Valuing Pharmaceutical Drug Innovations: An Event Study Approach *

Gaurab Aryal† Federico Ciliberto‡

Leland E. Farmer§ Ekaterina Khmelnitskaya¶

December 14, 2022

Abstract

We measure the *value of pharmaceutical drug innovations* by estimating the market values of drugs and their development costs. We rely on market responses to firms' drug development announcements to identify the values and costs. Using announcements by public U.S. pharmaceutical firms and their daily stock returns, we estimate the average market value of a successful drug at $1.62 billion. At the discovery stage, drugs are valued at $64.3 million, whereas their expected development cost is $58.5 million, on average. Furthermore, we estimate the costs of the three phases of clinical trials at $0.6, $30, and $41 million, respectively.

JEL: L65, O31, G14.

*We thank Lam K. Bui, Will K. Nehrboss and Ishaan Dey for outstanding research assistance, and graduate and undergraduate students at UVA enrolled in empirical industrial organization classes for helpful comments. We also thank Bart Hamilton, Derek Lemoine, Tim Simcoe and the participants and discussants at the UVA, SEA 2021, MaCCI 2022, ASHEcon 2022, EARIE 2022, SciencePo, Indiana University, and BU TPRI. We acknowledge financial support from the Batten Research Grant Program Quantitative Collaborative and the Bankard Fund for Political Economy at UVA.

†Department of Economics, Washington University in St. Louis, aryalg@wustl.edu

‡Department of Economics, University of Virginia, DIW and CEPR, ciliberto@virginia.edu

§Department of Economics, University of Virginia, lefarmer@virginia.edu

¶Sauder School of Business, University of British Columbia, ekaterina.khmelnitskaya@sauder.ubc.ca

1 Introduction

A central question in the economics of innovation is how to incentivize firms to develop new products or production processes, what Arrow (1972) calls *new knowledge*, when research and development (R&D) is costly and prone to failures. These incentive considerations are salient in the pharmaceutical industry, where finding a safe and effective drug is difficult, time-consuming, expensive, and financially risky (Munos, 2009; Pammolli et al., 2011; Scannell et al., 2012; Barberio et al., 2022).

A widely adopted solution to this incentive problem is to provide innovators with exclusive property rights, e.g., patents (Herrling, 2007). Nevertheless, pharmaceutical innovations in new drugs are lacking (Paul et al., 2010). Moreover, property rights are inefficient (Boldrin and Levine, 2008) and can even generate *negative innovation* (Feldman et al., 2021).

Thus, it is important to determine ways to incentivize drug innovation while minimizing inefficiencies associated with property rights. Several proposed solutions (e.g., Boldrin and Levine, 2008) include patent buyouts (Kremer, 1998) and advance market commitment (Kremer and Glennerster, 2004). Successfully implementing these approaches requires knowing the *value of pharmaceutical drug innovations*.[1] However, determining market values and development costs of drugs is difficult because good data on R&D expenditures are not available to researchers. Furthermore, firms develop multiple drugs concurrently over a long horizon using hard-to-measure processes.

We propose a way to value pharmaceutical innovations that is intuitive and easy to implement. In particular, we propose an *event study*-based approach to estimate the market value of drugs using stock market reactions to drug development announcements. We also estimate the expected cost at different stages of development. These cost estimates are important in and of themselves as they help inform regulations, including price interventions (e.g., U.S. House of Representatives, 2021).[2]

[1] Kremer (1998) suggests "modified" second-price auctions with common values to implement patent buyouts. However, as we discuss later, there are several difficulties in implementing this auction.

[2] Click here for the report. (Last accessed November 10, 2022.)

To illustrate the intuition behind our approach, let us consider a stylized example. Suppose a new firm announces the discovery of a drug compound to treat asthma. The jump in the firm's market value immediately following this "surprise" announcement is equal to the expected net present value of the drug. In other words, the discovery announcement informs the market about cash flows that may accrue to the firm from selling the new asthma drug, and the stock price adjusts to reflect this new information (Fama, 1965; Samuelson, 1965; Fama et al., 1969). The higher the probability of success and the shorter the time to success, the larger the firm's value post-announcement, and vice versa.

In practice, most firms develop more than one drug, and for each drug, they may make several announcements during its development cycle. By tracking announcements about drug discoveries, FDA applications, FDA approval decisions, and discontinuations at different stages and the subsequent changes in the announcing firm's market value, we can estimate the values of drugs and their development costs.[3,4]

The first estimation step is determining the *abnormal* returns associated with announcements. For information on drug announcements and the name of the firm developing the drug, we use the Cortellis dataset from Clarivate, which contains information for more than 70,000 drug candidates. Here we consider only drugs developed by publicly traded companies and use stock market data from the Center for Research in Security Prices (CRSP).

To estimate abnormal returns associated with announcements, we use the *unrestricted market model* (see Campbell et al., 1997, Chapter 4.3). The abnormal returns associated with an announcement for a particular firm represent the part of the firm's daily return that the market return alone cannot explain. Then, for every firm-drug-disease-announcement combination, we cumulate the abnormal returns one day before and two days after the announcement date to get the *cumulative abnormal return* (CAR).

The weak form of the efficient markets hypothesis implies that the CAR associated with

[3]The uncertain nature of drug development makes the announcement dates difficult to know in advance.

[4]Event study approach is widely used across different fields in economics, see, e.g., Eckbo (1983); Whinston and Collins (1992); Card and Krueger (1995); Fisman (2001); Dube et al. (2011); Kogan et al. (2017); Boller and Scott Morton (2020); Langer and Lemoine (2020); Jacobo-Rubio et al. (2020), and Känzig (2021).

an announcement multiplied by the firm's market capitalization equals the change in that firm's market value due to the event mentioned in the announcement. However, that event is associated with only a specific drug, so this change in the market value is also the change in the drug's expected net present discounted value.

A drug's expected value depends on the stream of future profits the firm expects to earn when the drug reaches the market, the probability of success, and the *time to success*. Specifically, the distribution of time it takes a drug to gain marketing approval affects the expected rate with which future profits are discounted. We estimate success probabilities, i.e., *transition probabilities*, as the shares of drugs that pass each stage of the development process, and use the competing risk model (Aalen, 1976) to estimate the distribution of time to success, and thus expected discount rates.

We begin with the last announcement in the development process, i.e., FDA approval, to identify the value of a drug. We start with this stage because it requires the fewest assumptions, and we isolate the required additional assumptions when we consider identifying value and costs at earlier stages. The only uncertainty after the FDA application and just before the FDA decision is whether the drug will be granted marketing approval. So the change in the firm's market value equals the value of the drug adjusted by the risk of failure. Therefore, to estimate the value, it is sufficient to have a consistent estimate of the transition probability from FDA application to FDA approval and the CAR associated with FDA approval.

We find that the average expected market value of an approved drug is \$1.62 billion. For comparison, based on the sales data for launched drugs, we estimate that the average ten-year and fifteen-year discounted revenue is \$1.4 and \$1.99 billion per drug, respectively.

Next we consider the discovery stage, which includes pre-clinical research on the drug's safety and efficacy. We find that the expected cost of developing a drug is \$58.51 million, and the expected value is \$63.37 million, on average. This small expected surplus at the early stage of development is consistent with a low probability of success and a long time

gap between discovery and market access. Thus, in addition to the transition probabilities and CAR estimates, we also need expected discount rates, which in turn rely on additional assumptions, to estimate the distribution of the time to success.

We also estimate the expected cost of clinical trials, evaluated at different development stages, using announcements about discontinuation. Besides our previous assumptions, we assume that, on average, there is no negative selection between discontinued drugs and their (ex-ante) profitability. This assumption is reasonable insofar as the main reason for discontinuations is negative clinical trial results (DiMasi, 2013; Khmelnitskaya, 2022).

We estimate that at discovery, the expected cost of clinical trials is approximately $12.43 million. This amount reflects a low likelihood of reaching the costlier Phase II and Phase III clinical trials. Using the discontinuation announcements at different phases of clinical trials across drugs, we find the expected cost of Phase I clinical trials is $0.62 million. In contrast, these costs increase to $30.48 million and $41.09 million for Phases II and III, respectively. Unlike the expected value of drugs, the expected costs for each drug cannot be identified. Instead, we estimate only average development costs across several drugs.

Our estimates can inform policymakers in designing systems to support drug development by making a few concrete proposals. First, the government could implement *drug buyouts* – buy the manufacturing rights to FDA-approved drugs and put them in the public domain – using the estimated market value to determine the price. For drugs in earlier stages of development that the firm may discontinue, the government could design a cost-plus contract to incentivize the firms to develop the drugs. Under this system, we envision the government covering development costs and paying a "bonus" if development is successful. The payments could be based on our cost estimates. If, however, the firm cannot develop the drug, it can sell the rights to the government, and the government can implement multi-stage R&D contests, similar to the U.S. government's Small Business Innovation Research programs (Howell, 2017; Bhattacharya, 2021).

Related Literature. Our paper contributes to several strands of literature. First,

it contributes to the literature that estimates the value of innovations, starting with the important works assessing the value of patents and R&D by Giliches (1981); Pakes (1985) and Austin (1993). More recently, Kogan et al. (2017) show that the private value of patents is positively correlated with the quality of patents, measured by future citations. Furthermore, McKeon et al. (2022) consider the problem of valuing pharmaceutical patent thickets. For more examples of valuing patents and R&D, see Chan et al. (2001); Hall et al. (2005), and Munari and Oriani, eds (2011). Unlike the extant literature, we focus on drugs rather than patents and propose an easy-to-implement and intuitive approach to valuing drugs.

Second, by providing estimates of R&D costs, this paper contributes to the literature that evaluates the cost of bringing new drugs to the market. For an overview of the problem, see Congressional Budget Office (2021).[5] There is, however, considerable variation in reported estimates, ranging from $314 million to $2.8 billion (Wouters et al., 2020). Most of these studies, including DiMasi et al. (2016), use confidential survey data from a handful of pharmaceutical firms, and others, like Sertkaya et al. (2016), use data on expenses associated with running a clinical trial. While Sertkaya et al. (2016) calculates accounting costs associated with running a clinical trial study, we provide new estimates of the total discounted expected *economic* cost associated with these trials, which are more relevant costs when one is considering innovation incentives. Furthermore, we provide new cost estimates for different development stages, not just clinical trials, based on a more representative sample than the extant literature, without relying on confidential data.

Lastly, our paper complements the idea of patent buyouts (Kremer, 1998) because policymakers can use our method to price drugs and use the information for drug buyouts. The idea of replacing intellectual property with a system of prizes in which the innovator receives a lump sum and the innovation is put in the public domain dates back to at least Marshall (1890). However, our method of estimating the value of drugs requires weaker assumptions

[5]The importance of costs is captured by the CEO of Teva Pharmaceuticals during his testimony to the 116th U.S. Congress: *"In order for any pharmaceutical company to [R&D] new drugs or improve old ones, the price of successful medicines must reflect the high cost of ongoing [R&D] projects."*

than the modified auctions proposed by Kremer (1998) and has a few practical advantages over auctions. Auctions rely on voluntary participation and competitive bidding by firms with pertinent information about the value of the patent. In contrast, we rely on stock prices that aggregate information dispersed among a large pool of self-interested investors (Milgrom, 1981). Thus our approach may provide a more accurate valuation than the auction. Furthermore, drugs are often "protected" by patent thickets (Gupta, 2021; McKeon et al., 2022). So, valuing a drug, instead of valuing patents first and then aggregating, is easier than implementing multi-unit auctions with a common value. The interested reader should consult (Ausubel, 2004) for more about challenges in designing such auctions. One important caveat of both our approaches is that if there is a systematic plan to use market to estimate the appropriate premium ex post then it is likely that the market would eventually realize that its reactions or bids determines the premium, leading to indeterminacy.

2 Examples of Drug Development Announcements

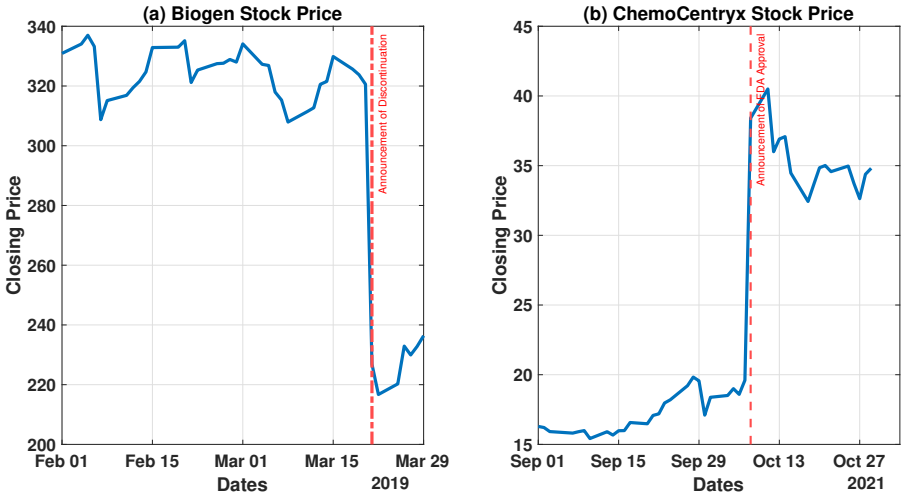
Here, we present examples of announcements, subsequent changes in stock prices, and timelines for five successful drugs in our sample. Before 2019, Biogen was developing what would have been a blockbuster drug called aducanumab to treat Alzheimer's disease. However, on March 21, 2019, it announced the discontinuation of Phase III clinical trials for aducanumab.[6] After this negative announcement, Biogen's stock price dropped sharply; see Figure 1-(a). We interpret this drop in Biogen's market value as what the market expected Biogen to earn from selling aducanumab minus the development cost no longer required.

In Figure 1-(b), we plot the share price of ChemoCentryx around a positive announcement. On October 8, 2021, ChemoCentryx announced FDA approval for its ANCA-associated vasculitis therapy, and the market responded positively to this "news shock."[7]

In Table 1, we present examples of five successful drugs, including information about the

[6] Click here for Biogen's announcement document. (Last accessed October 21, 2022.)
[7] Click here for Chemocentryx's (archived) announcement document. (Last accessed October 23, 2022.)

Figure 1: **Examples of Drug Announcements**



Note: Panels (a) and (b) display the time series of stock prices for Biogen and Merck around the announcement dates, respectively. On March 21, 2019, Biogen discontinued the Phase III clinical trial for a drug to treat Alzheimer's disease. On October 8, 2021, ChemoCentryx announced its FDA approval for a vasculitis drug.

dates associated with three development milestones. We can determine the time it takes for a drug to reach the market, and in some cases, we also have sales data that will allow us to evaluate our estimates of drugs' values. In the rest of the paper, we summarize institutional details and the data, and we formalize the idea that the change in a firm's market value in a tight window around announcements identifies drug value and cost.

Table 1: **Examples of Timeline**

	(1)	(2)	(3)	(4)	(5)
Drug Name	*caspofungin*	*dimethyl fumarate*	*evolocumab*	*telotristat etiprate*	*ziprasidone*
Firm	Merck & Co Inc	Biogen Inc	Amgen Inc	Lexicon-Pharma Inc	Pfizer Inc
Indication	Fungal Infection	Multiple Sclerosis	Hyper-cholesterolemia	Carcinoid Syndrome	Bipolar Disorder
Discovery	Jun 12, 1996	Nov 12, 2003	Jun 16, 2009	Feb 21, 2007	Jan 1, 2002
FDA Application	Dec 13, 2000	Feb 28, 2012	Aug 28, 2014	Mar 20, 2016	Aug 31, 2003
FDA Decision	Jan 1, 2001	Mar 27, 2013	Jul 17, 2015	Feb 28, 2017	Aug 31, 2004
Sales (15 years)	\$1.504 bil	\$45.6 bil	\$0.856 bil	\$0.268 bil	\$2.26 bil

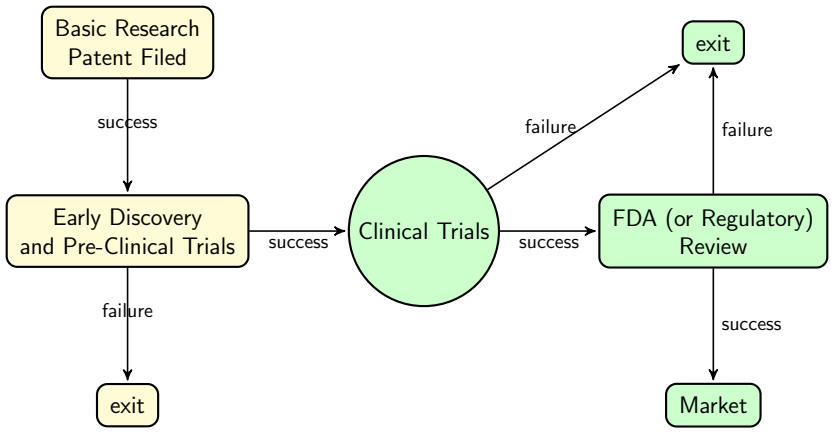
Note: This table displays key information for five successful drugs in our sample. For each drug, we observe its name, the firm developing the drug, the indication for which the drug is being developed, the dates for three stages of development, and discounted 15 years of sales revenue. The sales figures are real and expressed in December 2020 dollars.

3 Institutional Background and Data

3.1 Drug R&D and Announcements

The R&D process in the U.S. pharmaceutical industry consists of distinct stages defined by the FDA. Figure 2 is a schematic description of the drug development process. The first stage is a pre-clinical stage that includes creating a new molecule (or a system of molecules) and testing it (*in vitro* and *in vivo*) in the laboratory. We refer to this stage as "discovery."

Figure 2: **Drug Development Process.**



Note: This is a schematic representation of the drug development process, where the entries "success" and "failure" correspond to possible announcements by a firm.

If successful, the firm can test the drug candidate in humans during three phases, known as clinical trials. In Phase I of the clinical trials, the firm screens the drug for possible toxicity using a small sample of healthy subjects. In Phase II, the firm tests the drug for efficacy on a larger sample with targeted diseases. Finally, in Phase III, the firm uses double-blinded tests to assess the drug's effectiveness on a large sample.

After the successful completion of clinical trials, the firm can apply for FDA approval. The FDA has a group of internal and external experts who review the results from clinical trials and the manufacturing capacity of the applicant firm, then decide whether to approve the application. If the FDA approves the drug, the drug is launched in the market.

However, only some drug candidates reach the market. In Table 2, we present transition probabilities based on the frequency distributions, assuming, for simplicity, that the transition probabilities are homogenous across firms and diseases. On average, only about 11% of drug candidates are successful unconditionally. In contrast, 89% of drug candidates that apply for FDA approval are successful.[8] Conditional transition probabilities are estimated as the share of drugs that reached the next stage out of all the drugs in the current stage. Thus, the conditional probability of a drug reaching Phase II clinical trials, given that it has reached Phase I, is the ratio of the number of drugs that have reached Phase II to those that have reached Phase I.

Table 2: **Transition Probabilities**

	Probability of Reaching a Stage	
Stages	Marginal	Conditional
Phase I Clinical Trials	51.2%	51.2%
Phase II Clinical Trials	31.9%	62.4%
Phase III Clinical Trials	16.7%	52.4%
FDA Application	12.1%	72.3%
FDA Approval	10.8%	89%

Note: The unit of observation is a development project, i.e., a specific firm-drug-disease combination, associated with at least one announcement. The column labeled *Marginal* denotes the shares of all the initiated development projects, and the column labeled *Conditional* denotes the shares of the development projects that made it to the next stage. For example, 16.7% of all projects reached Phase III, and conditional in reaching Phase II, 52.4% made it to Phase III.

Announcements. Three regulatory paradigms affect firms' announcements about the success and failure of drug candidates as they transition through different development stages. First, the Security and Exchange Commission (SEC) requires all public companies to disclose all material information to their investors via the annual 10-K, quarterly 10-Q, and current 8-K forms. *Regulation Fair Disclosure*, instituted in 2000, requires publicly traded firms to disclose all material information on time.

Second, the FDA controls what a firm can and cannot announce about its drugs dur-

[8] A drug can be concurrently in development for multiple indications–reasons a medication might be used–by more than one firm. We define an observation to be at the individual firm-drug-indication level. Our estimate of 10.8% is comparable to estimates from the literature (Paul et al., 2010; Hay et al., 2014; Mullard, 2016; DiMasi et al., 2016; Wong et al., 2019), which range from 8% to 13.8%.

ing development. For instance, the *Food and Drug Administration Modernization Act* of 1997 established the centralized registry https://www.clinicaltrials.gov, and firms are required to register clinical trials within 21 days of enrolling the first subject. The FDA also requires firms to disclose information about their clinical trials and their FDA application processes whenever relevant. Furthermore, these announcements cannot be *materially misleading*.

Third, under the *Sarbanes-Oxley Act* of 2002, the SEC monitors firms' announcements about their FDA review process. Since 2004, the FDA can make direct referrals to the SEC if "in the normal course of their activities, they come to believe that a company may have made a false or misleading statement to the investing public."

These regulations incentivize firms to correctly and promptly inform the market and the general public. Nevertheless, it is up to companies to decide what is *material* and what is *not misleading*. This ambiguity is more pronounced when the results from clinical trials are "small" and for large companies developing multiple drugs. In such cases, firms may either delay announcements or bundle negative announcements with positive announcements to try and "soften" the market reaction.

However, we remain agnostic about the incentives for bundling because we focus on days with only one announcement by a firm. Furthermore, we consider only *major* announcements that are most certainly *material*. In particular, we use information only about discovery, whether a firm applies for FDA authorization to market the drug, the FDA's decision, and discontinuations at different stages.

International Announcements. Although most announcements in our sample are made in the U.S. and involve the FDA, U.S. firms also market drugs elsewhere, e.g., in the E.U. and Canada. We use announcements in all of the "western" countries. They all have similar rules governing drug development and announcements; see Chapter 7 of Ng (2015).

For instance, the drug development process in the E.U. is generally similar to that of the U.S. (Van Norman, 2016). The FDA counterpart in the E.U. is the European Medical

Association (EMA), which manages a centralized list of diseases. Firms seeking to develop drugs against those diseases must follow the centralized procedure to get approval. Firms submit a single authorization application to the EMA, and the centralized procedure allows marketing a drug based on a single E.U.-wide assessment of the drug.

However, most drugs authorized in the E.U. are approved by the member states. In such cases, when a company wants authorization to sell the drugs in several member states, it can apply for simultaneous approval in more than one member state if it has yet to be authorized anywhere in the E.U. The company can apply for this authorization to be recognized in other E.U. countries if authorized in one member state. Irrespective of the authorization routes, E.U. legislation requires that each member state apply the same rules regarding the authorization and monitoring of medicines. Thus, we expect U.S. subsidiaries operating in other western countries to behave similarly to their parent companies in the U.S. We further assume that all drugs must go through the same stages across all countries.[9] Henceforth, we use "FDA" as an omnibus term to refer to all relevant regulatory agencies.

3.2 Announcements Data

Our primary dataset on drug development comes from Cortellis, which is owned and managed by Clarivate Analytics. The database includes detailed information on more than 70,000 drug candidates in the development process worldwide. The database tracks each drug candidate's progression through different development stages using information from academic peer-reviewed articles, patents, press releases, financial filings, presentations, conferences, and FDA submissions. Almost half of all announcements in our final sample are from press releases and corporate publications.

The database also contains information about every development milestone for all drugs in the development process. Notably, it records the date when the information about each milestone was announced, the drug's names, the associated firm, and the target disease. The

[9]This assumption, while reasonable for drugs, does not hold for Class III medical devices, e.g., coronary artery stents, where the E.U. has a less stringent approval system than the FDA (Grennan and Town, 2020).

Table 3: **Announcements, by Development Stage**

	Announcements		Dates		Single Annoucement Dates	
	N	%	N	%	N	%
Discovery	12,053	62.2	7,828	67.6	5,582	67.4
Discontinued during Discovery	1,083	5.6	552	4.8	266	3.2
Discontinued during Phase I	1,168	6.0	635	5.5	232	2.8
Discontinued during Phase II	1,648	8.5	910	7.9	414	5.0
Discontinued during Phase III	565	2.9	435	3.8	234	2.8
FDA Application	1,277	6.6	1,017	8.8	763	9.2
Discontinued after Application	98	0.5	84	0.7	49	0.6
FDA Approval	1,473	7.6	987	8.5	741	8.9
Total	19,365	100.0	11,576	100.0	8,281	100.0

Note: The table displays the count (and share) of different types of announcements. The column "Dates" refers to unique dates associated with announcements, and "Single Announcement Dates" refers to unique dates with a single announcement.

database is regularly updated by professional analysts working for Clarivate Analytics, who ensure the consistency and accuracy of the data.

In Table 3, we show the number of different announcements in our sample. For almost 30% of the dates, firms make more than one announcement. In Table 4, we display the distribution of the number of announcements per day. The median number of announcements is 1, but there is a long right tail because the number of announcements depends on firms' drug portfolios.

Table 4: **Summary Statistics for Daily Announcements**

Announcements	Mean	Med	90%	Std. Dev.	Min	Max
All	1.67	1	3	2.24	1	88
Discovery	1.04	1	2	1.81	0	86
Discontinued during Discovery	0.09	0	0	0.62	0	20
Discontinued during Phase I	0.1	0	0	0.57	0	13
Discontinued during Phase II	0.14	0	0	0.74	0	29
Discontinued during Phase III	0.05	0	0	0.28	0	5
FDA Application	0.11	0	0	0.44	0	17
Discontinued after Application	0.01	0	0	0.11	0	6
FDA Approval	0.13	0	0	0.69	0	38

Note: Summary statistics for the number of daily announcements by type.

Due to data limitations, we do not use announcements about the start of clinical trials.

The database provides information about the dates associated with the start and end of clinical trials, but usually *not* about the dates when the firm announces the *results* from the clinical trials. For example, a firm may reveal that it completed the Phase II clinical trial for a drug at a conference, but we may not know if this was the first announcement about Phase II results. Similarly, we may observe the start date of (for example) a Phase II clinical trial, but the firm may have already announced the date when disclosing Phase I results. We cannot determine firms' market values without the exact date information was revealed to the public.

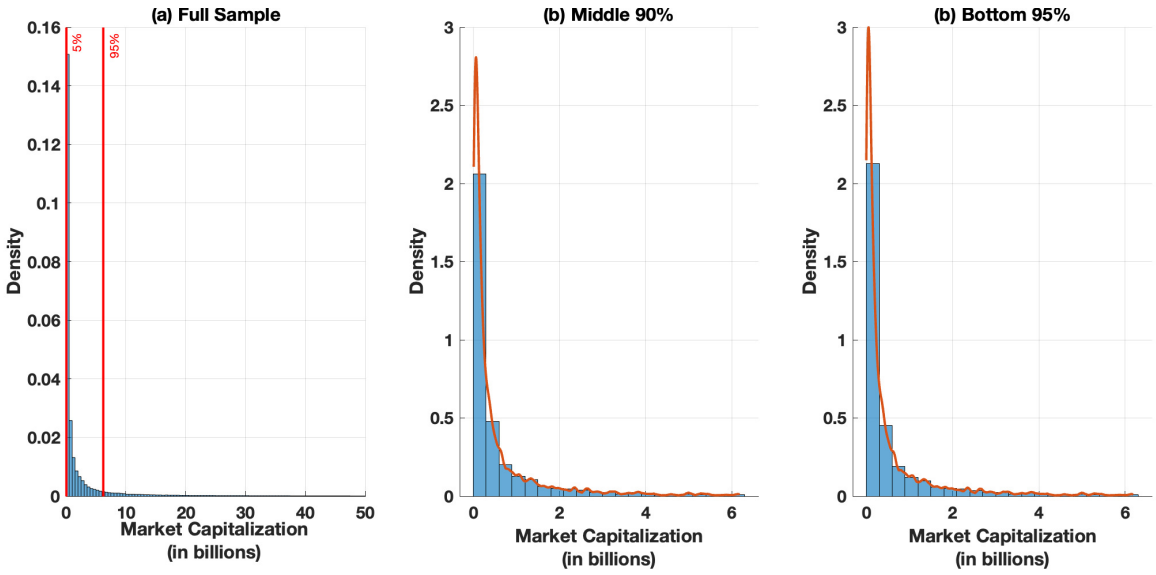
3.3 Market Value of a Firm

We observe daily returns for all biomedical and pharmaceutical companies publicly listed on U.S. stock exchanges from CRSP. We merge the stock market data with the Cortellis data by matching firms' names. However, firms may change their names, merge with another firm, or be acquired by another firm. To keep track of unique firms, we use CRSP-generated permanent I.D. numbers. We match if any name the firm has had in its history (i.e., any name associated with the CRSP-generated permanent I.D.) matches the name in the Cortellis data. Our matching process results in an unbalanced panel.

We define the "market return" as the return on the CRSP value-weighted portfolio (including dividends). We use a firm's market capitalization as a measure of its size. We determine each firm's median real market capitalization across years.[10] Figure 3 plots histograms of firm sizes for different subsamples. In the full sample of firms in Figure 3-(a), market capitalization exhibits a long right tail. The vertical lines denote the 5th ($11.6 mm) and 95th ($6.2 bil) percentiles of market capitalization, respectively. Figures 3-(b) and (c), respectively, display the density of firms restricted to be between the 5th and 95th percentiles of market capitalization, i.e., the "Middle 90%" sample, and below the 95th percentile of market capitalization, i.e., the "Bottom 95%" sample.

[10] We use December 2020 consumer price index to deflate all dollar amounts.

Figure 3: **Market Capitalization**



Note: The figure shows histograms of (real) market capitalization (in billions of U.S. dollars) of all the firms in panel (a), for the firms whose real market capitalization is between the 5th ($11.6 million) and 95th ($6.2 billion) percentiles in panel (b), and for the firms whose real market capitalization is below the 95th percentile in panel (c).

There is a long-standing comparison between the innovation performance of large and small firms; see Arrow (1983); Holmström (1989); Arora et al. (2009) and Akcigit and Kerr (2018). For our approach, heterogeneity between small and large firms can be crucial in several ways. Large firms might have more expertise and resources in conducting clinical development, increasing their chances of success and implying heterogeneity in transition probabilities. For the same reasons, large firms might be able to conduct their clinical trials faster, implying differences in the time investors have to wait to realize their returns.

Similarly, because of the differences in scope from other firms, they may select and develop certain types of drugs (MCockburn and Henderson, 2001; Krieger et al., 2022). Finally, heterogeneity in market capitalization implies that an announcement about a drug with a specific value would result in a different percent change in the price of a single stock. Therefore, including large firms in our sample with all other firms may affect our estimates. So, in the empirical exercises below, we focus on the middle 90% and bottom 95% samples, although for completeness, we also show the drug values estimated using the full sample.

4 Drug Valuation and the Cost of Development

To illustrate the main intuition behind our approach, let us re-consider the example mentioned in the introduction. Suppose a firm without any drugs in development announces the *discovery* of a new drug candidate to treat asthma. The product of the CAR and market capitalization (henceforth, `mktcap`) on the announcement day is the change in the firm's market value. Because the only "news" that pertains to the firm was the discovery announcement, the change in the firm's market value equals the expected net present discounted value of the asthma drug. The latter is the difference between the expected revenue and the R&D cost.

The expected net present discounted value of the drug at the time of discovery equals the discounted expected profit from the market adjusted for the probability that the drug will transition from discovery to FDA approval. A drug generates profits only after it is authorized to go on the market. However, the time of approval is uncertain (see Table (1) for examples). We use a stochastic rate to discount the expected stream of profits to the time of drug discovery. Ultimately, under the assumption that the expected profit is constant, we can express the expected net present discounted value of a drug as the product of (a) the present discounted profit from the market onwards, (b) the probability that the drug will transition from discovery to the market, and (c) the stochastic discount rate.

Next, suppose the firm initiates the FDA review process. At this stage, the market updates the expected value of the drug because even though the expected cash flow has not changed since the discovery announcement, the probability of FDA approval has increased. Moreover, the time until the drug may be authorized to enter the market is much shorter. Here too the product of CAR and `mktcap` at the time of announcement of FDA application is equal to the change in the probability of success times the drug's expected profit discounted to the time of the FDA application.

Finally, the FDA determines whether to approve the drug for distribution. If it is approved, all uncertainty surrounding the drug development process is resolved. The product

of CAR and `mktcap` is equal to the product of the probability of failure and the expected present discounted sum of the variable profit.

Most drugs do not succeed and are discontinued at different stages of the development process. If the firm announces that it is stopping the development, the product of the CAR and the `mktcap` should be negative, reflecting the decrease in the firm's value following this "bad news." This change in the firm's market value is informative about the remaining expected development cost.

Our approach is to determine changes in market value for each announcement and use the relationships described above to infer drugs' values and development costs. The first step is to estimate the CAR associated with each announcement. Then, in the second step, we use a linear regression model to predict the CAR as a function of the announcement type. These predicted values for CAR, along with the `mktcap`, allow us to determine the expected change in the firm's value immediately following an announcement.

4.1 Model

Let the random variable V denote a firm's market value and C the firm's development cost. There are three stages for which we have reliable data on the dates when an announcement was made: discovery, FDA application, and FDA decision, respectively, denoted by $k = \texttt{disc}, \texttt{appl}$, and $\texttt{appr}$. We also have reliable data on drug discontinuations, denoted by `drop`. The cost is paid at different stages of development. To capture that, we use notations such as $C_{\texttt{disc}\rightarrow}$ and $C_{\texttt{disc}\rightarrow\texttt{appl}}$ to denote the cost from discovery to approval and from discovery to FDA application, respectively.

We begin by introducing some additional notation. Let $S_k \in \{0, 1\}$ denote the stage $k \in \{\texttt{disc}, \texttt{phase I}, \texttt{phase II}, \texttt{phase III}, \texttt{appl}, \texttt{appr}\}$ announcement about the success or failure of a drug, with the interpretation that $S_k = 1$ denotes success in stage k and $S_k = 0$ denotes failure. Our sample starts from the discovery stage, so by definition, $S_{\texttt{disc}} \equiv 1$. For any two consecutive stages k and k', let $p_{k'|k}$ denote the conditional probability that the

drug is successful in stage k', given that it is successful in the previous stage k. For example, $p_{\texttt{appl|phase III}}$ is the probability that a drug reaches the FDA application stage conditional on completing Phase III clinical trials. Likewise, $(1 - p_{\texttt{appl|phase III}})$ is the probability that the firm discontinues before submitting an FDA application after finishing Phase III trials.

We now show how we can use changes in a firm's market capitalization following an announcement about a drug to learn about the drug's value. When a firm announces that the FDA has approved a drug, the firm's value should jump immediately after the announcement, with the size of the jump equaling the increase in the *expected* profits from selling the drug. However, the only relevant change in the drug's status since the last announcement is no more uncertainty about the development process. In other words, the value of the drug, which is the expected present value of market profits conditional on regulatory approval, equals the increase in the firm's market value adjusted by the ex-ante probability of success. In particular, we have the following equation associated with announcements about FDA approval:

$$\begin{aligned}\mathbb{E}\left(\texttt{CAR}_{\texttt{appr}}\right) \times \texttt{mktcap} = & \underbrace{\mathbb{E}(V|S_{\texttt{appr}}=1)}_{\texttt{after the announcement}} - \underbrace{\mathbb{E}(V|S_{\texttt{appr}}=1) \times p_{\texttt{appr|appl}}}_{\texttt{just before the announcement}} \\ & = \mathbb{E}(V|S_{\texttt{appr}}=1)(1 - p_{\texttt{appr|appl}}). \end{aligned} \tag{1}$$

We can estimate the expected CAR using stock prices and announcements data, as we show later. Furthermore, we observe the market capitalization from CRSP. Substituting $p_{\texttt{appr|appl}}$ from Table 2 in (1), we can identify the value of a drug, $\mathbb{E}(V|S_{\texttt{appr}} = 1)$ at the time of approval. Notice that there is no cost at approval because the development costs would have been paid by then. The costs necessary to launch and market the drug are included in this estimate.

It is helpful to introduce new notation regarding discounting rates and profits. Let $\mathbb{P}_{\texttt{appr}}(\tau|S_k = 1)$ be the probability that a drug will get FDA approval within the next $\tau \in \mathbb{N}$ years conditional on reaching stage k of development. For example, $\mathbb{P}_{\texttt{appr}}(5|S_{\texttt{appl}} = 1)$ is the

probability that a drug gets FDA approval in the next 5 years, given that it has successfully submitted the FDA application. Hence, $\mathbb{P}_{\texttt{appr}}(5|S_k = 0) = 0$ for all k. Let $\delta \in (0,1)$ be the annual discount factor, and let $\mathbb{E}(\delta^{\tau_k \rightarrow}) := \sum_{\tau \geq 0} \delta^\tau \times \mathbb{P}_{\texttt{appr}}(\tau|S_k = 1)$ be the expected discount rate when the drug is at stage k. In Section 5.3, we use a competing-risk model to estimate the time-to-success $\mathbb{P}_{\texttt{appr}}(\tau|S_k = 1)$ for all k.

To help interpret the expected value of a drug at approval, we introduce new notation and make some simplifying assumptions. Let π be the expected (average) yearly profit from selling the drug *after* FDA approval. The per-period profits do not include the sunk costs faced by the firm after the drug is approved but before it is marketed. Furthermore, suppose there is no change in the fundamental value of a drug between discovery and approval. Then, the expected value of a drug is the present discounted profits, i.e., $\mathbb{E}(V|S_{\texttt{appr}} = 1) := \frac{\pi}{1-\delta}$.[11]

So the difference in the expected drug value at discovery and approval is entirely due to uncertainty about success and discounting. Then we can show that, given the transition probabilities and the expected discount rate, we can estimate the expected value of the drug when the discovery announcement is made. The expected value of the drug at discovery is given by

$$\begin{aligned}
\mathbb{E}(V|S_{\texttt{disc}} = 1) &= \left(\sum_{\tau \geq 0} \left(\sum_{t=\tau}^{\infty} \delta^t \pi\right) \times \mathbb{P}_{\texttt{appr}}(\tau|S_{\texttt{disc}} = 1)\right) \times p_{\texttt{appr}|\texttt{appl}} \times p_{\texttt{appl}|\texttt{disc}} \\
&= \left(\sum_{\tau \geq 0} \pi \times \left(\sum_{t=\tau}^{\infty} \delta^t\right) \times \mathbb{P}_{\texttt{appr}}(\tau|S_{\texttt{disc}} = 1)\right) \times p_{\texttt{appr}|\texttt{appl}} \times p_{\texttt{appl}|\texttt{disc}} \\
&= \frac{\pi}{1-\delta} \times \left(\sum_{\tau \geq 0} \delta^\tau \times \mathbb{P}_{\texttt{appr}}(\tau|S_{\texttt{disc}} = 1)\right) \times p_{\texttt{appr}|\texttt{appl}} \times p_{\texttt{appl}|\texttt{disc}} \\
&= \mathbb{E}(V|S_{\texttt{appr}} = 1) \times \mathbb{E}(\delta^{\tau_{\texttt{disc}} \rightarrow}) \times p_{\texttt{appr}|\texttt{appl}} \times p_{\texttt{appl}|\texttt{disc}}. \qquad (2)
\end{aligned}$$

We rely on two implicit assumptions for the derivation: profits when the drug is marketed are constant across periods and firms receive these constant profits forever. We make

[11] The constant per period profit assumption is to simplify the interpretation of the expected value, and can be relaxed without affecting our procedure.

these assumptions for exposition and to simplify the interpretation of the expected value. These assumptions can be relaxed to allow profits to vary across years without affecting our procedure.

Once we have $\mathbb{E}(V|S_{\texttt{disc}} = 1)$, we can use the discovery announcements to infer the expected development cost from discovery to FDA approval, i.e., $\mathbb{E}(C_{\texttt{disc}\rightarrow}|S_{\texttt{disc}} = 1)$. In particular, the change in the market value of a firm immediately following a discovery announcement is equal to the difference in the expected value of a drug at discovery and the expected cost of drug development from discovery until FDA approval. Hence, we get

$$\mathbb{E}\left(\texttt{CAR}_{\texttt{disc}}\right) \times \texttt{mktcap} = \mathbb{E}(V|S_{\texttt{disc}} = 1) - \mathbb{E}(C_{\texttt{disc}\rightarrow}|S_{\texttt{disc}} = 1), \tag{3}$$

which then we can use to estimate $\mathbb{E}(C_{\texttt{disc}\rightarrow}|S_{\texttt{disc}} = 1)$.

In order to evaluate the costs associated with each stage of drug development separately, we use negative announcements about discontinuations at these different stages (see Table 3). For example, firms may announce that a drug has been discontinued after Phase I clinical trials and will no longer be developed. We can use these announcements to determine the remaining costs of development.

Let us consider identifying the costs of FDA application and FDA review. We defer the discussion of what is included in the costs for FDA review to Section 6.2. We can use discontinuation announcements after Phase III clinical trials to identify these costs. The intuition is as follows. Following a discontinuation announcement, the firm's market value falls by the amount that the market had expected the drug to earn (during the previous successful announcements) and increases by the cost savings that need not be incurred.

More formally, we begin with the following relationship:

$$\mathbb{E}\left(\texttt{CAR}_{\texttt{drop after phase III}}\right) \times \texttt{mktcap} = \underbrace{-\mathbb{E}(V|S_{\texttt{phase III}} = 1)}_{\texttt{value lost after discontinuation}} + \underbrace{\mathbb{E}(C_{\texttt{appl}\rightarrow}|S_{\texttt{appl}} = 1) \times p_{\texttt{appl|phase III}}}_{\texttt{cost savings}}, \tag{4}$$

Then, we note that the expected value of a drug at Phase III is given by $\mathbb{E}(V|S_{\texttt{phase III}} = 1) = \mathbb{E}(V|S_{\texttt{appr}} = 1) \times \mathbb{E}(\delta^{\tau_{\texttt{appl}\rightarrow}}) \times p_{\texttt{appr}|\texttt{appl}} \times p_{\texttt{appl}|\texttt{phase III}}$, where all the variables on the RHS are known. Then, plugging this back in (4) we can identify $\mathbb{E}(C_{\texttt{appl}\rightarrow}|S_{\texttt{appl}} = 1)$.

Although we do not show the derivations here, following similar steps, we can use discontinuations at earlier stages to identify the expected costs of Phase I, II, and III clinical trials using discontinuation announcements after discovery, and Phase I and Phase II clinical trials, respectively. Furthermore, we note that to identify the costs of earlier stages of development (for example) Phase II clinical trials, we rely on our estimates of values, transition probabilities, and the discount rates *and* the estimated costs of later stages, e.g., Phase III clinical trials, FDA application, and FDA review.

5 CAR and Expected Discount Rates

Our event study methodology is based on estimating the impact of each drug development announcement on the value of a firm. If there is only one announcement, the change in the firm's value, which we estimate, is also the expected net present value of the drug.

This relationship between the change in the firm's value on the day of the announcement and the expected net present value of the drug continues to hold even when a firm is developing more than one drug so long as there is only one announcement on a given day. In the remainder of this section, we formalize this idea and discuss how we estimate expected discount rates.

5.1 Cumulative Abnormal Returns

In this section, we show how we estimate the cumulative abnormal return (CAR) associated with each drug-firm-announcement combination. CAR accounts for the effect of predictable variation in the firm's value (using the market return) and cumulates over a window around the event to account for partial adjustment of stock prices to the new information.

Then we use a linear regression model to decompose CAR as a function of different types of announcements, i.e., discovery, FDA application, FDA approval, and discontinuations at different stages. Using the estimated regression coefficients, we determine the expected change in the firm's market value given an announcement and apply it to determine the value of drugs and the average cost of drug development using the method detailed in Section 4.

It is helpful to introduce some notation and then define CAR. Let $r_{i,t}$ denote the stock return of firm $i \in \{1, \ldots, I\}$, at date $t \in \{1, \ldots, T\}$ and r_t denote the market-wide return at date t, which we proxy for using the return on the CRSP value-weighted portfolio for that day. Let J_i denote the set of announcements by firm i, and we let $t_{i,j}$ denote the date when firm i makes its j^{th} announcement. If the announcement date is not a trading day, $t_{i,j}$ denotes the first trading date after the announcement.

We use an unrestricted market model and posit that firm i's log of returns $(r_{i,t})$ is a linear function of the log of the market return (r_t), i.e.,

$$r_{i,t} = \alpha_i + \beta_i r_t + \varepsilon_{i,t}. \tag{5}$$

For each firm-announcement date pair $(i, t_{(i,j)})$, we determine $\{r_{i,t}, r_t\}$ for a 200-day window that ends ten days before the announcement date $t_{(i,j)}$, and we fit (5) using linear regression.[12]

For each firm i, we estimate as many of these regressions as the number of announcements in J_i. From this estimation exercise, we obtain estimates $\{\hat{\alpha}_{i,j}, \hat{\beta}_{i,j} : j = 1, \ldots, J_i\}$ of $\{\alpha_i, \beta_i\}$ respectively, where we index the estimates with i and j to denote that these estimates differ by the firm and by the announcement. The abnormal returns associated with the $j \in J_i$ announcement are then the fitted residuals from (5), i.e., $\hat{\varepsilon}_{i,j,t} \equiv r_{i,t} - \hat{r}_{i,j,t}$.

Then the CAR associated with announcement j, by firm i on $t_{i,j}$ is the cumulative sum

[12]Stopping ten days before the announcement lowers the chance of having the announcement and the lead-up to the announcement contaminate abnormal returns. We use a 200-day estimation sample to account for possible time variation in the relationship between a firm's returns and market returns.

of $\hat{\varepsilon}_{i,j,t}$ around a pre-specified window $[w_l, w_u]$ and is given by

$$\widehat{\texttt{CAR}}_{i,j,t_{(i,j)}} = \sum_{t=t_{(i,j)}-w_l}^{t_{(i,j)}+w_u} \hat{\varepsilon}_{i,j,t}, \tag{6}$$

where $w_l > 0$ and $w_u > 0$ are the lower and upper window lengths of the event study, respectively.

For our estimation, we set $w_l = 1$ and $w_u = 2$, which means we aggregate the abnormal returns one trading day before the announcement and two trading days after the announcement. We estimate (5) for every firm-announcement pair separately and use the estimated $\hat{\varepsilon}$ to determine the associated CAR using (6). Table 5 reports the summary statistics of the CAR across all firms and all announcements. The average CARs have the expected signs: on average, the CAR is positive for good news and negative for bad news.

Table 5: **Summary Statistics of CAR**

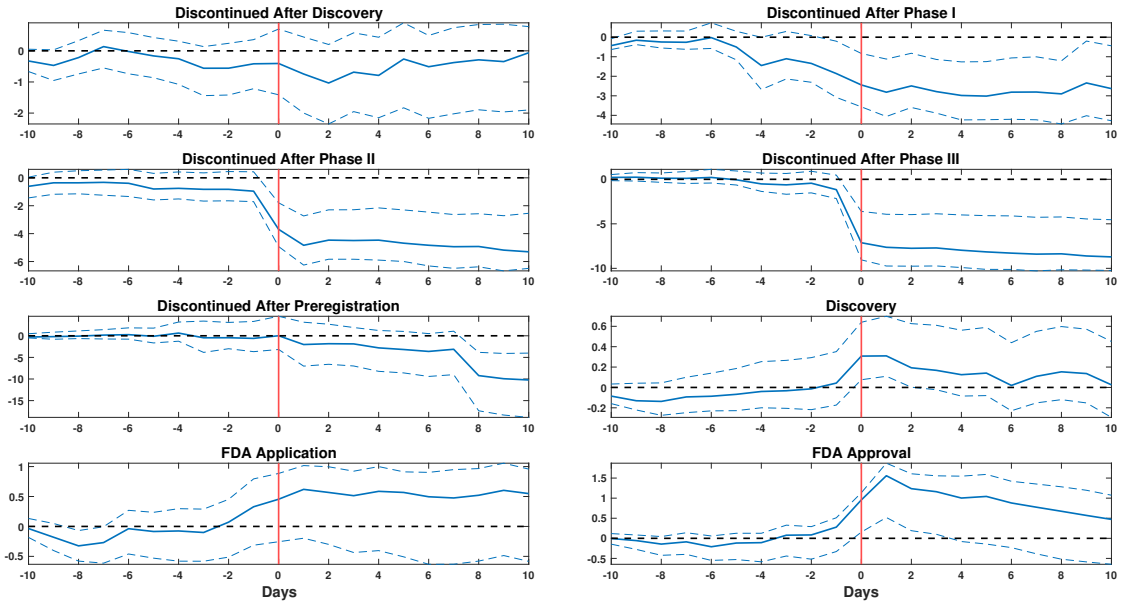
	Mean	**Median**	**Std. Dev.**	**Min**	**Max**	**N**
Overall	-0.225	0.010	10.917	-180.528	207.120	11,576
Discovery	0.221	0.031	9.105	-137.464	165.573	7,828
Discontiued During Discovery	-0.159	0.314	10.915	-126.511	57.733	552
Discontinued During Phase I	-1.333	-0.097	12.964	-124.009	66.176	635
Discontinued During Phase II	-4.002	-0.513	18.911	-175.680	66.176	910
Discontinued During Phase III	-6.556	-0.377	24.213	-180.528	36.380	435
FDA Application	0.290	0.154	6.913	-44.419	116.284	1,017
Discontinued After Application	-0.667	-0.007	8.014	-43.724	18.584	84
FDA Approval	1.039	0.291	9.793	-68.960	207.120	987

Note: Summary statistics for CAR as defined in (6), by each type of announcement.

We observe that the magnitudes of both positive and negative announcements increase as we move further in the development process. For example, the average CAR associated with the FDA approval is larger than that associated with FDA applications or discovery announcements. This pattern is consistent with the fact that in the later stages of development, the higher the chance of success, the smaller the remaining development costs, and the sooner the market launch. A mean CAR of 0.22 for Discovery announcements means that,

on average, a firm's share price increases by 22 basis points after announcing the discovery of a new drug. The interpretation for other types of announcements is similar.

Figure 4: **Cumulative Abnormal Return across Time**



Note: The figures display CAR, expressed in percentage points on the y-axes, for ten days' window around the event day (denoted as 0 and marked by red vertical lines), with a 90% confidence interval. Each figure corresponds to the type of announcement. We use those events with one announcement.

One concern with the announcements is possible information leakage before the official date of the announcement. Figure 4 investigates the possibility of such leakage in our sample. Each panel corresponds to a different type of announcement. The x-axis denotes event time, where 0 corresponds to the day an announcement occurred. Negative numbers are the days leading up to the announcement, and positive numbers are the days after the announcement. The y-axis measures CAR in percentage points. The solid blue line is the average CAR across all announcements of the specified type from 10 days before to 10 days after the announcement, the dotted blue lines are 90% bootstrapped confidence intervals, and the dotted black line is at 0.

According to the efficient markets hypothesis, we would only expect a non-zero CAR on

an announcement day because of the new information about the prospects of the relevant drug, e.g., increased probability of success. We would also expect the CAR to "jump" at time 0. Systematic increases or decreases in the CAR before an announcement suggest information leakage. As shown in the figure, the CAR in the days leading up to the announcements is statistically insignificant. There appears to be little evidence of information leakage across announcement types, as all the panels suggest that the primary change in CAR occurs on the day of the announcement.

If a firm develops more than one drug simultaneously, it may make multiple announcements about different drugs on the same day. In such cases, all the announcements have the same CAR, and we cannot separate the effect of each announcement on CAR. So, throughout the paper, we focus only on dates associated with single announcements.

5.2 Effects of Announcements on CAR

From Table 5, we see that the type of announcement affects the CAR. Instead of using the sample mean of the CAR, we use OLS by pooling across all firms, drugs, and time, to determine the marginal effects of different types of announcements on CAR, assuming that the effects are homogeneous across firms and drug candidates.

In particular, to determine the expected CAR for each type of announcement, we use the following linear model

$$\begin{aligned}\widehat{\texttt{CAR}}_{i,j,t_{(i,j)}} &= \beta_{\texttt{disc}} \times \texttt{disc}_{i,j,t_{(i,j)}} + \beta_{\texttt{appl}} \times \texttt{appl}_{i,j,t_{(i,j)}} + \beta_{\texttt{appr}} \times \texttt{appr}_{i,j,t_{(i,j)}} \\ &\quad + \beta_{\texttt{drop}} \times \texttt{discontinuations}_{i,j,t_{(i,j)}} + \omega_{i,j,t_{(i,j)}}, \end{aligned} \tag{7}$$

where the dependent variable is from (6), and `discontinuations` is a vector that includes separate indicators for discontinuation during discovery, Phase I clinical trials, Phase II clinical trials, Phase III clinical trials, and FDA applications. For example, if firm i's j-th announcement was made on date $t_{(i,j)}$ and if the announcement was the discovery of a drug,

Table 6: **Effects of Announcements on CAR**

	Full Sample	**Middle 90%**	**Bottom 95%**
Discovery	0.213	0.37	0.401
	[0.029, 0.420]	[0.029, 0.420]	[0.029, 0.420]
Discontinued during Discovery	-0.921	-2.429	-2.43
	[-2.239, 0.255]	[-2.238, 0.254]	[-2.238, 0.254]
Discontinued during Phase I	-1.150	-2.33	-2.319
	[-2.191, -0.157]	[-2.191, -0.157]	[-2.191, -0.157]
Discontinued during Phase II	-3.637	-7.63	-7.813
	[-5.199, -2.252]	[-5.198, -2.252]	[-5.198, .2.252]
Discontinued during Phase III	-7.310	-15.8	-15.809
	[-9.963, -4.626]	[-9.962, -4.625]	[-9.962, -4.625]
FDA Application	0.496	0.672	0.683
	[0.047, 0.953]	[0.047, 0.953]	[0.047, 0.953]
Discontinued after FDA Application	-1.384	-3.451	-3.451
	[-3.736, 0.850]	[-3.736, 0.849]	[-3.736, 0.849]
FDA Approval	1.158	4.017	4.017
	[0.547, 1.836]	[0.546, 1.836]	[1.836, 1.985]
Observations	8,281	3,968	4,032
$\overline{R}^2$	0.021	0.047	0.048

Note: The table presents estimated coefficients from Equation (7) using only single announcements. Each coefficient is followed by a 90% bootstrap confidence interval estimated using 1,000 bootstrap samples.

$\texttt{disc}_{i,j,t_{(i,j)}}$ is equal to 1 and the other right-hand-side variables in (7) for $t_{(i,j)}$ are equal to zero.

Thus each coefficient measures the marginal effect on the average CAR of a specific announcement. For instance, $\beta_{\texttt{appl}}$ is the change in expected CAR associated with the announcement that a firm has applied for FDA approval. The estimated coefficients from (7) are in the first column (full sample) of Table 6. To capture the uncertainty in the estimated coefficients, particularly the error in the estimation of $\widehat{\texttt{CAR}}$, we also present the 90% bootstrapped confidence intervals based on 1,000 bootstrap samples. The coefficients have the expected signs in line with our results for average CAR in Table 5.

To help interpret the coefficients from the regression, consider the following example. Suppose a firm with a market value of $100 mm announces discovering a new drug compound.

After the announcement, its market value is, on average, $\$100 \times (1 + \frac{0.21}{100}) = \$100.21$ mm.

In columns two and three of Table 6, we present estimates using the middle 90% and bottom 95% samples. While the estimates for these samples have the expected signs, compared to the full sample, we find that these restricted samples have larger estimated effects for all announcements. However, the discontinuations have particularly larger negative effects. These estimates are consistent with, all else equal, announcements having larger effects on smaller firms than larger ones, and the theory (Shin, 2003) that predicts that the return variance following a poor disclosed outcome is higher than it would have been if the disclosed outcome were good.

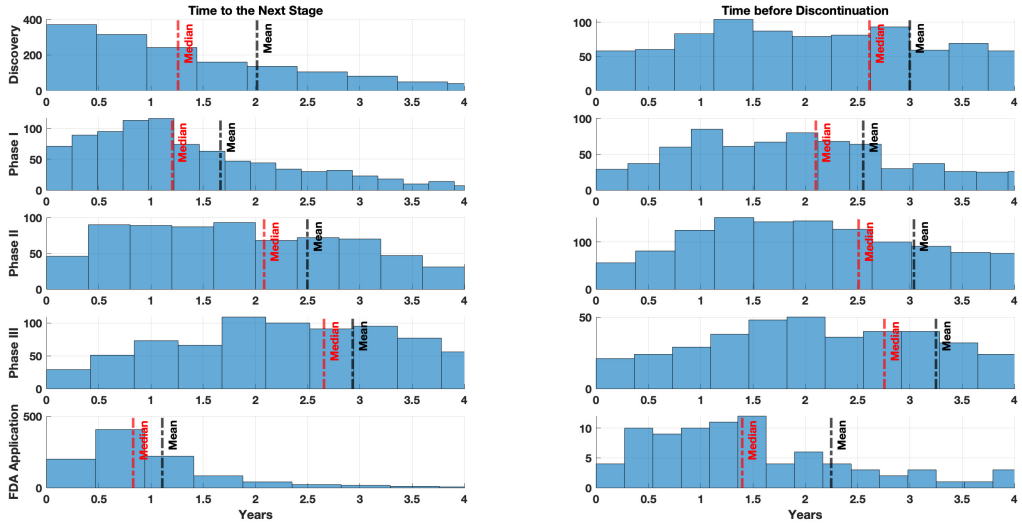
5.3 Expected Discount Rates

In this section, we present the estimates of discount rates for different stages. As we showed earlier in (2), to determine the expected value of a drug *at discovery*, $\mathbb{E}(V|S_{\texttt{disc}} = 1)$, we begin with its value *at approval*, $\mathbb{E}(V|S_{\texttt{appr}} = 1)$. Then we discount it to the discovery stage using transition probabilities and the discount rate $\mathbb{E}(\delta^{\texttt{disc}\rightarrow})$, where the expectation is taken at discovery with respect to the distribution of the time for FDA approval.

Let $\mathbb{P}_{\texttt{phase I}}(t|S_{\texttt{disc}} = 1)$ denote the probability that a drug will move to phase I clinical trials by year t given that it is starting at the discovery stage. Suppose we know $\mathbb{P}_{\texttt{phase I}}(t|S_{\texttt{disc}} = 1)$ and all subsequent development-stage time probabilities. In that case, we can determine the probability of *time to success* from discovery to FDA approval, denoted by $\mathbb{P}(\cdot)$. Then we can estimate the expected discount rate using Monte Carlo simulation, i.e., $\mathbb{E}(\delta^{\texttt{disc}\rightarrow}) = \mathbb{E}(\delta^{\tau}) \approx \frac{1}{L}\sum_{\ell=1}^{L}\delta^{\tau_\ell}$, where $\tau_\ell \sim \mathbb{P}(\cdot)$ is the time it takes for a drug to get approval from discovery. Therefore, to estimate the expected discount rate, we have to estimate the probabilities of time to success for all five stages.

To estimate the probabilities, we follow Aalen (1976) and use the observed time it takes for drugs to transition to the next stage (which we call "success"). Figure 5 shows the histograms of time (in years) it takes to move either to the next stage–time to success–or

Figure 5: **Time to Success and Discontinuation**



Note: Histograms of the time (in years) it takes for drugs to transition to the next stage, i.e., time to success, are in the first column, and when drugs are discontinued, i.e., time to failure, are in the second column. Each row corresponds to one stage, starting from Discovery in the first row to FDA application in the last. The two vertical lines denote the mean (in black) and the median (in red) of the time for that stage and event. For clarity of the presentation, the x-axes have been limited to only four years.

to discontinuation. Thus, there are "competing risks" at any time; a drug can be in its current state until it is either discontinued or succeeds and moves to the next stage. We refer to these three states as status-quo, failure, and success, by $\kappa = 0, 1, 2$, respectively. The idea is the same for $\mathbb{E}(\delta^{\texttt{clinic}\rightarrow})$ and $\mathbb{E}(\delta^{\texttt{appl}\rightarrow})$, except now we start from `clinic` and `appl`, respectively.

Let $Q_\kappa(t), \kappa = 0, 1, 2$ denote the probability that a drug stays in the state κ at time $t \in [0, \infty)$ given that it started at state 0 at time 0. The probability of transition from state 0 to state 2 by time t is given by $W(t) = 1 - \exp\left(-\int_0^t \frac{-Q_2'(y)}{Q_0(y)} dy\right)$.

Next, let $N_0(t)$ and $N_2(t)$ be the number of drugs in the development stage and for which the firms applied for approval at time t, respectively. Suppose that we partition $[0, t]$ into small intervals such that there is at most one transition in each subinterval. In an interval $(y, y+\eta]$, the conditional probability of one transition from development to FDA application, given that there are $N_0(y)$ drugs in the development stage, is equal to $N_0(y) \times \eta \times h_2(y)$.

Thus, if there is only one transition, we can estimate the transition probability $\eta \times h_2(y)$ by $\frac{1}{N_0(y)}$. Aalen (1976) shows that this intuition applies more generally and that we can estimate the hazard rate that defines $W(t)$ above by $\widehat{\frac{-Q_2'(t)}{Q_0(t)}} = \int_0^t \frac{1}{N_0(s)+1} dN_2(s)$.

In practice, we apply this method separately to five different subsamples: (i) drugs at the discovery stage, where the failure is discontinuation and the success is transition to Phase I clinical trial; (ii) drugs in Phase I clinical trials, where the failure is discontinuation and the success is transition to Phase II clinical trial; (iii) drugs in Phase II clinical trials, where the failure is discontinuation, and the success is transition to Phase III clinical trial; (iv) drugs in Phase III clinical trial, where failure is discontinuation and success is FDA application; and (v) drugs after FDA application, where success is FDA approval. These five exercises give us the transition probabilities from a preceding stage to the next. Using these estimates, we determine the expected discount rates using $\delta = 0.98$, which are in Table 7.

6 Drug Values and Costs Estimates

In this section, we use the estimates in the third column of Table 6 combined with regression Equation (7) to determine the change in the market value of a firm associated with each announcement. Then we apply the methodology presented in Section 4 to determine the value of a drug and the average cost of drug development.

6.1 Expected Value of Drugs

To determine the expected value of drugs at approval, i.e., $\mathbb{E}(V|S_{\texttt{appr}} = 1)$, we use all drugs that reach the market in Equation (1). Figure 6 shows the histograms of the expected drug values measured in billions of dollars for the sample of drugs for which firms submitted applications for FDA approval.[13]

[13]We could also use announcements about discontinuations after the FDA application to inform the estimates of drug values. However, there are only a few such announcements, and the CAR estimate associated with that event is noisier, so we do not use them to estimate drug values.

Table 7: **Expected Discount Rates**

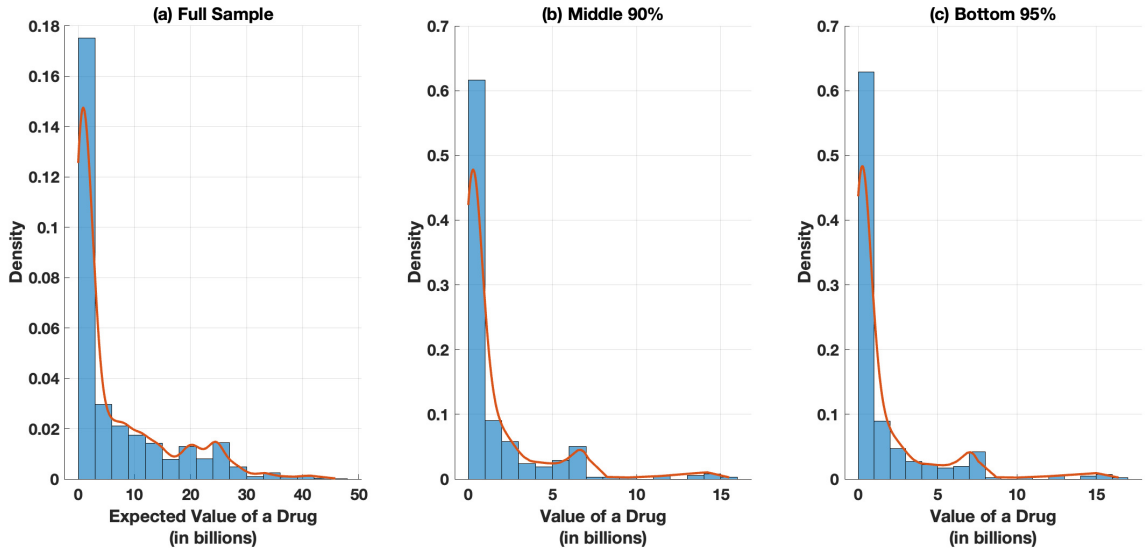
	Full Sample	**Middle 90%**	**Bottom 95%**
Discovery			
Discovery to Phase I	0.820	0.821	0.823
Discovery to Phase II	0.733	0.737	0.738
Discovery to Phase III	0.608	0.637	0.637
Discovery to FDA Application	0.507	0.554	0.555
Discovery to Market	0.448	0.482	0.483
Clinical Trials Phase I			
Phase I to Phase II	0.893	0.897	0.897
Phase I to Phase III	0.740	0.775	0.775
Phase I to FDA Application	0.618	0.674	0.675
Phase I to Market	0.546	0.586	0.587
Clinical Trials Phase II			
Phase II to Phase III	0.828	0.863	0.863
Phase II to FDA Application	0.692	0.751	0.752
Phase II to Market	0.611	0.653	0.655
Clinical Trial Phase III			
Phase III to FDA Application	0.835	0.870	0.871
Phase III to Market	0.737	0.756	0.757
FDA Application			
FDA Application to Market	0.883	0.869	0.870

Note: Expected discount rates for different stages, estimated using competing risk models, with a yearly discount rate of 0.98. The columns use full sample, middle 90% sample, and bottom 95% sample, respectively.

Figure 6-(a) considers only the drugs that received FDA approval in the full sample. There are 1,561 such drugs, and for these drugs, Equation (1) identifies their values. The estimated drug values range from $110,300 to (approximately) $45.8 billion, with a mean of $6.83 billion and a standard deviation of $8.88 billion. Column 1 of Table 8 reports the two estimates for the mean expected value at approval is $6.83 billion.

In light of the pronounced right skewness in the firm size distribution, we estimate the value of drugs after removing the outlier firms. The data restrictions are motivated by our discussion of the distribution of market capitalization in Section 3.3. We are concerned that large and small firms are fundamentally different from other firms. So, we remove firms with a market capitalization of either less than the 5th or more than the 95th percentile. We call

Figure 6: **Expected Value of Individual Drugs**



Note: The figure shows histograms and densities of $\mathbb{E}(V|S_{\text{appr}} = 1)$ (in billions of U.S. dollars) for drugs with FDA approval. Panel (a) corresponds to the full sample (1,561 drugs received FDA applications); Panel (b) corresponds to the middle 90% sample (378 drugs), and Panel (c) corresponds to the bottom 95% sample (401 drugs). To calculate values, we use the transition probabilities from Tables A.1, effects of announcements on CAR from Table 6, and the discount rates from Tables 7.

this sample "middle 90%." We also consider the sample where we remove only the large firms with a market capitalization larger than the 95th percentile. We call this sample "bottom 95%."

For these samples, we re-estimate the transition probabilities, the discount rates, and Equation (7). The estimates from CAR regressions are in Table 6, the expected discount factors are in Table 7, and the transition probabilities are in Table A.1. Figures 6-(b) and (c) show the histograms of valuations of successful drugs for the middle 90% (378 drugs) and bottom 95% samples (401 drugs). Unlike the full sample, these samples have fewer outliers.

The estimated values are reported in Table 8. We find that the mean of the expected values is \$1.62 billion among the 378 drugs developed by mid-90% firms, which is considerably smaller than in the full sample. For the bottom 95% sample, the average value of a drug is \$1.6 billion.

For a subsample of 84 drugs, we observe the "complete path" from discovery to FDA approval, and we can determine their values and costs without combining estimates across

Table 8: **Expected Value of Drugs, at Approval**

	Full Sample	Middle 90%	Bottom 95%
At Approval			
All Drugs	\$6.83 bil	\$1.62 bil	\$1.6 bil
Drugs with Complete Path	\$7.43 bil	\$1.89 bil	\$1.99 bil
At Discovery			
All Drugs	\$331.12 mm	\$63.37 mm	\$65.97 mm
Drugs with Complete Path	\$360.16 mm	\$74.05 mm	\$82.21 mm

Note: The table presents the mean of the expected value of individual drugs at the time of approval, $\mathbb{E}(V|S_{\texttt{appr}} = 1)$, and at discovery, $\mathbb{E}(V|S_{\texttt{disc}} = 1)$. The 90% sample refers to the drugs developed by firms with real market capitalization between 5% and 95% of the entire sample. The row, "Drugs with Complete Path" refers to the sample of drugs for which we observe discovery, FDA application, and FDA approval announcements. Of the 84 such drugs, 29 belong to the Middle 90% and Bottom 95% samples. The row "Average" refers to drugs for which we observe only a few stages.

other drugs for which we do not observe complete paths. Although we prefer to use all the drugs, for sensitivity analysis, we also consider this particular subsample of drugs and present the estimates in Table 8 in the row labeled "Drugs with Complete Path." The estimated values of 29 drugs with complete paths at \$1.89 billion and \$1.99 billion are similar to the estimates associated with all the approved drugs for the middle 90% and bottom 95% samples, respectively.

Once we have the expected value of the drug at the time of approval, we can also determine the value at the time of discovery using Equation (2). In particular, the value at discovery is the value at approval adjusted for the likelihood of a drug getting approval and discounted to discovery. These estimates are shown in the second half of Table 8. When we use the entire sample, we estimate the expected value at discovery to be \$331.12 million. When we consider the middle 90% and bottom 95% sample, the mean of the expected value at discovery is \$63.37 million and \$65.97 million, respectively.

To verify that the magnitude of the estimated values is reasonable, we have collected additional data on sales. The expected value is the NPV of the cash flow that accrues to the firm, so our estimates should be similar to what we would get from discounted sales data.

We use sales data from the Cortellis Competitive Intelligence database, which includes information on yearly drug-level total (worldwide) sales. The data were obtained from firms'

SEC filings and are relatively sparse.

The sales data are at the drug level, but the announcements are at the drug-disease level. We do not observe the breakdown of sales by disease, so we allocate the sales equally across all the diseases associated with the drug. So if a drug targets three diseases, each would be allocated one-third of the total sales.

Table 9: **Summary Statistics for the Sales**

Variable	N	Mean	Median	Std. Dev.	Min	Max
Full Sample						
Average yearly sales	764	268.24	69.90	519.57	0.10	6,655.11
# of years data available	764	11.57	10	8.84	1	34
Middle 90% and Bottom 95%						
Average yearly sales	148	155.79	32.94	317.31	0.10	1,791.74
# of years data available	148	6.50	5	5.97	1	23

Note: Sales (in millions of U.S. dollars) data on the firm-drug-disease level.

Table 9 presents the descriptive statistics for the sales data. Typically, the data are available for only about ten years, which is short for pharmaceutical sales, as many drugs stay on the market for over 20 years. To correct the short panel duration, we average sales across all the years available for a given drug-and-disease pair for a given firm. Assuming that this average sales value is received yearly, we calculate the discounted total sales for this drug-disease-firm pair.[14]

Table 10 presents the results for different time horizons after aggregating average sales. We find that the total discounted sales of the drug if the drug had been in the market for ten years is \$2.4 billion for the full sample and \$1.4 billion for the middle 90% and bottom 95% samples. Considering that generic drugs may drive the prices down, the observed sales presented in Table 10 support our approach to estimating the expected value of a drug.

[14]For example, suppose we observe sales for Wyeth's insomnia drug called zaleplon for three years. The sales data for these three years are \$54 million, \$133 million, and \$109 million (in 2020 U.S. dollars). The average sales value is then \$(54+133+109)/3=\$98.7 million. We then use these average sales to calculate the total discounted sales for a time horizon of length T employing as $\$98.7 \times (\sum_{t=0}^{T} \delta^t)$.

Table 10: **Total Discounted Sales**

Sample	10 yrs	15 yrs	20 yrs	25 yrs	30 yrs
Full Sample	\$2.4 bil	\$3.44 bil	\$4.37 bil	\$5.21 bil	\$5.97 bil
Middle 90% and Bottom 95%	\$1.4 bil	\$1.99 bil	\$ 2.54 bil	\$3.03 bil	\$3.47 bil

Note: Each entry shows total discounted sales averaged across drug-firm-diseases for a specific time horizon. The Full Sample refers to those drug-firm-diseases for which we have the sales data and are also present in our Full Sample for the value estimation. Middle 90% refers to those drug-firm-diseases present in the Middle 90% sample for the value estimation (drugs developed by firms whose real market capitalization is between 5% and 95% of the entire sample).

6.2 Cost of Drug Development

Total Cost at Discovery. Next we present the estimate of the expected *total* development cost at discovery, which includes the expected costs of clinical trials, FDA application, and the FDA review process. For that, we use Equation (3). In particular, by averaging both sides of (3) and substituting the average of the expected value of drugs at discovery, we can determine the expected cost of drug development at discovery for the middle 90% sample to be \$58.51 million, and for the bottom 95% sample, \$60.72 million.

Table 11: **Total Development Costs, at Discovery**

	Middle 90%	Bottom 95%
All Drugs	\$58.51 mm	\$60.72 mm
Drugs with Complete Path	\$69.24 mm	\$77.01 mm

Note: The table presents the mean of the expected cost of clinical trials and the FDA application and review process (in millions of U.S. dollars) at the time of discovery. The 90% sample refers to the drugs developed by firms with real market capitalization between 5% and 95% of the entire sample. The row "Drugs with Complete Path" refers to the sample of drugs for which we observe discovery, FDA application, and FDA approval announcements. There are 84 such drugs, out of which 29 belong to the Middle 90% and Bottom 95% samples. The row "Average" refers to drugs for which we do not observe the complete path but only a subset.

Notice that this cost is the risk-adjusted cost at the time of discovery. In other words, this cost factors in that with a high probability, the drug will not make it to the later, more expensive stages of the development, implying that such high costs likely are not incurred.

Cost of the FDA Review Process. Next, we use Equation (4) associated with discontinuations after Phase III clinical trials to estimate the expected cost of FDA application and review. Although the scientific experiments are mostly completed at the time of appli-

cation, there would still be additional expenses to set up manufacturing capacity and legal and administrative fees.[15]

The estimates are presented in Table 12. We estimate the expected cost to be $638.75 million for the middle 90% sample and $648.04 million for the bottom 95% sample. This cost estimate should be interpreted as the remaining cost the firm must face between the time the application to the FDA is submitted and when the drug is approved.

Table 12: **Cost of FDA Review and Application**

Middle 90%	Bottom 95%
$638.75 mm	$648.04 mm

Note: The table presents the mean of the expected cost of FDA review and FDA application at the time of discovery. Middle 90% sample refers to the drugs developed by firms with real market capitalization between 5% and 95% of the entire sample.

Clinical Trials. The last step is determining the costs the firm expects to pay for the clinical trials. To determine these costs, we can follow the steps outlined in Equation (4) and present the results in Table 13. Specifically, using the discontinuation announcements after discovery, we estimate the expected cost of Phase I clinical trials for the two samples to be $620.51 thousand and $219.24 thousand, respectively. The costs for Phase II clinical trials are substantially higher at $30.48 million (middle 90%) and $34.46 million (bottom 95%). To estimate these costs, we use the discontinuation announcements after Phase I. Similarly, using the discontinuation announcements after Phase II, we estimate the cost of Phase III clinical trials to be $41.09 million (middle 90%) and $39.71 million (bottom 95%).

Using these estimates of the costs of clinical trials, the transition probabilities (Table A.1), and expected discount rates (Table 7), we can determine the expected cost of running clinical trials where the expectation is at the time of drug discovery. We estimate the average cost of running clinical trials to be $11.8 million for the mid 90% sample and $12.43 million

[15]The FDA has prepared a set of instructions for drugs to receive approval. The instructions in Food and Drug Administration (2010) clarify that the "*FDA may approve an NDA or an ANDA only if the methods used in, and the facilities and controls used for, the manufacture, processing, packing, and testing of the drug are found adequate to ensure and preserve its identity, strength, quality, and purity.*"

Table 13: **Costs of Clinical Trials**

	Middle 90%	Bottom 95%
Phase I	\$0.62 mm	\$0.22 mm
Phase II	\$30.48 mm	\$34.46 mm
Phase III	\$41.09 mm	\$39.71 mm

Note: The table presents the mean of the expected cost of the three phases of clinical trials. Middle 90% sample refers to the drugs developed by firms with real market capitalization between 5% and 95% of the entire sample.

for the bottom 95% sample.

Cost Estimates in the Literature. It is helpful to compare our costs with prior estimates from the literature. For example, suppose we use the estimates of the transition probabilities, average durations for each phase, and average costs for each stage from DiMasi et al. (2016). In that case, the discounted expected cost of clinical development will be approximately \$40 million. While this estimate is larger than the comparable estimate of \$11.8 million we get using our preferred specification, estimates from DiMasi et al. (2016) are known to be larger than even other estimates in the literature (Wouters et al., 2020).

However, Sertkaya et al. (2014) suggest that the cost of bringing a new drug to a market is between \$161 million and \$2 billion. Relatedly, using proprietary cost data on a small sample of drugs developed from 2004 to 2012, Sertkaya et al. (2016) estimate the average non-risk-adjusted (accounting) cost of the three phases of clinical trials range from \$1.4 million to \$6.6 million, \$7.0 million to \$19.6 million and \$11.5 million to \$52.9 million, respectively. Finally, for drugs entering human clinical trials for the first time between 1989 and 2002, Adams and Brantner (2006) estimate the cost per new drug to be \$868 million.

7 Applying Our Estimates to Policy

One of the policy goals of a (democratic) government is to ensure widespread access to drugs at lower prices. One way to achieve that goal is by encouraging research on new drugs. However, using 60 years of data on FDA-approved drugs, Munos (2009) finds that despite

large outlays on R&D, the number of new drugs approved remains low. He suggests,

> Overcoming these difficulties will require bold initiatives, such as *open innovation*, that may take companies far from their comfort zone.

One such idea for "open innovation" is *drug buyouts*, where the government buys the drug and puts it in the public domain. Next, we explore how our estimates may be useful in improving drug R&D. For simplicity, we use estimates from the middle 90% sample with all drugs and focus only on average costs and values, but the core message extends to any of our other estimates.

Recall that the expected value and development costs depend on where the drug is in its development life cycle. So we consider two scenarios: (i) a drug is at its last stage and is approved by the FDA, and (ii) a drug is at the earlier stage of development, e.g., the discovery stage, but the firm may not develop the drug further.

The first case is the closest to patent buyouts (Kremer, 1998) because, as with a patent, after FDA approval, all uncertainty about the drug's efficacy is resolved, all costs are paid, and the only question remaining is the price of the drug buyout. Following our approach, one can learn about the expected value of the drug. Our estimates (Table 8) suggest that on average, the payment for the manufacturing rights should be at least \$1.62 billion.[16]

Next, let us consider the second case, when a firm announces the discovery of a new drug candidate. At that point, the firm faces considerable uncertainty about the possibility of success. Suppose further that during pre-clinical research, the firm decides not to develop the drug because of low expected net present discounted value. Suppose the social value of the drug is high. In that case, the government can use multi-stage prizes to incentivize the firm to continue the R&D process.[17] The government can use the cost estimates to

[16] If the total payment is less than the full value of the drug and excess burden of taxation (e.g., Hendren, 2020), this policy is Pareto improving.

[17] Kremer and Glennerster (2004)'s advance market commitment is another mechanism to stimulate drugs R&D. It has been successfully used to distribute pneumococcal vaccines and fund COVID-19 vaccines. Also see Kremer et al. (2020); Athey et al. (2020) for more. Kremer et al. (2022) study optimal design of AMC. Several other policies can promote innovation, e.g., compulsory licensing, used in the past to promote innovations successfully; see Moser (2013) for more.

design multi-stage prizes to support the development of the drug. Below, we discuss one such approach and keep the discussion informal.

Abstracting from potential asymmetric information and moral hazard between the government and the firm, we envision a system where the government offers a take-it-or-leave-it contract to develop the drug. Such a contract would promise to cover the development costs and award a prize after each successful development stage. In return, the government owns the drug's manufacturing rights, even if the firm is unsuccessful.

Suppose the firm successfully develops the drug and gets FDA approval. In that case, the government can buy the manufacturing rights at an average price of $1.62 billion minus the costs incurred during development. In other words, the government can make the firm the residual claimant of any gain, incentivizing the firm not to overrun costs.

Based on our estimates in Table 8, on average, the expected value of a drug at discovery is $63.37 million, and the development cost is $58.51 million. So, the government can offer a contract that pays $58.51 million to defray development costs at the start and an additional $4.86 million if the drug is successful, and additional amounts (determined below) after each successful step.

Next, suppose the drug completes pre-clinical trial research and is at the point of starting Phase I clinical trials. Then the firm is expected to incur $11.8 million in costs on clinical trials (see the discussion of Table 13), which the government can pay. Similarly, the cost estimates of clinical trials (Table 13) suggest minimum payments for late-stage clinical trials.

If the firm is unsuccessful, we envision that it will surrender its drug ownership. At that point, the government can set up an R&D contest among several other firms to develop it further. For more on designing procurement auctions for innovation, see Che and Gale (2003); Halac et al. (2017), and Ely et al. (Forthcoming) and references therein. An example where such contests have been used is the Small Business Innovation Research programs.

Despite the promise and the potential to improve R&D incentives, such a system is insufficient. For one, government interventions can crowd out private investment (Wallsten,

2000). Furthermore, instead of encouraging fundamental research, it may encourage socially wasteful R&D (Finkelstein, 2004) but not necessarily fundamental research, such as discovering new drug candidates.

Determining an optimal system that addresses these issues is beyond the paper's scope. Analysis of this type would require a significant extension to the theory and additional data to estimate the social value of drugs. However, we hope we have provided a starting point encouraging other researchers to work on these topics.

8 Conclusion

In this paper, we use the event study approach to estimate the market value of a drug and the costs to develop that drug, at different stages of development, from pre-clinical research to FDA approval. Our approach relies on the market responses to drug development announcements. In particular, the weak form of the efficient market hypothesis implies that the change in the market value of a firm from an announcement reflects the change in the expected net present value of a drug.

Using data from public firms, we estimate that on average, the market value of a drug at the time of FDA approval is approximately $1.62 billion. At the time of discovery, the value is $63.37 million, whereas the development cost is $58.51 million. We also estimate the costs of the three phases of clinical trials at approximately $0.62, $30.48, and $41.09 million, respectively. Thus we propose a tractable framework to value pharmaceutical innovations.

Our estimates suggest important areas for future research. For example, we can relax the assumption of homogenous transition probabilities and the distributions of time to success. These probabilities likely depend on the firm and disease characteristics (e.g., cancer vs. hypertension). Progress on this question could be made by estimating a semiparametric competing-risk model (e.g., Han and Hausman, 1990), which allows parametric dependence of those probabilities on firms' and drugs' characteristics.

Second, we have excluded the drugs developed by large pharmaceutical firms with a market valuation above the 95th percentile of the firm size distribution. Insofar as these firms develop different types of drugs than others, our approach fails to capture those drugs. We could consider large firms separately and keep track of announcements about acquired drugs to relax this assumption.

Third, a competitor's announcements can affect a firm's drug valuation. For instance, if two firms are developing competitive drugs, the impact of an announcement by one firm on the market value of other firms would inform about the expected effect of competition. We can adapt our approach by including "competitive announcements" in our framework.

In this paper, we have focused only on single announcements and ignored days with more than one announcement. The difficulty is determining the change in a firm's market value in response to each announcement when there are multiple announcements on any given day. One way to progress would be to rely on additional functional form assumptions. For instance, we could use a log of the number of announcements for each stage as regressors to determine the average value of each announcement.

References

Aalen, Odd, "Nonparametric Inference in Connection with Multiple Decrement Models," *Scandinavian Journal of Statistics*, 1976, *3* (1), 15–27.

Adams, Christopher P. and Van V. Brantner, "Estimating the Cost of New Drug Development: Is It Really $802 Million?," *Health Affairs*, 2006, *25* (2), 420–428.

Akcigit, Ufuk and William R. Kerr, "Growth through Heterogeneous Innovations," *Journal of Political Economy*, 2018, *126* (4), 1374–1443.

Arora, Ashish, Alfonso Gambardella, Laura Magazzini, and Fabio Pammolli, "A Breath of Fresh Air? Firm Type, Scale, Scope, and Selection Effects in Drug Development," *Management Science*, 2009, *55* (10), 1638–1653.

Arrow, Kenneth J., "Economic Welfare and the Allocation of Resources for Invention," in Charles K. Rowley, ed., *Readings in Industrial Economics: Volume Two: Private Enterprise and State Intervention*, London: Macmillan Education UK, 1972, pp. 219–236.

_ , "Innovation in Large and Small Firms," in R. R. Nelson, ed., *Entrepreneurship*, Lexington Books, 1983.

Athey, Susan, Michael Kremer, Christopher Snyder, and Alex Tabarrok, "In the Race for a Coronavirus Vaccine, We Must Go Big. Really, Really Big," *New York Times* May 4 2020.

Austin, David H., "An Event-Study Approach to Measuring Innovative Output: The Case of Biotechnology," *American Economic Review P&P*, 1993, *83* (2), 253–258.

Ausubel, Lawrence M., "An Efficient Ascending-Bid Auction for Multiple Objects," *American Economic Review*, 2004, *94* (5), 1452–1475.

Barberio, Joseph, Jacob Becraft, Zied Ben Chaouch, Dimitris Bertsimas, Tasuku Kitada, Michael L. Li, Andrew W. Lo, Kevin Shi, and Qingyang Xu, "Accelerating Vaccine Innovation for Emerging Infectious Diseases via Parallel Discovery," *NBER Working Paper 30126*, 2022.

Bhattacharya, Vivek, "An Empirical Model of R&D Procurement Contests: An Analysis of the DOD SBIR Program," *Econometrica*, 2021, *89*, 2189–2224.

Boldrin, Michele and David K. Levine, *Against Intellectural Monopoly*, Cambridge University Press, 2008.

Boller, Lysle and Fiona M. Scott Morton, "Testing the Theory of Common Stock Ownership," *NBER Working Paper 27515*, 2020.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, *The Econometrics of Financial Markets*, Princeton University Press, 1997.

Card, David and Alan B. Krueger, *Myth and Measurement: The New Economics of Minimum Wage*, Princeton University Press, 1995.

Chan, Louis K. C., Josef Lakonishok, and Theodore Sougiannis, "The Stock Market Valuation of Research and Development Expenditure," *Journal of Finance*, 2001, *56* (6), 2431–2456.

Che, Yeon-Koo and Ian Gale, "Optimal Design of Research Contests," *American Economic Review*, 2003, *93* (3), 646–671.

Congressional Budget Office, "Research and Development in the Pharmaceutical Industry," Technical Report, Congressional Budget Office 2021.

DiMasi, Joseph A., "Causes of Clinical Failures Vary Widely by Therapeutic, Phase of Study.," *Impact Report*, 2013, *15* (5).

_ , **Henry G. Grabowski, and Ronald W. Hansen**, "Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs," *Journal of Health Economics*, 2016, *47*, 20–33.

Dube, Arindrajit, Ethan Kaplan, and Suresh Naidu, "Coups, Corporations, and Classified Information," *Quarterly Journal of Economics*, 2011, *126*, 1375–1409.

Eckbo, B. Espen, "Horizontal Mergers, Collusion, and Stockholder Wealth," *Journal of Financial Economics*, 1983, *11* (1-4), 241–273.

Ely, Jeffrey C., George Georgiadis, Sina Khorasani, and Luis Rayo, "Optimal Feedback in Contest," *Review of Economic Studies*, Forthcoming.

Fama, Eugene F., "The Behavior of Stock-market Prices," *Journal of Business*, 1965, *38* (1), 34–105.

_ , **Lawrence Fisher, Michael C. Jensen, and Richard Roll**, "The Adjustment of Stock Prices to New Information," *International Economic Review*, 1969, *10* (1), 1–21.

Feldman, Robin C., David A. Hyman, W. Nichoolson Price II, and Mark J. Ratain, "Negative Innovation: When Patents Are Bad for Patients," *Nature Biotechnology*, 2021, *39*, 914–916.

Finkelstein, Amy, "Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry," *Quarterly Journal of Economics*, 2004, *119* (2), 527–564.

Fisman, Raymond, "Estimating the Value of Political Connections," *American Economic Review*, 2001, *91* (4), 1095–1102.

Food and Drug Administration, "Compliance Program Guidance Manual 7346.832," https://www.fda.gov/media/71498/download 5 2010.

Giliches, Zvi, "Market Value, R&D, and Patents," *Economics Letters*, 1981, *7* (2), 183–187.

Grennan, Matthew and Robert J. Town, "Regulating Innovation with Uncertain Quality: Information, Risk, and Access in Medical Devices," *American Economic Review*, 2020, *110* (1), 120–161.

Gupta, Charu N., "One Product, Many Patents: Imperfect Intellectual Property Rights in the Pharmaceutical Industry," *Working Paper UCLA Anderson*, 2021.

Halac, Marina, Navin Kartik, and Qingmin Liu, "Contests for Experimentation," *Journal of Political Economy*, 2017, *125* (5), 1523–1569.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg, "Market Value and Patent Citations," *RAND Journal of Economics*, 2005, *36* (1), 16–38.

Han, Aaron and Jerry A. Hausman, "Flexible Parametric Estimation of Duration and Competing Risk Models," *Journal of Applied Econometrics*, 1990, *5* (1), 1–28.

Hay, Michael, David W. Thomas, John L. Craighead, Celia Economides, and Jesse Rosenthal, "Clinical Development Success Rates for Investigational Drugs," *Nature Biotechnology*, 2014, *32* (1), 40–51.

Hendren, Nathaniel, "Measuring Economic Efficiency Using Inverse-Optimum Weights," *Journal of Public Economics*, 2020, *187*, 104198.

Herrling, Paul, "Patent Sense," *Nature*, 2007, *449* (174-175).

Holmström, Bengt, "Agency Costs and Innovation," *Journal of Economic Behavior and Organization*, 1989, *12* (3), 305–327.

Howell, Sabrina T., "Financing Innovation: Evidence from R&D Grants," *American Economic Review*, 2017, *107* (4), 1136–1164.

Jacobo-Rubio, Roben, John L. Turner, and Jonathan W. Williams, "The Distribution of Surplus in the US Pharmaceutical Industry: Evidence from Paragraph iv Patent-Litigation Decisions," *Journal of Law and Economics*, 2020, *63*, 203–238.

Känzig, Diego R., "The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements," *American Economic Review*, 2021, *111* (4), 1092–1125.

Khmelnitskaya, Ekaterina, "Competition and Attrition in Drug Development," *Working Paper*, 2022.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological Innovation, Resource Allocation, and Growth," *Quarterly Journal of Economics*, 2017, *132* (2), 665–712.

Kremer, Michael, "Patent Buyouts: A Mechanism for Encouraging Innovation," *Quarterly Journal of Economics*, 1998, *113* (4), 1137–1167.

_ **and Rachel Glennerster**, *Strong Medicine: Creating Incentives for Pharmaceutical Research on Neglected Diseases*, Princeton University Press, 2004.

_ **, Jonathan D. Levin, and Christopher M. Snyder**, "Advance Market Commitments: Insights from Theory and Experience," *American Economic Review P&P*, 2020, *110*, 269–273.

_ **,** _ **, and** _ , "Designing Advance Market Commitments for New Vaccines," *Management Science*, 2022, *68* (7), 4786–4814.

Krieger, Joshua, Danielle Li, and Dimitris Papanikolaou, "Missing Novelty in Drug Development," *The Review of Financial Studies*, 2022, *35* (2), 636–679.

Langer, Ashley and Derek Lemoine, "What Were the Odds? Estimating the Market's Probability of Uncertain Events," *NBER Working Paper 28265*, 2020.

Marshall, Alfred, *Principles of Political Economy*, New York: Macmillan, 1890.

McKeon, John, Felix Poege, and Tim Simcoe, "Valuing Pharmaceutical Patent Thickets," *Working Paper*, 2022.

MCockburn, Iain and Rebecca M. Henderson, "Scale and Scope in Drug Development: Unpacking the Advantages of Size in Pharmaceutical Research," *Journal of Health Economics*, 2001, *20* (6), 1033–1057.

Milgrom, Paul R., "Rational Expectations, Information Acquisition, and Competitive Bidding," *Econometrica*, 1981, *49* (4), 380–391.

Moser, Petra, "Patents and Innovation: Evidence from Economic History," *Journal of Economic Perspectives*, 2013, *27* (1), 23–44.

Mullard, Asher, "Parsing Clinical Success Rates," *Nature Reviews Drug discovery*, jun 2016, *15* (7), 447.

Munari, Federico and Raffaele Oriani, eds, *The Economic Valuation of Patents: Methods and Applications*, Edward Elgar, 2011.

Munos, B, "Lessons from 60 Years of Pharmaceutical Innovation," *Nature Reviews Drug Discovery*, 2009, *8*, 959–968.

Ng, Rick, *Drugs: From Discovery to Approval*, Wiley, 2015.

Pakes, Ariel, "On Patents, R&D, and the Stock Market Rate of Return," *Journal of Political Economy*, 1985, *93*, 390–409.

Pammolli, F., L Magazzini, and M. Riccaboni, "The Productivity Crisis in Pharmaceutical R&D," *Nature Reviews Drug Discovery*, 2011, *10*, 428–438.

Paul, Steven M., Daniel S. Mytelka, Christopher T. Dunwiddie, Charles C. Persinger, Bernard H. Munos, Stacy R. Lindborg, and Aaron L. Schacht, "How to Improve R&D Productivity: The Pharmaceutical Industry's Grand Challenge," *Nature Reviews Drug Discovery*, 2010, *9*, 203–214.

Samuelson, Paul, "Proof That Properly Anticipated Prices Fluctuate Randomly," *Industrial Management Review*, 1965, *6*, 41–49.

Scannell, Jack W., Alex Blanckley, Helen Boldon, and Brian Warrington, "Diagnosing the Decline in Pharmaceutical R&D Efficiency," *Nature Reviews Drug Discovery*, 2012, *11*, 191–200.

Sertkaya, Aylin, Anna Birkenbach, Ayesha Berlind, and John Eyraud, "Examination of Clinical Trial Costss and Barriers for Drug Development," Technical Report, U.S. Department of Health and Human Services 2014.

_ , Hui-Hsing Wong, Amber Jessup, and Trinidad Beleche, "Key Cost Drivers of Pharmaceutical Clinical Trials in the United States," *Clinical Trials*, 2016, *13* (2), Epub.

Shin, H. S., "Disclosures and Asset Returns," *Econometrica*, 2003, *71*, 105–133.

U.S. House of Representatives, "Drug Pricing Investigation," Technical Report, Committee on Oversight and Reform 2021.

Van Norman, Gail A., "Drugs and Devices: Comparison of European and U.S. Approval Processes," *JACC: Basic to Translational Science*, 2016, *1* (5), 399–412.

Wallsten, Scott J., "The Effects of Government-Industry R&D Programs on Private R&D: The Case of the Small Business Innovation Research," *RAND Journal of Economics*, 2000, *31* (1), 82–100.

Whinston, Michael D. and Scott C. Collins, "Entry and Competitive Structure in Deregulated Airline Markets: An Event Study Analysis of People Express," *RAND Journal of Economics*, 1992, *23* (4), 445–462.

Wong, Chi Heem, Kien Wei Siah, and Andrew W. Lo, "Estimation of Clinical Trial Success Rates and Related Parameters," *Biostatistics*, 2019, *20* (2), 273–286.

Wouters, Olivier J, McKee, Martin, and Luyten, Jeroen, "Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018," *Journal of American Medical Association*, 2020, *323* (9), 844–853.

Appendix

Table A.1: **Transition Probabilities**

Sample	Stages	Probability of Reaching a Stage	
		Marginal	Conditional
Middle 90%	Phase I Clinical Trials	0.501	0.501
	Phase II Clinical Trials	0.290	0.579
	Phase III Clinical Trials	0.137	0.471
	FDA Application	0.092	0.674
	FDA Approval	0.081	0.878
Bottom 95%	Phase I Clinical Trials	0.501	0.501
	Phase II Clinical Trials	0.292	0.583
	Phase III Clinical Trials	0.140	0.479
	FDA Application	0.096	0.685
	FDA Approval	0.085	0.884

Note. Transition probabilities of reaching three stages for all drug candidates in our sample. The column labeled *Marginal* denotes the shares (as a percent) of all the initiated development projects (where a development project refers to a specific firm-drug-disease), and the column labeled *Conditional* denotes the shares (as a percent) of the development projects that made it to the previous stage. Middle 90% refers to the drugs developed by firms whose real market capitalization is between 5% and 95% of the entire sample; and Bottom 95% refers to the drugs developed by firms whose real market capitalization is below 95% of the entire sample.