOUTERS, O. J., M. MCKEE, AND J. LUYTEN (2020): "Estimated research and development investment needed to bring a new medicine to market, 2009-2018," *Journal of the American Medical Association*, 323, 844–853.

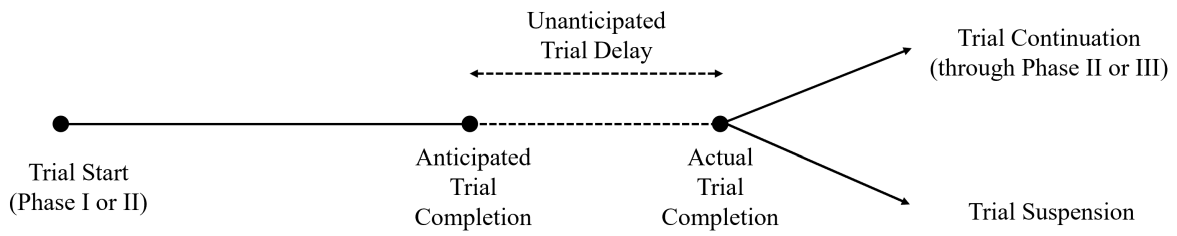
Figures

Figure 1: Geographic Distribution of Trial Sites



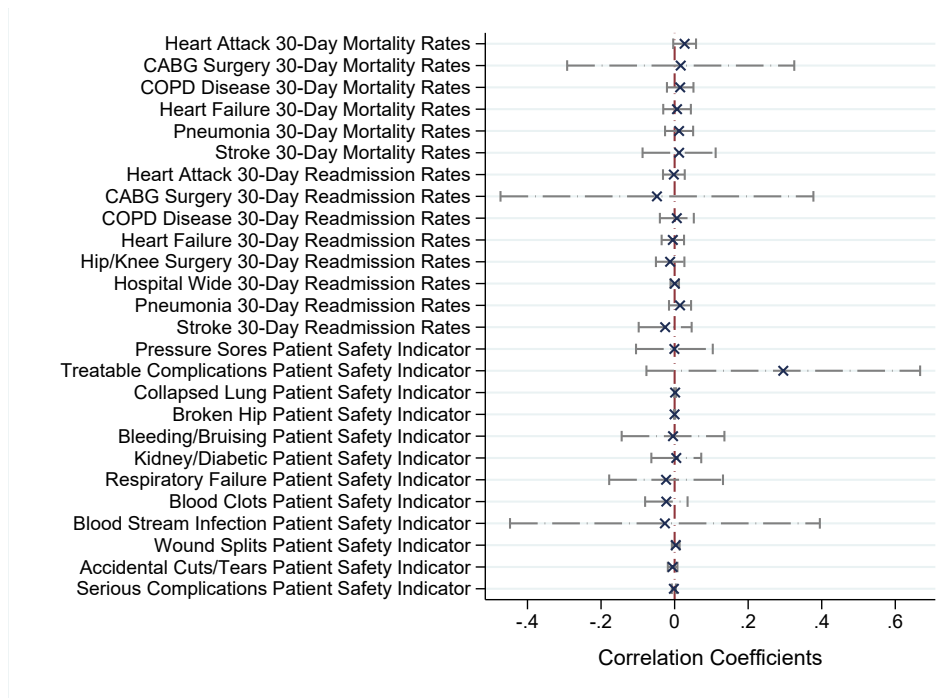
This figure exhibits the trial-site geographic distribution of clinical trials included in the analysis sample. For each country, we calculate the natural logarithm of the number of trial sites.

Figure 2: Project Timeline



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

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

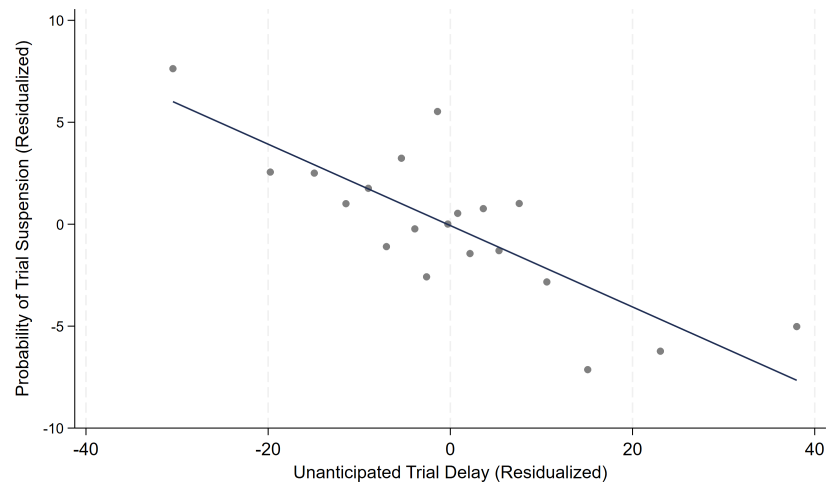


This figure exhibits the correlation coefficients between the trial site congestion measure and the average hospital care quality measures in a zip code. The correlation coefficients are estimated by the regression

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

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

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



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

Tables

Table 1: Summary Statistics

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

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

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

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

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

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

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

	DV: Delay		DV: Suspension	
	First Stage		Second Stage	
	(1)	(2)	(3)	(4)
<i>Trial Site Congestion</i>	0.360*** (4.26)	0.358*** (4.37)		
<i>Delay in Trial Completion</i>			-0.337** (-2.33)	-0.530*** (-2.97)
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	18.13	19.13	–	–
Observations	9,397	9,161	9,397	9,161

Table 4: Exchange Rate Variation and Project Continuation Versus Suspension

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

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

Table 5: Firm- and Project Characteristics-Related Channels

This table investigates underlying firm- and project characteristics-related channels. The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records. The dependent variable, $Suspension_i$, is an indicator variable (multiplied by 100 for ease of exposition) for whether project i was suspended after completing Phase I or II clinical trials. The independent variable, $Delay in Trial Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). See Section 4.5 for further details. Standard errors are two-way clustered at the ICD category and quarter levels. We report t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

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

Table 6: Real Options Channel

This table investigates real options as an underlying channel. Panel A examines the relationship between the log of expected drug sales (Million U.S. dollars adjusted to 2017 by GDP 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. Columns (1) to (4) use four different sales measures (all take logs), including the expected first-year sales after launch, the sum of first three-year sales, the sum of first five-year sales, and the average annual sales. Panel B examines the relationship between the percentage of adverse reactions among trial participants for project i that already completed Phase I or II clinical trials and $Delay in Trial Completion_i$. Columns (1) to (4) use four different dependent variables, including the total percentage of adverse reaction reported in the completed trial (a participant might experience multiple adverse events such as headache, nausea, fatigue) relative to the total number of enrolled participants, the maximum percentage of adverse reaction among various adverse symptoms, the total percentage of severe adverse reaction (life-threatening reaction, for example, acute myelogenous leukemia, basal cell carcinoma, brain hemorrhage, etc.), and the maximum percentage of severe adverse reaction among various adverse symptoms. We adopt the specification following Equation (3.1). Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD category (taking logs). Standard errors are two-way clustered at the ICD category and quarter levels. We report t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Expected Drug Sales

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

Panel B: Adverse Reaction in Trials

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

Table 7: CEO-Related Channels

This table investigates underlying CEO-related channels. The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records. The dependent variable, $Suspension_i$, is an indicator variable (multiplied by 100 for ease of exposition) for whether project i was suspended after completing Phase I or II clinical trials. The independent variable, $Delay\ in\ Trial\ Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. $CEO\ Delta$ is the natural logarithm of delta as in Coles et al. (2006), standardized to have a mean of zero and standard deviation of one for expositional purposes. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). See Section 4.5 for further details. Standard errors are two-way clustered at the ICD category and quarter levels. We report t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

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

Table 8: Effects on New Drug Launches With No or Few Existing Medications (Orphan Drugs)

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

	DV: Delay	DV: FDA Approval
	First Stage	Second Stage
	(1)	(2)
<i>Trial Site Congestion</i>	0.551*** (5.17)	
<i>Delay in Trial Completion</i>		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 9: Effects on Adverse Events Related to Approved Drugs

This table investigates the effects of delay-induced R&D investment distortions on adverse events related to approved drugs, using trial site congestion as an instrument for delay in clinical trial completion. The unit of observation is a drug project. The sample contains all approved projects that completed Phase I or II trials with detailed clinical trials records. The second-stage dependent variables are the count of adverse events reported by the FDA (taking logs) within the first and the first three years since drug launch, respectively. The instrument, $TrialSiteCongestion_i$, is the normalized change in the number of participants in the universe of trials on ClinicalTrials.gov between trial start and end date, taking the maximum across all trial sites of a given trial in the sample. The second-stage independent variable, $DelayinTrialCompletion_i$, 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 5.2 for further details on the econometric specification and estimation. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company’s pipeline (taking logs), and number of competing projects in the same ICD (taking logs). We report t–statistics, based on standard errors accounting for the generated regressors problem by bootstrapping, in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	DV: Delay First Stage	DV: log(Adverse Events in 1 Year) Second Stage	DV: log(Adverse Events in 3 Years) Second Stage
	(1)	(2)	(3)
<i>Trial Site Congestion</i>	0.903*** (5.25)		
<i>Delay in Trial Completion</i>		0.014 (0.57)	0.014 (0.58)
<i>Inverse Mills Ratio</i>	−1.572 (−0.12)	1.101 (0.54)	1.467 (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

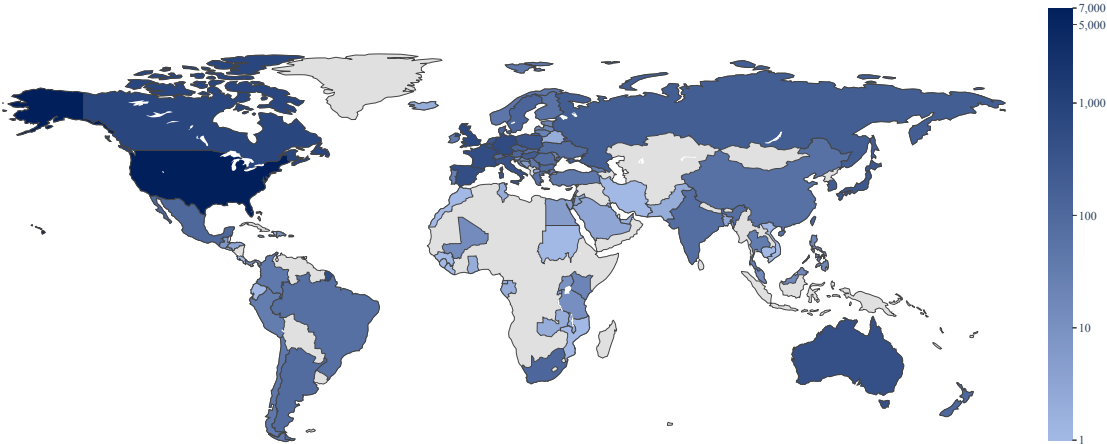
Marius Guenzel Tong Liu
Wharton MIT Sloan

Not For Publication

I. Supplementary Results

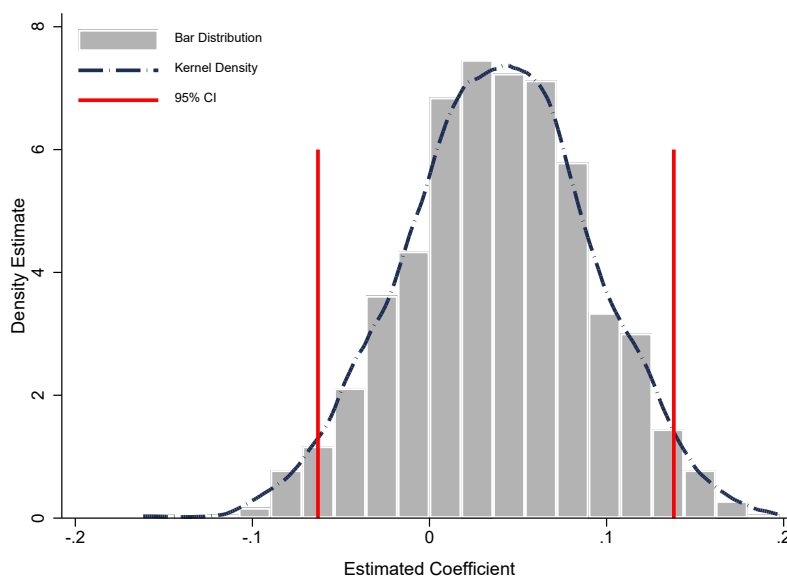
I.A Figures

Figure OA.1: Geographic Distribution of Clinical Trials

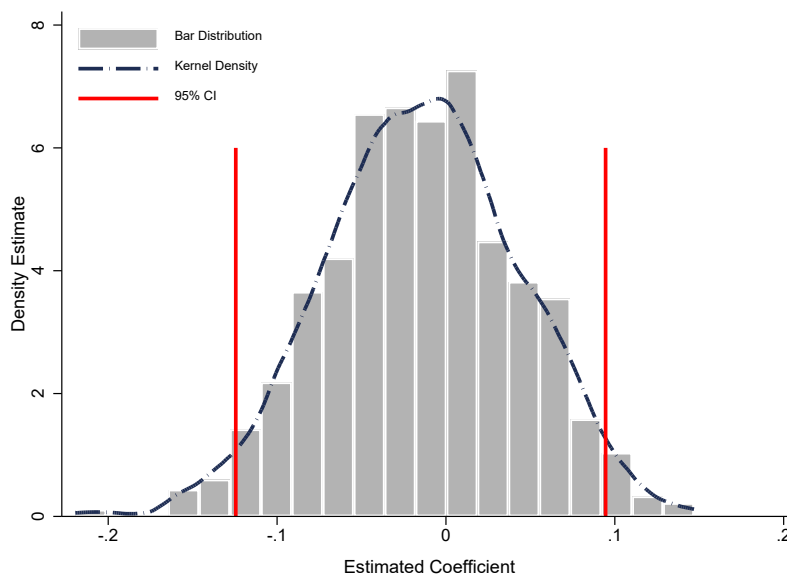


This figure exhibits the geographic distribution of clinical trials in our analysis sample. For each country, we calculate the natural logarithm of the number of trials.

Figure OA.2: Placebo Test for Variation in Exchange Rates



(a) Distribution of Estimated Coefficient of ΔFX without Controls



(b) Distribution of Estimated Coefficient of ΔFX with Controls

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

I.B Tables

Table OA.1: Delays in Trial Completion and Project Continuation Versus Suspension: Further Evidence on Financial Constraints

This table further explores the role of financial constraints for explaining the relationship between trial suspension and delay in completion of the preceding trial phase, adding to the evidence in the final column in Panel A of Table 5. The unit of observation is a drug project. The sample contains all projects that completed Phase I or II trials with detailed clinical trials records. The dependent variable, $Suspension_i$, is an indicator variable for whether project i was suspended after completing Phase I or II clinical trials. The independent variable, $Delay\ in\ Trial\ Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD (taking logs). See Section 4.5 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)
<i>Delay in Trial Completion</i>	-0.196*** (-4.08)	-0.179*** (-3.91)	-0.197*** (-4.10)	-0.197*** (-4.84)	-0.195*** (-4.69)	-0.193*** (-4.65)
<i>Constrained (WW)</i>	0.238 (0.06)					
<i>Constrained (HP)</i>		13.722*** (3.20)				
<i>Constrained (KZ)</i>			-0.766 (-0.27)			
<i>Constrained-Trial Start (WW)</i>				-0.905 (-0.25)		
<i>Constrained-Trial Start (HP)</i>					4.572 (1.00)	
<i>Constrained-Trial Start (KZ)</i>						3.429 (1.18)
Controls	Y	Y	Y	Y	Y	Y
Year \times 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.2: FDA Clinical Holds and Trial Duration

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

DV:	Trial Duration	Delay in Completion	Instrument for Delay
	(1)	(2)	(3)
<i>Had Clinical Hold</i>	−6.032 (−1.65)	1.122 (0.32)	5.465 (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

II. Data and Estimation Appendix

II.A Details on Information from `clinicaltrials.gov`

In a first step, we use the official and short titles of observations in our main dataset to find the corresponding National Clinical Trial (NCT) identifier. For this, we use a combination of searching the `clinicaltrials.gov` database for the official and short titles, and fuzzy-string matching between the titles in our dataset and those on `clinicaltrials.gov` (using the `jellyfish` package in Python). After matching, we add additional information from `clinicaltrials.gov` (number of participants, trial start and end dates), in order to compare the added information to that from our main dataset. This allows us to determine whether we matched the correct trial. Before determining the correctness of the match (described further below), we also scrape the anticipated completion date as well as primary anticipated completion date from `clinicaltrials.gov`'s "History of Changes" records associated with each trial (i.e., each NCT identifier). We use the information from the earliest available historical record, and use the anticipated completion date (rather than the primary anticipated completion date) when available.¹²

In a second step, we further assess the correctness of the matched trial from `clinicaltrials.gov`. Whenever it is not clear whether a match is correct (e.g., we conclude the match is correct when the information on number of participants as well as trial start and end dates in both datasets coincides, and the Levenshtein distance between titles is at most 10, or the `jellyfish` similarity score is above 0.99) or incorrect (e.g., we conclude the match is incorrect when the anticipated completion date from `clinicaltrials.gov` precedes the start date of the trial in our main dataset), we manually assess the match (e.g., we assess whether the sponsors and collaborators match, taking into account acquisitions, joint ventures, and name changes). Table OA.3 below contains several examples of correctly or incorrectly matched trials between our main dataset and `clinicaltrials.gov`. The first (third) example contains a clear correct (incorrect) match. The second example contains a slightly more subtle incorrect match. Finally, for incorrect matches that were based on searching the `clinicaltrials.gov` database for the official and short titles, we also implement the fuzzy-string matching, and repeat the subsequent steps for the new matches (adding information from `clinicaltrials.gov`, scraping NCT identifier, assessing correctness of each match).

¹²For one observation in the dataset, we use the primary anticipated end date, as the anticipated completion date is implausibly high in the first record (year 2087) and corrected in subsequent records.

Table OA.3: Examples

		Main Dataset				clinicaltrials.gov			
Company	Start Date	Completion Date	No. of Participants	Title	Sponsor/ Collaborator	Start Date	Completion Date	No. of Participants	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 Corporation	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 I, 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 Pharmaceuticals 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

II.B Matching with FDA Drug Adverse Events Database

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

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

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

II.C Estimation Procedure

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

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

where $g(z_i, \mathbf{X}_i, Y_i^{app}) = E[v_i | z_i, \mathbf{X}_i, Y_i^{app}]$ and $e_i = v_i - E[v_i | z_i, \mathbf{X}_i, Y_i^{app}]$. Since $E[e_i | z_i, \mathbf{X}_i, Y_i^{app}] = 0$ by construction, we could 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,

$$\begin{aligned} g(z_i, \mathbf{X}_i, Y_i^{app} = 1) &= E[v_i | z_i, \mathbf{X}_i; Y_i^{app} = 1] \\ &= E[v_i | z_i, \mathbf{X}_i; \beta_1 c_i + \beta_2' \mathbf{X}_i + u_i > 0] \\ &= E[v_i | z_i, \mathbf{X}_i; \beta_1 (\gamma_1 z_i + \gamma_2' \mathbf{X}_i + \xi_i) + \beta_2' \mathbf{X}_i + u_i > 0] \\ &= E[v_i | z_i, \mathbf{X}_i; \beta_1 \xi_i + u_i > -\beta_1 \gamma_1 z_i - (\beta_1 \gamma_2' + \beta_2') \mathbf{X}_i]. \end{aligned} \quad (8)$$

Given the assumption in Equation (7), we have

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

Therefore, the conditional expectation of v_i is

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

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

Applying iterated expectations on Equation (8),

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

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.