Excess Commitment in R&D

March 2023

Please do not distribute without permission.

Abstract

High levels of commitment to R&D activities can facilitate breakthrough innovations, but can also turn into excess commitment to previously chosen actions. Using project-level R&D data on clinical trials by pharmaceutical firms, we study how *unanticipated* variation in firms' commitment to trials affects subsequent firm decision-making and R&D outcomes. Unexpected trial completion delays, as well as unexpected trial cost increases due to exchange rate fluctuations, significantly increase the likelihood that firms advance trials to the next trial phase. Consumers may, in fact, in some ways *benefit* from firm-induced distortions in new drug development. Marginally-launched drugs because of commitment distortions are associated with insignificantly more adverse events, but are significantly more likely to target diseases for which there are no or only few existing medications in the marketplace (orphan drugs).

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

1 Introduction

Innovation is central to economic growth and, undoubtedly, research and development (R&D) is an essential ingredient in the process of innovation. Therefore, it is important to understand the factors that determine the intensity of R&D activity. One defining feature of innovation and R&D is that to reach breakthroughs, it is oftentimes necessary to demonstrate long-term commitment to a given course of action. 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.

Conceptually, excess investment is different from overinvestment in the economy. There is a large literature on, and active debate about, to what extent there is adequate innovation activity in the aggregate economy. For instance, seminal prior work has uncovered many reasons for why there might be aggregate underinvestment in innovation, including low spillover effects, low competition, patent protection, infringement lawsuits, and risk aversion (Hall and Lerner, 2010). Our focus is different from this prior work, in that we examine the level of commitment to *existing, in-process* R&D endeavours, rather than whether at the extensive margin there is sufficient adoption of R&D projects.

Empirically, little is known to date about to what extent R&D activity involves adequate versus excess commitment. Furthermore, it is even less clear to what extent any potential excess commitment to R&D is harmful, or beneficial, from an aggregate welfare perspective. For example, does elevated commitment lead to worse or even unsafe innovation, or does it spur the creation of otherwise nonexisting innovation that benefits consumers?

Studying to what extent commitment and path-dependent decision-making with respect to innovation is efficient or distorted, as well as resulting welfare consequences, is challenging for at least two reasons. First, it requires granular data on R&D activities at the project level, with observable information on R&D milestone achievements and final R&D outcomes. Second, it requires not solely reliable measures of the level of commitment to a certain R&D activity, but additionally a setting in which there is plausibly-exogenous variation in the intensity of commitment to R&D.

In this paper, we overcome these twin challenges related to data and identification in the important setting of new drug development by pharmaceutical firms. Clinical trial drug projects provide a near-ideal,

project-level setting of R&D investment with detailed observable and quantifiable outcomes. Additionally, in this setting, we are able to isolate variation in R&D commitment to drug projects. We study whether *unanticipated* trial delays or cost increases distort firms' decision to continue versus suspend clinical trials, and how this affects ultimate R&D outcomes, in particular patient outcomes of approved drugs. Consistent with the existence of firm-level frictions inducing path dependence and excess commitment, we find that unexpectedly delayed trials, and trials that become costlier due to exchange rate fluctuations, are less likely to be suspended by firms. Marginally-launched drugs as a result of commitment distortions come with slightly, but insignificantly, more adverse events in patients such as death, hospitalization, disability, and congenital anomaly after consuming the drug. At the same time, distorted R&D commitment leads to more approvals of orphan drugs for rare diseases which commonly have no or only few other treatment options.[1] One broad-level takeaway of these results is that firm distortions can potentially have *positive* externalities on consumers.

To establish these findings, 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, 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. First, we leverage the information on drug firms' self-reported anticipated trial end dates—to our knowledge, these data are not explored in the 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). We uncover a strong link between the unanticipated trial

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

delays and firms' post-trial-completion decision to suspend versus advance the trial to the next clinical trial phase. Consistent with path dependence in decision-making, and elevated commitment by firms in response to prior commitments made, more delayed trials are less likely to be subsequently suspended. A one standard deviation increase in delay reduces the suspension probability by 4 percentage points, or 15% relative to the baseline suspension probability of 28%.

To further strengthen a causal and investment-distorting interpretation of the effect of unanticipated delays on subsequent project decisions, we apply an instrument variable (IV) approach that exploits the nature of limited capacity for a trial site to accommodate clinical trials. When more trials are taking place simultaneously in a given location, it is more likely to lead to bottlenecks. Our IV for trial completion delay is a measure capturing *trial site congestion*, which is the normalized change in the average patient enrollment speed of a zipcode where the trial takes place between trial start and end date. The main identification assumption is that changes in trial site congestion affect firms' decisions to advance trials to the next phase only through the effect of trial completion delay. We argue this exclusion restriction of our IV plausibly holds since other potential factors, for example, trial site quality proxied by hospital care quality provided in a region, seem uncorrelated with the congestion measure.

We find that trial site congestion and delay are significantly positively related, confirming the existence of a strong first stage. With respect to 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. Moreover, the IV-based economic magnitude of the effect is in the same ballpark as the OLS estimates.

Second, we shift the focus from delay and time-based commitment measures to trial cost measures, 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.[2] 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 to the next phase. In the exchange rate context, we estimate a path-dependence effect amounting to a 6% reduction in trial

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

suspension probability in response to a one standard deviation increase in trial costs through exchange rates.

Third, we investigate potential underlying channels and moderators for the effects of trial delay and exchange rate changes on trial suspension decisions. Neither set of results is driven by early-stage firms with no or few alternative drugs as in Guedj and Scharfstein (2004). Most firms in our sample have many viable drug candidates and restricting the estimation to such firms yields nearly identical effect sizes. The path dependence in decision-making is also not merely the result of firms not being able to pivot due to financial constraints as gauged by a series of common financial constraint measures. Moreover, differential sorting by firms into trials with different lengths or delays based on initial expectations is unlikely to be an underlying mechanism of our delay findings, since we can directly control for trial length expectations in the analysis without this having any effect on the results. Instead, we find evidence that both unexpected trial delay and the exchange rate effects on subsequent trial commitment are substantially larger when the CEO in charge does not change between trial start and end, consistent with senior-management-induced frictions in firms' project decision-making.

Finally, we assess welfare implications of our findings, and in particular implications for consumers. We are able to shed light on different welfare-relevant outcomes, including how distortions in new drug development affect the availability of drugs in the marketplace as well as adverse events in patients. With respect to adverse effects, we find that marginally-approved drugs due to delay-induced distorted firm behavior are 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. Taken together, these findings suggest a nuanced relation between firm distortions and welfare implications for consumers.

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.[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 challenges 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 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

[3] With respect to biomedical innovation, Lo and Thakor (2022) 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, Liu (2021) on the IPO decisions of biotech, Krieger et al. (2022b) on profit erosion of existing products, and Li et al. (2022) on common ownership among venture capital firms.

(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 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 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. Favorable trial results are required to move from one phase to the next. 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

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

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 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 certain trial phase if it is explicitly coded as suspended, discontinued, withdrawn, or coded as "no development" for over three

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

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., 2022).

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.

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). 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 US exchanges from Compustat

and Execucomp. We get the expected drug revenues once launching in the market from the Cortellis Competitive Intelligence. We collect the quality information of US hospitals from various CMS-managed quality programs, including 30-day mortality rates, 30-day readmission rates, and patient safety indicators.

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 I trials, the remainder are Phase II trials. About one in three trials are suspended, i.e., not advanced to Phase II or III, respectively. The average trial in our data is conducted at 13 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 3.2 to 3.4, 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 demonstrates the geographic distribution of clinical trial sites in our sample. As shown, clinical trials have gone global in the recent decades as the trial sites are widely spread across various countries. Though a large share of clinical trials are clustered in the US, accounting for 63% of the total trial sites (59,617 out of 94,261 clinical trial sites), other countries including Canada, Germany, France and Japan contribute another 12% of the total trial sites. The pattern is robust if we tabulate the number of trials across countries, as depicted in Figure OA.1 of the Online Appendix. In terms of the trial number, the US takes up 45% of the total trials (6,847 out of 15,222 clinical trials). The other four countries (Candada, Germany, France and Japan) contribute 16% of clinical trials in total.

3 Conceptual Framework and 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 and increases in the cost of foreign-based trials due to exchange rate fluctuations induce firms to be more

likely to "stay committed" to a clinical trial. 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. Before going into the empirical design in more detail, we first discuss how we conceptually think about "excess commitment" in our setting.

3.1 Conceptual Framework

To fix ideas, consider a firm i that has finished a given phase of a clinical trial and decides whether to continue or suspend the trial. As a benchmark scenario (which we refer to the "no excess commitment" scenario), we can model the firm's decision as a threshold criterion such that it continues if and only if

$$f(X) > T_{Bench}$$

where $X = (x_1, x_2, ..., x_n)$ is a vector of decision inputs observable to the firm (but potentially not to the econometrician), f aggregates the inputs, and T is the continuation threshold. We like to think of X as including all variables that a "standard," frictionless economic framework would deliver should matter for the decision, such as previous trial outcomes, prospects about future clinical trials, private information about the drug safety and efficacy, forecasted demand for the drug, degree of competition, firm or managerial degree of risk aversion, etc. Note that the above decision criterion is not necessarily optimal from a welfare perspective. For example, if firm is risk averse, it may produce less innovation than socially optimal. In other words, while we use the above case as the benchmark, no-excess-commitment case, we do *not* benchmark against the socially optimal decision.

We can conceptualize the idea that variables included in X matter for the continuation decision by saying that $\frac{\partial f(X)}{\partial x_i} \neq 0$ for all i (assuming f is differentiable). This also easily allows us to spell out what we mean by excess commitment. We refer to a firm as showing excess commitment with respect to clinical trial decisions if there exists an observable variable $w \in W, W \cap X = \emptyset$, such that the firm continues if and only if

$$f(X,w) > T \quad \text{with} \quad \frac{\partial f(X,w)}{\partial (w)} \neq 0.$$

That is, excess commitment is characterized by nonstandard observables affecting the decision to

suspend a trial.[6] This yields the testable prediction that under the null (alternative) hypothesis, variation in w across different trials should not (should) be related to firms' continuation decision.

There are, of course, many candidate observables falling into the set of W. We focus in particular on variables related to unexpected R&D costs, both monetary and time costs, given their importance in the R&D investment and innovation process, and given the prior evidence that initial costs can affect commitment levels (Staw, 1976; Guenzel, 2023). In Section 4.4, we will also draw links to firm and managerial incentives and frictions, including real options, financial constraints, and agency conflicts, to further explore potential underlying mechanisms.

3.2 Delay in Trial Completion

We first examine how firms factor unexpected delays in clinical trial completion into their R&D decision of the follow-on phase of trial, by estimating:

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

Suspension_i is an indicator variable for whether project i was suspended in the follow-on phase trial (Phases II or III clinical trials) 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.

[6] We do not take a stance on whether T in the excess commitment case is the same as or different from T_{Bench}.

3.3 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, 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 site/location hosts too many trials, it is challenging to recruit patients on time and monitor multiple trials simultaneously, possibly leading to a slowdown in the 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 ClinicialTrials.gov. We define each unique ZIP code in a country as one trial location, dropping any trials with missing addresses, number of participants, and start and completion dates. In a first step. we then construct a congestion measure at the ZIP code level as follows. For a trial i conducted in N_i sites with a start year τ_1 and a completion year τ_2, we compute the average patient enrollments per year per site E_i as trial i's total number of enrolled patients divided by the number of sites N_i and the number of years to complete the trial, $(\tau_2 - \tau_1)$. The the congestion measure of year t for location z, denoted by G_{zt}, is the sum of E_i for all trials that have a site located in z and year t lying between the trial start year τ_1 and the trial completion year τ_2:

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

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

congestion at location z. Our instrument for delay is then

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

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

The key identifying assumption is that 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 to 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 zipcode (aggregate conditions) is unlikely to be correlated with unobserved shocks to firms' decisions to advance a specific clinical trial (firm-level condition). One potential concern is that the more crowded trial site signals better service quality (e.g. better site monitoring, data collection, implementation of clinical protocol) it could potentially provide in trials. Clinical trials conducted in these crowded sites, though with a delayed schedule, tend to have trial results with higher quality. Therefore, firms are more likely to advance the trials to 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 zipcode. Figure 2 exhibits the correlation coefficients of 26 hospital quality measures devised by the CMS between 2005 to 2017.

3.4 Exchange Rate Variation in Foreign Trials

As a second measure of plausibly-exogenous prior commitments to a clinical trial that is independent of the time-based delay and congestion measures, we leverage the institutional feature that a large portion of clinical trials are conducted outside of the US. For a trial involving multiple sites internationally, the contracts and actual payments are conducted in other currencies rather than the US dollar. When foreign currencies become more expensive relative to the US 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 $\Delta FX^c_{[\tau]}$ as the percentage change in the exchange rate (unit of dollars in exchange for one unit of foreign currency in country c) τ years after the clinical trial was initiated. Then we construct an average exchange rate change for clinical trial i as

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

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

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

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

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

4 Results

4.1 Delay in Trial Completion

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 percentage points, or 15% relative to the baseline suspension probability of 28% (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).[7]

Overall, the results in Table 2 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 3.3, with and without the controls included in the last column of Table 2. The columns reveal a strong association between congestion and delay. Economically, a one standard deviation increase in the congestion instrument is associated with an increase in trial completion delay of close to four months. Statistically, the Kleibergen and Paap (2006) $F-$statistic is clearly above the common threshold for weak instruments of 10 (Stock et al., 2002).

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% continuing to use two-way clustered standard errors, now corrected for using a two-stage estimation approach.

These IV results further strengthen the interpretation that unanticipated, plausibly-exogenous delays in

[7] The coefficients on controls (unreported) are mostly 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.

trial completion induce distortions in firms' subsequent behavior with respect to clinical trial continuation versus suspension.

4.3 Exchange Rate Variation in Foreign Trials

Table 4 shows the results for how firms' decision to continue versus suspend a clinical trial after a given completed phase depends on exchange rate fluctuations experienced in foreign-based trials between trial start and completion. 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. Column (1) corresponds to Column (4) of Table 2, i.e., the specification with the full set of fixed effects but without additional controls. Column (2) then adds the controls, as in Column (5) of Table 2. In both columns, the coefficient on exchange rate changes is negative, implying exchange-rate-driven cost increases make trial continuation more likely, and highly statistically significant. In terms of magnitudes, a one standard deviation increase in trial costs through exchange rates is estimated to reduce the likelihood of subsequent trial suspension by 1.75 percentage points, or 6% relative to the baseline. These estimated magnitudes are slightly lower compared to the effects of unexpected delays on suspension decisions, but still economically meaningful.

Columns (3) and (4) add an "interaction" between the exchange rate variable and whether the trial is foreign-based and control for the main effect of foreign-based trial on trial suspension probability.[8] Doing so has little effect on the exchange rate variable of interest. If anything, the effect becomes slightly stronger when controlling for the main effect of foreign-based trial. Now, a one standard deviation exchange rate increase is associated with a reduction in suspension likelihood of 7% relative to the baseline.

The results in Table 4 provide additional evidence in favor of the idea that plausibly-exogenous increases in prior commitment to a chosen action induce distortions and path dependence firms' R&D decision-making.

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

4.4 Underlying Channels

The empirical findings thus far show that pharmaceutical firms' R&D decisions exhibit excess commitment with respect to R&D projects. In this section, we discuss potential underlying channels that might be related to path dependence and (excess) commitment in firms' drug project decision-making over time. We discuss several hypotheses in detail below, and gather additional empirical evidence to help distinguish between the various potential channels.

4.4.1 Potential Underlying Channels

Table 5 focuses on four potential underlying channels and moderators for the previous findings in Tables 2 to 4: agency conflicts, management changes, initial expectations about trial length, and financial constraints. Panel A presents mechanism results for the effect of unexpected trial delays on subsequent R&D commitment. Panel B presents mechanism results for the effect of exchange rate fluctuations.

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 (labeled *Many Projects* in the table).

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% in either Panel A or Panel B). That is, the findings in Tables 2 to 4 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) in both panels shows that unexpected delays and exchange rate changes 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).

Management Changes. If firms distort drug project decisions by taking past actions taken into account, it is natural to ask who the responsible decision-makers are. In a different context (mergers and

acquisitions), Guenzel (2023) shows that firms engage in sunk-cost thinking with respect to acquired targets, distorting their subsequent decisions to abandon acquired businesses through divestiture. Guenzel (2023) finds 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 and exchange rate variables. We emphasize the test of interest is *not* whether an incumbent is more or less likely to suspend a trial, but how the effects of delay and exchange rate fluctuations vary depending on the CEO regime.

Indeed, we find that the commitment effects we uncover are more pronounced when the CEO at trial end, when the decision to advance or suspend the trial is to be made, is the same CEO that was at the helm at trial start. Both the effect of unexpected trial delay on subsequent commitment to the trial and that of exchange rate variation are more than three times as large when there is no CEO change between trial start and end, in Columns (2) of Table 5. (We note that while economically large, the incumbent-CEO interaction effect with respect to exchange-rate-driven trial commitment is not statistically significant.) Overall, these results are consistent with Guenzel (2023) and the existence of senior-management-induced frictions in firms' project decision-making.

Expectations of Trial Duration. One possible channel most relevant to our delay results 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 Columns (3) of Table 5, we obtain very similar results, most importantly with respect to the effect of delay, but also with respect to the effect of exchange rates.[9] 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.

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

Financial Constraints. Finally, 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.[10] Specifically, the *Constrained* variable in the final columns 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 (4) in both panels shows that the distortive effect of both delay and exchange rates 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).[11]

Real Option We also explore the potential role of real option considerations in Table 6. The R&D process involves a sequence of actions and hence involves valuable 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). In a survey paper, McAfee et al. (2010) propose a stylized real option model such that more past investments might reveal better prospect about the option value of further investment, which might help explain why unexpected delay in a trial is positively correlated with the likelihood of the firm to advance it to the next phase trial. If indeed the trial delay reveals 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 less adverse reaction among trial participants so that it has a higher chance to survive all stipulated phases of clinical trials (a higher probability of obtaining the FDA approval).

To test this channel, we examine if delays in trial completion predict higher expected drug sales conditional on launching in the market or less adverse reaction among trial participants. We adopt a specification following Equation (1) and Table 6 reports the regression results. Panel A examines the outcome variables pertaining to expected drug sales. We use four different measures (all take logs),

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

[11] We show robustness to other measures of financial constraints, in particular the Hadlock and Pierce (2010) (HP) and Kaplan and Zingales (1997) (KZ) indices, as well as to constraint measures defined as of the time of trial start in Appendix Tables OA.2 and OA.3. Irrespective of the measure we use, we continue to find a strong effect of trial delay and exchange rates on subsequent suspension probabilities.

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 approved by the FDA. All sales are million US dollars in 2017 adjusted by GDP deflators. Contrary to the real option hypothesis, we find delays in trial completion fail to predict the expected sales of a drug, both in statistical significance and economic significance. Panel B looks into the outcome variables pertaining to adverse reaction reports in a completed trial. Columns (1) and (2) use the total and maximum percentages 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 document a significantly positive relation between delays in trial completion and the number of adverse reaction reports. Notice that much of light adverse reaction, e.g., headache, nausea, fatigue, are included in the counts, and these are of less importance for drug evaluation by the FDA. Therefore, in columns (3) and (4), we redo our analysis by only counting life-threatening adverse reaction in a trial, such as acute myelogenous leukemia, basal cell carcinoma, brain hemorrhage, etc. Again, the results are inconsistent with the real option hypothesis and the slopes on $\text{DelayInTrialCompletion}_i$ are statistically and economically insignificant. All these results suggest that the real option value might not be able to explain the excess commitment in pharmaceutical R&D.

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 two 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

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

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

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

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

where v_i is the residual term. The sample selection is reflected in the fact that we can observe $\tilde{Y}_i^{out}$ only if

drug project i is approved by the FDA. Therefore, we define the observable outcome $Y_i^{out} = \tilde{Y}_i^{out} \times \tilde{Y}_i^{app}$.

In our setting, the endogenous sample selection is embodied in the possible correlation between c_i and u_i while the endogenous explanatory variable (or endogenous treatment) is captured by the fact that c_i might be correlated with v_i. Following a similar strategy as in Section 3.3, we use an exogenous variable to instrument for c_i. Specifically, assume c_i is generated according to the model

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

where z_i is the instrumental variable.

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

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

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

We detail our estimation procedure as follows:

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

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

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

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

Our estimation procedure is similar to that proposed in Wooldridge (2010). We provide a detailed derivation in the Online Appendix (OA Section 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 3.3, 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 7. 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 8 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 7, we continue to observe a strong relation between trial site congestion and delay despite the reduction in sample size. (In Table 8, 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 ranging from project timelines to

project outcomes. Clinical trials that experience unanticipated increases in commitment, due do being unexpectedly delayed or unexpectedly costlier as a result of exchange rate fluctuations, are significantly more likely to advance to 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.

The findings of this paper 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.

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.

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.

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.

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

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*.

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 (2022): "Common ownership and innovation efficiency," Working paper, University of Pennsylvania.

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.

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.

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.

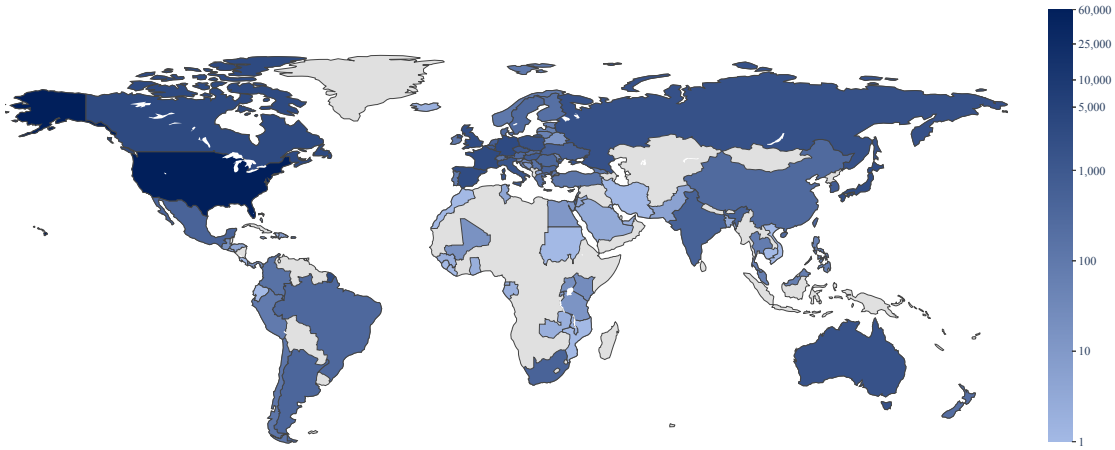
WONG, H.-H., A. JESSUP, A. SERTKAYA, A. BIRKENBACH, A. BERLIND, AND J. EYRAUD (2014): "Examination of clinical trial costs and barriers for drug development final," *Reports Submitted to Office of the Assistant Secretary for Planning and Evaluation*.

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.

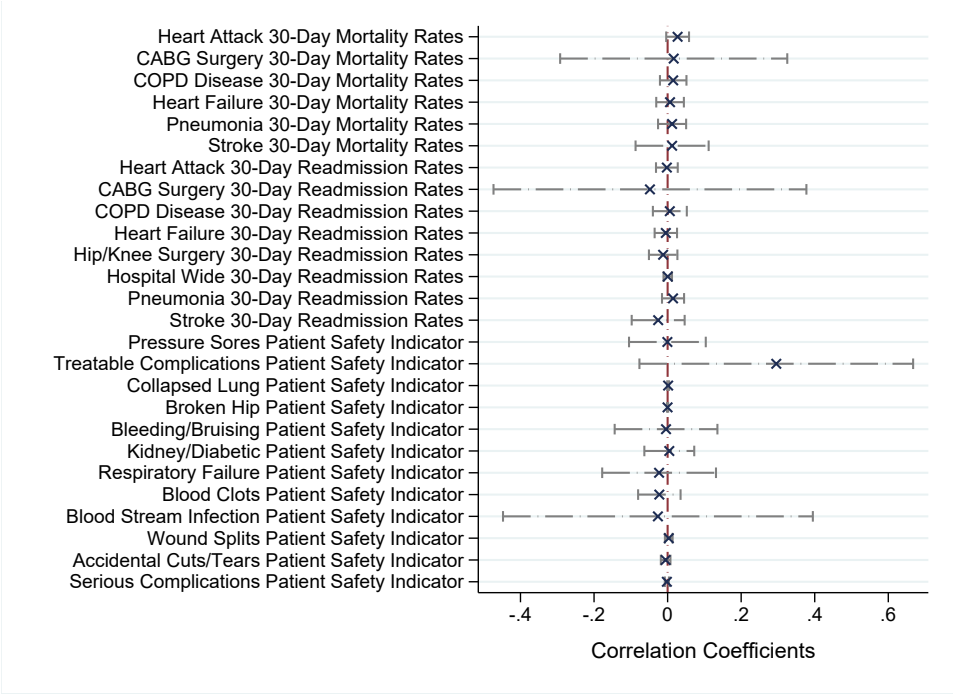
Figures

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: 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 zipcode. 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 zipcode z in year t. G_{zt} is the congestion measure at zipcode z in year t. γ_z and τ_t are zipcode 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 zipcode and year levels. Capped spikes represent 95% confidence intervals.

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
Clinical Trial Phase (I or II)	11,228	1.59	0.49	1.00	2.00	2.00
Suspension	11,228	0.28	0.45	0.00	0.00	1.00
Number of Clinical Trial Sites	11,179	13.65	40.62	1.00	3.00	13.00
Trial Participants (in hundreds)	11,195	1.00	1.62	0.25	0.49	1.06
Number of Drug Projects	10,998	136.14	179.07	11.00	41.00	230.00
Number of Competing Drug Projects (in logs)	11,182	5.16	1.34	4.39	5.33	6.01
Trial Duration (in months)	11,228	35.65	29.45	12.00	29.00	51.00
Anticipated Trial Duration	9,948	22.83	17.79	9.00	20.00	32.00
Delay in Trial Completion	9,948	11.78	19.19	0.00	3.00	19.00
Trial Site Congestion	9,948	5.28	10.04	0.00	0.93	4.19
Foreign Trial	11,228	0.35	0.48	0.00	0.00	1.00
$\Delta FX \mid$ *Foreign Trial*	3,918	-1.40	6.35	-4.60	-0.92	1.58

Table 2: Project Outcome and Delays in Trial Completion

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

	Dependent Variable: Suspension			
	(1)	(2)	(3)	(4)
Delay in Trial Completion	-0.220^{***}	-0.211^{***}	-0.195^{***}	-0.188^{***}
	(-5.29)	(-6.36)	(-5.60)	(-5.30)
Controls	N	N	N	Y
Year × Quarter FE	Y	Y	Y	Y
Firm FE	N	Y	Y	Y
ICD FE	N	N	Y	Y
Trial Phase FE	N	N	Y	Y
Observations	9,912	9,433	9,397	9,161
Adj. R-squared	0.1262	0.2592	0.3050	0.3060

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

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

	DV: Delay First Stage		**DV: Suspension** Second Stage	
	(1)	(2)	(3)	(4)
Trial Site Congestion	0.360^{***}	0.358^{***}		
	(4.26)	(4.37)		
Delay in Trial Completion			-0.334^{**}	-0.422^{**}
			(-2.42)	(-2.37)
Controls	N	Y	N	Y
Year × Quarter FE	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y
F-stat	18.13	19.13	–	–
Observations	9,397	9,161	9,397	9,161

Table 4: Project Outcome and Change in Exchange Rates

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

	Dependent Variable: Suspension			
	(1)	(2)	(3)	(4)
ΔFX	-0.279^{***}	-0.276^{***}		
	(-2.98)	(-2.83)		
$\Delta FX \times$ *Foreign Trial*			-0.309^{***}	-0.307^{***}
			(-3.24)	(-3.14)
Foreign Trial			-1.857^{*}	-2.331^{**}
			(-1.92)	(-2.04)
Trial Duration	-0.179^{***}	-0.170^{***}	-0.179^{***}	-0.172^{***}
	(-6.46)	(-6.03)	(-6.52)	(-6.14)
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	10,600	10,363	10,600	10,363
Adj. R-squared	0.3112	0.3104	0.3114	0.3108

Table 5: Underlying Channels

This table investigates underlying channels with respect to the results in Tables 2 to 4. 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. In Panel A, 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. In Panel B, the independent variable, ΔFX_i, is the average exchange rate change for clinical trial i defined in Equation (2). 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 4.4 for further details. Standard errors are two-way clustered at the ICD category and quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Delay in Trial Completion

	Dependent Variable: Suspension			
	Many Projects	CEO	Vs. Expect.	Vs. Fin. Constr.
	(1)	(2)	(3)	(4)
Delay in Trial Completion	−0.196***	−0.083	−0.206***	−0.185***
	(−5.09)	(−1.31)	(−5.49)	(−3.64)
× *Same CEO*		−0.262**		
		(−2.28)		
Same CEO		6.521***		
		(2.78)		
Anticipated Trial Duration			−0.085*	
			(−1.95)	
Constrained (WW)				2.961
				(0.69)
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	7,050	3,399	9,159	4,745
Adj. R-squared	0.2756	0.2552	0.3064	0.2616

Table 5: *Continued.*

Panel B: Change in Exchange Rates				
	Dependent Variable: Suspension			
	Many Projects	CEO	Vs. Expect.	Vs. Fin. Constr.
	(1)	(2)	(3)	(4)
ΔFX	-0.269^{**}	-0.091	-0.256^{**}	-0.301^{**}
	(-2.42)	(-0.40)	(-2.35)	(-2.16)
$\times$ *Same CEO*		-0.211		
		(-0.59)		
Same CEO		2.604		
		(1.39)		
Trial Duration	-0.181^{***}	-0.165^{***}	-0.202^{***}	-0.203^{***}
	(-6.32)	(-4.39)	(-5.47)	(-5.60)
Anticipated Trial Duration			0.112^{**}	
			(2.55)	
Constrained (WW)				2.699
				(0.70)
Observations	7,992	3,752	9,157	5,283
Adj. R-squared	0.2800	0.2522	0.3067	0.2644
Controls	Y	Y	Y	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

Table 6: Examining the Real Option Channel

This table investigates the hypothesis of real option values. Panel A examines the relationship between the log of expected drug sales (Million US dollars adjusted to 2017 by GDP delfators) after obtaining the FDA approval and $Delay in Trial Completion_i$, the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Columns (1) to (4) use four different sales measures (all take logs), including the expected first-year sales after launch, the sum of first three-year sales, the sum of first five-year sales, and the average annual sales. Panel B examines the relationship between the percentage of adverse reactions among trial participants for project i that already completed Phase I or II clinical trials and $Delay in Trial Completion_i$. Columns (1) to (4) use four different dependent variables, including the total percentage of adverse reaction reported in the completed trial (a participant might experience multiple adverse events such as headache, nausea, fatigue) relative to the total number of enrolled participants, the maximum percentage of adverse reaction among various adverse symptoms, the total percentage of severe adverse reaction (life-threatening reaction, for example, acute myelogenous leukemia, basal cell carcinoma, brain hemorrhage, etc.), and the maximum percentage of severe adverse reaction among various adverse symptoms. We adopt the specification following Equation (3.2). Fixed effects are indicated in the bottom rows. Control variables include the number of sites in a clinical trial (taking logs), count of enrolled participants (in hundreds), number of drug projects in the company's pipeline (taking logs), and number of competing projects in the same ICD category (taking logs). Standard errors are two-way clustered at the ICD category and quarter levels. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Expected Drug Sales

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

Panel B: Adverse Reaction in Trials

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

Table 7: Effects on New Drug Launches With No or Few Existing Medications (Orphan Drugs)

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

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

Table 8: Effects on Adverse Events Related to Approved Drugs

This table investigates the effects of delay-induced R&D investment distortions on adverse events related to approved drugs, using trial site congestion as an instrument for delay in clinical trial completion. The unit of observation is a drug project. The sample contains all approved projects that completed Phase I or II trials with detailed clinical trials records. The second-stage dependent variables are the count of adverse events reported by the FDA (taking logs) within the first and the first three years since drug launch, respectively. The instrument, $TrialSiteCongestion_i$, is the normalized change in the number of participants in the universe of trials on ClinicalTrials.gov between trial start and end date, taking the maximum across all trial sites of a given trial in the sample. The second-stage independent variable, $Delay in Trial Completion_i$, is the gap in months between the realized trial completion date and the expected trial completion date for project i to complete Phase I or II clinical trials. Please see Section 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)
Trial Site Congestion	0.903***		
	(5.25)		
Delay in Trial Completion		0.014	0.014
		(0.57)	(0.58)
Inverse Mills Ratio	−1.572	1.101	1.467
	(−0.12)	(0.54)	(0.66)
Controls	Y	Y	Y
Year × Quarter FE	Y	Y	Y
Firm FE	Y	Y	Y
ICD FE	Y	Y	Y
Trial Phase FE	Y	Y	Y
Observations	438	438	438

Online Appendix

Excess Commitment in R&D

Marius Guenzel
Wharton

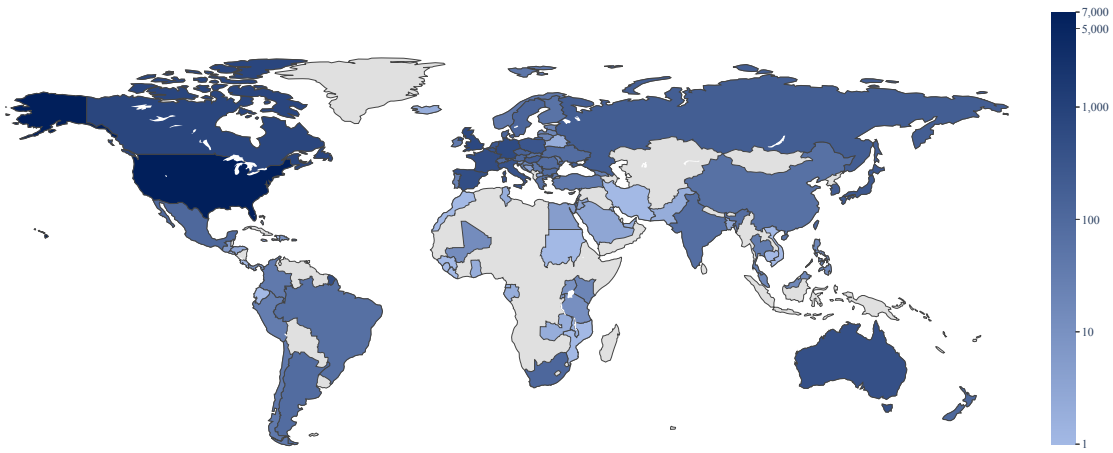
Tong Liu
MIT Sloan

Not For Publication

I. 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.

I.B Tables

Table OA.1: Phase I Versus Phase II Trials

This table separates the results in Tables 2 and 4 by trial phase. 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 variables, $Delay in Trial Completion_i$ and ΔFX_i, are 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, and the average exchange rate change for clinical trial i defined in Equation (2), respectively. 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.

	Dependent Variable: Suspension			
	Phase I	Phase II	Phase I	Phase II
	(1)	(2)	(3)	(4)
Delay in Trial Completion	-0.152^{**}	-0.181^{***}		
	(-2.63)	(-4.68)		
ΔFX			-0.284	-0.280^{**}
			(-1.10)	(-2.14)
Trial Duration			-0.143^{***}	-0.167^{***}
			(-3.01)	(-6.48)
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	3,356	5,414	4,055	5,890
Adj. R-squared	0.3653	0.2877	0.3545	0.2937

Table OA.2: Project Outcomes and Delays in Trial Completion: 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 4.4 for further details. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: Suspension					
	(1)	(2)	(3)	(4)	(5)	(6)
Delay in Trial Completion	−0.185***	−0.168***	−0.187***	−0.183***	−0.181***	−0.179***
	(−3.64)	(−3.51)	(−3.65)	(−4.20)	(−4.06)	(−4.05)
Constrained (WW)	2.961					
	(0.69)					
Constrained (HP)		17.184***				
		(3.91)				
Constrained (KZ)			−1.937			
			(−0.79)			
Constrained–Trial Start (WW)				1.867		
				(0.59)		
Constrained–Trial Start (HP)					8.369	
					(1.63)	
Constrained–Trial Start (KZ)						2.761
						(1.12)
Controls	Y	Y	Y	Y	Y	Y
Year × Quarter FE	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y	Y	Y
Observations	4,745	4,854	4,758	4,985	5,172	4,994
Adj. R-squared	0.2616	0.2683	0.2632	0.2551	0.2585	0.2564

Table OA.3: Project Outcomes and Delays in Trial Completion: Further Evidence on Financial Constraints

This table further explores the role of financial constraints for explaining the relationship between trial suspension and exchange rate changes in the preceding trial phase, adding to the evidence in the final column in Panel B 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, ΔFX_i, is the average exchange rate change for clinical trial i defined in Equation (2) 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 4.4 for further details. We report t−statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: Suspension					
	(1)	(2)	(3)	(4)	(5)	(6)
ΔFX	−0.301**	−0.283**	−0.307**	−0.260*	−0.228*	−0.254*
	(−2.16)	(−2.08)	(−2.22)	(−1.94)	(−1.69)	(−1.89)
Trial Duration	−0.203***	−0.187***	−0.205***	−0.183***	−0.179***	−0.180***
	(−5.60)	(−5.42)	(−5.54)	(−5.55)	(−5.33)	(−5.47)
Constrained (WW)	2.699					
	(0.70)					
Constrained (HP)		14.365***				
		(3.60)				
Constrained (KZ)			−0.778			
			(−0.35)			
Constrained–Trial Start (WW)				3.766		
				(1.18)		
Constrained–Trial Start (HP)					8.499*	
					(1.78)	
Constrained–Trial Start (KZ)						3.835
						(1.52)
Controls	Y	Y	Y	Y	Y	Y
Year × Quarter FE	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y
ICD FE	Y	Y	Y	Y	Y	Y
Trial Phase FE	Y	Y	Y	Y	Y	Y
Observations	5,283	5,405	5,298	5,573	5,792	5,581
Adj. R-squared	0.2644	0.2690	0.2658	0.2581	0.2595	0.2594

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 preceeds 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.4 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).

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

Table OA.4: Examples

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

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

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

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

By definition,

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

Given the assumption in Equation (7), we have

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

Therefore, the conditional expectation of v_i is

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

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

Applying iterated expectations on Equation (8),

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

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