Valuing Pharmaceutical Drug Innovations*

Gaurab Aryal† Federico Ciliberto‡

Leland E. Farmer§ Ekaterina Khmelnitskaya¶

April 12, 2024

Abstract

We propose a methodology to estimate the market value of pharmaceutical drugs. Our approach combines an event study with a model of discounted cash flows and uses stock market responses to drug development announcements to infer the values. We estimate that, on average, a successful drug is valued at $1.62 billion, and its value at the discovery stage is $64.3 million, with substantial heterogeneity across major diseases. Leveraging these estimates, we also determine the average drug development costs at various stages. Furthermore, we explore applying our estimates to design policies that support drug development through drug buyouts and cost-sharing agreements.

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, Christopher Adams, Derek Lemoine, Maxim Engers, Tim Simcoe and the participants and discussants at the UVA, SEA 2021, MaCCI 2022, ASHEcon 2022, EARIE 2022, SciencesPo, Indiana University, BU TPRI, NASM-ES 2023, MaCCI-EPoS: Frontiers in Empirical IO 2023 and JHU. 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

Measuring the value of pharmaceutical drug innovations is essential in determining the appropriate level of incentives for innovators, assessing the effectiveness of current policies, and guiding decision-making in the industry. Specific estimates of drug profitability are of particular interest, as they frequently influence policy choices, such as the Medicare Drug Price Negotiation program of the *Inflation Reduction Act 2022*. However, research on drug valuation is limited and challenging due to a lack of drug-specific R&D expenditure data and the complex, long-term nature of drug development.

We develop a systematic approach to valuing drugs by combining the *event study* approach with a model of *discounted cash flows*. Using our approach, we identify the value of drugs using stock market reactions to drug development announcements for a large sample of drugs under development. While event studies are widely used, we are the first to combine them with a model of discounted cash flows, which captures risk associated with drug R&D, to propose a practical way to estimate the market value of drugs using public data.[1]

Our work seeks to produce reliable estimates of the market value of drugs in order to enable informed decision-making. Ultimately, and as shown below, we hope this will lead to more efficient and effective drug development processes that optimize societal benefits while ensuring sufficient incentives for firms to invest in R&D (Paul et al., 2010).

To estimate the values of drugs, we propose a two-step methodology and apply it to the Cortellis dataset from Clarivate, which covers more than 70,000 drug candidates. We focus on drugs developed by publicly traded companies and supplement the data with daily stock prices and market capitalization data from the Center for Research in Security Prices. The first step is to determine the change in the market value of the firm following a drug development announcement. We use the *unrestricted market model* to estimate the *cumulative abnormal returns* (CAR) associated with each announcement about discoveries, FDA appli-

[1] For some applications of the event study method, see Whinston and Collins (1992); Hwang (2013); Kogan et al. (2017); Langer and Lemoine (2020); Jacobo-Rubio et al. (2020) and Singh et al. (2022).

cations, approvals, and discontinuations at different stages. The product of the (predicted) CAR and market capitalization equals the change in the firm's market value.

The second step relies on the weak form of the *efficient markets hypothesis* (Fama, 1965; Samuelson, 1965; Fama et al., 1969), which implies that the change in the firm's market value following an announcement equals the change in the market value of the drug in question.

In particular, when a firm announces FDA approval of a drug under review, the only relevant change in the drug's status moments before this announcement is the resolution of the uncertainty about the FDA's decision. Thus, the change in the firm's market value upon the approval announcement equals the drug's value adjusted by the risk of failure. We consider two cases: one with a homogeneous probability of success and another with heterogeneous probabilities across 11 major indications (e.g., cancers, rare diseases, immune disorders) and use them to recover the value of the drug at approval.

Next, we discount and adjust the estimated value at approval to determine the drug's value at the discovery stage. To account for the random *time to success* from discovery to approval, we use the competing risk model (Aalen, 1976) to estimate the distribution of time to success and use it to determine the expected discount rate. We need additional assumptions to implement the discounting, such as constant per-period profits for the drug post-approval. We discuss these assumptions in our empirical analysis.

We find that under homogeneous probabilities of success, the average expected market value of an approved drug is \$1.62 billion. We compare this estimate with discounted sales revenue to verify that it is sensible. The Cortellis Competitive Intelligence database includes information on yearly drug-level total sales for a subset of drugs. The average ten-year and fifteen-year discounted revenues for those drugs are \$1.40 and \$1.99 billion per drug, respectively, suggesting that our approach is reasonable.

We find substantial differences in drug values when we consider heterogeneous probabilities of success across several major diseases and estimate their values separately. For instance, on average, a successful drug for cancers, rare diseases, and endocrine diseases is

valued at $2.66, $7.83, and $15.30 billion, respectively.

Next, we use the value estimates to determine the average cost of drug development. While this exercise is distinct from the first two steps of our valuation approach and relies on the structure of our discounted cash flow model, it provides crucial insights into the financial implications of drug development and investment required to bring a drug to market.

Under the assumption of homogeneous probabilities of success, we estimate the average cost of developing a drug to be $58.51 million. This high cost suggests that drug development is risky compared to the expected value of $63.37 million at the discovery stage. Furthermore, by leveraging discontinuation announcements at various stages, something we did not need to use until now, we estimate the expected cost of clinical trials, where the expectation is taken at the discovery stage, to be approximately $12.43 million. Breaking down the clinical trial costs by phase, we find that the expected cost of Phase I clinical trials is $0.62 million, which increases significantly to $30.48 million and $41.09 million for Phases II and III, respectively. These estimates suggest escalating costs with each successive phase.

Our cost estimates allow us to assess the average cost of bringing new drugs to the market and inform regulations such as price interventions (U.S. House of Representatives, 2021).[2] The best-known cost estimates either rely on confidential surveys and a sample of a few firms (DiMasi et al., 2016) or estimate the accounting cost of a single trial (Sertkaya et al., 2016). Our work complements this literature by providing economic, risk-adjusted cost estimates based on a representative sample of drugs, leveraging the valuation approach we developed.

Finally, having explained our methodology and presented the estimation results, we explore how these estimates can inform policymakers in designing systems to support drug development and boost declining productivity (Munos, 2009; Pammolli et al., 2011; Scannell et al., 2012). To this end, we consider drug buyout schemes for FDA-approved drugs and drugs at the discovery stage and cost-sharing agreements between the government and firms.

Under a drug buyout program, much like Kremer (1998)'s patent buyouts, we envision

[2]For more on the topic from the Congressional Budget Office, see here and here. Also, see Dubois and Kyle (2016) for estimates of the costs per statistical life saved from cancer drugs.

that the government buys a drug's manufacturing rights and makes them publicly accessible. When buying a successful drug, the government must address Lucas' critique: the market will adjust the drug's valuation based on the expected government payment. We propose a solution under which the government commits to implementing the scheme with a small but exogenous probability. Conditional on buying, the government pays the optimal price (Myerson, 1981) determined using the historical distribution of values. Under this approach, the government pays the optimal price known in advance with a specified probability, and the market adapts to the new expected value. This solution addresses Lucas' critique and allows us to identify the values of new drugs and update the historical distribution. The solution also highlights an inherent tradeoff: the restriction to small probabilities of drug buyouts limits the effectiveness of drug buyouts as incentives for drug R&D. Our estimates are crucial in determining the expected cost of this program and enabling policymakers to make informed decisions when implementing such initiatives.

When buying out a drug at the early–discovery–stage, the government faces different challenges. If the government puts the drug in the public domain, it must address who will pay development costs. To incentivize R&D, the government can supplement the buyout program using advanced market commitment (Kremer and Glennerster, 2004; Kremer et al., 2020, 2022). Under this scheme, the government commits to buy a certain number of drugs at a pre-fixed price after FDA approval, acting as a prize for early developers. Our valuation exercise helps inform the level of commitment needed to incentivize R&D.

In addition to the drug buyout program, our cost estimates can be useful for designing cost-sharing agreements, where the government provides funding to cover development costs in exchange for the firm selling its approved drugs to the government at a pre-fixed price. The government could pay $58.51 million of R&D costs at discovery, the expected cost of drug development at this stage, in return for the firm committing to provide a certain number of units for free at a pre-negotiated price. Alternatively, the government can adopt a development stage contingent cost-sharing agreement, offering to pay $638.75 million towards FDA

review and application costs upon successful completion of clinical trials. This sequential payment can lower the program costs. If the drug is approved, the firm will deliver a specific number of units for free based on government investment and the pre-negotiated price.

Our paper contributes to several strands of literature. First, it adds to the literature that estimates the value of innovations (e.g., Giliches, 1981; Pakes, 1986; Austin, 1993; Chan et al., 2001; Hall et al., 2005; Munari and Oriani, eds, 2011; Azoulay et al., 2018; McKeon et al., 2022). Unlike the extant literature on patents, we focus on products under development, particularly novel drugs, and propose a practical strategy to identify their commercial values. We leverage the institutional features of the pharmaceutical industry that give rise to certain types of public announcements (e.g., FDA approval), allowing us to separately estimate the values of novel drugs and the cost of drug development. Thus, our valuation method can be applied to other regulated industries, such as medical devices, agrochemicals, agricultural biotechnology, and green technology sectors.

A recent study closely related to ours is Singh et al. (2022), which uses an event study methodology to evaluate the stock market's reaction to clinical trial announcements in the pharmaceutical industry. They allow for rich heterogeneity across drugs and correlate abnormal returns with clinical trial results, providing an important insight into characteristics that increase the stock market's valuation of an experimental drug. However, they only compare the values (e.g., a drug with a better safety profile is commercially valuable) and do not estimate their magnitudes, which is one of our main objectives. In addition, our approach can be generalized by including more types of announcements and disease heterogeneity.

Further, by providing estimates of R&D costs for different stages, this paper contributes to the small literature that evaluates the cost of bringing new drugs to the market (DiMasi et al., 2016; Sertkaya et al., 2016; Wouters et al., 2020; Congressional Budget Office, 2021).

Lastly, our paper complements the literature that seeks methods to improve market efficiency without hurting firms' R&D incentives, such as patent buyouts Kremer (1998) and transferable patents Dubois et al. (2022). Our drug buyout approach is closer to Kremer

(1998), but our method of estimating drug values requires weaker assumptions and has practical advantages over Kremer (1998)'s proposed auctions. We rely on stock prices that aggregate information dispersed among a large pool of self-interested investors (Milgrom, 1981), providing a more accurate valuation than an auction. Furthermore, as drugs are often covered by multiple patents (Gupta, 2021; McKeon et al., 2022), valuing a drug is better than combining the values of patents.

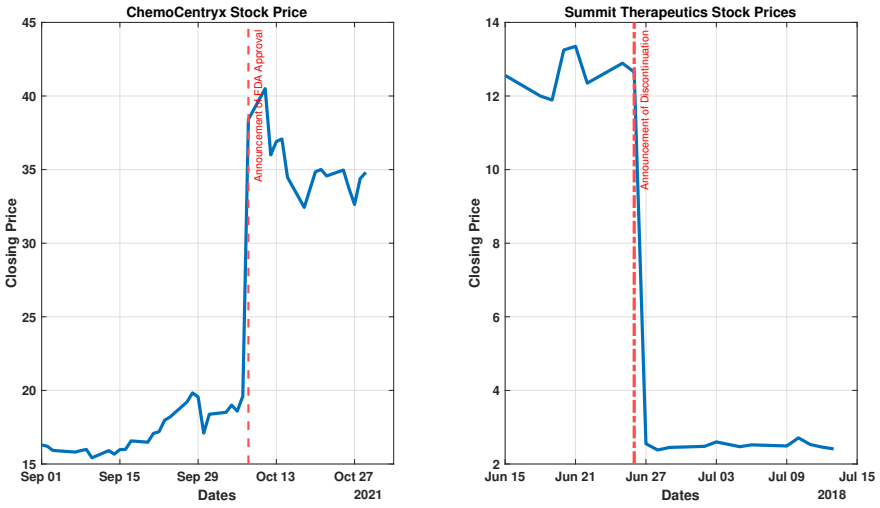
2 Examples of Drug Development Announcements

To illustrate the impact of FDA approvals on pharmaceutical companies, we consider the case of ChemoCentryx, which was developing ANCA-associated vasculitis therapy. On October 8, 2021, the company announced that the FDA approved its application. Figure 1-(a) shows that the market responded positively to this news, and its stock price increased sharply, resulting in a rise in the company's market value.[3] Because it was the only relevant news on that day, we can attribute this increase to the abnormal returns associated with the announcement. These abnormal returns reflect the market's revised expectation of the firm's future earnings from selling the drug now that it has received FDA approval.

Next, Summit Pharmaceuticals, a company focused on developing treatments for rare genetic disorders, was working on a drug called ezutromid for Duchenne muscular dystrophy. On June 27, 2018, Summit announced that it had discontinued the development of ezutromid following the results of a Phase II clinical trial conducted the previous day. As shown in Figure 1-(b), the market responded negatively to this news, reflected in the company's stock price.[4] The drop in Summit Pharmaceuticals' market value can be attributed to the loss of expected future revenue from the drug and the elimination of future development costs that would have been incurred to bring the drug to market. These examples illustrate that drug development announcements can affect pharmaceutical companies' market values.

[3] Click here for Chemocentryx's (archived) announcement document. Last accessed November 14, 2023.
[4] Click here for Summit's (archived) announcement document. Last accessed November 14, 2023.

Figure 1: **Examples of Drug Development Announcements**



Note: Panels (a) and (b) display the time series of stock prices for ChemoCentryx pharmaceutical and Summit therapeutics around announcement dates, respectively. On October 8, 2021, ChemoCentryx announced FDA approval for its vasculitis drug. On June 27, 2018, Summit announced the discontinuation of its Duchenne muscular dystrophy drug after the Phase II trial.

Next, in Table 1, we present examples of five successful drugs, including information about the dates associated with three key development milestones: discovery, FDA application, and FDA approval. These milestones inform us of the time it takes for a drug to reach the market from its initial discovery. In some cases, we also have sales data available, which allows us to evaluate the accuracy of our estimates of the drugs' values. In the rest of the paper, we first summarize the institutional details and the data used in our analysis. Then, we formalize the idea that a firm's change in market value within a *tight* window around these milestone announcements can be used to identify the value and cost of the drug under development.

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, as illustrated in Figure 2. The first stage is the pre-clinical stage, which includes

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.

creating a new molecule (or a system of molecules) and testing it in the laboratory using *in vitro* and *in vivo* methods. This stage is referred to as the discovery stage.

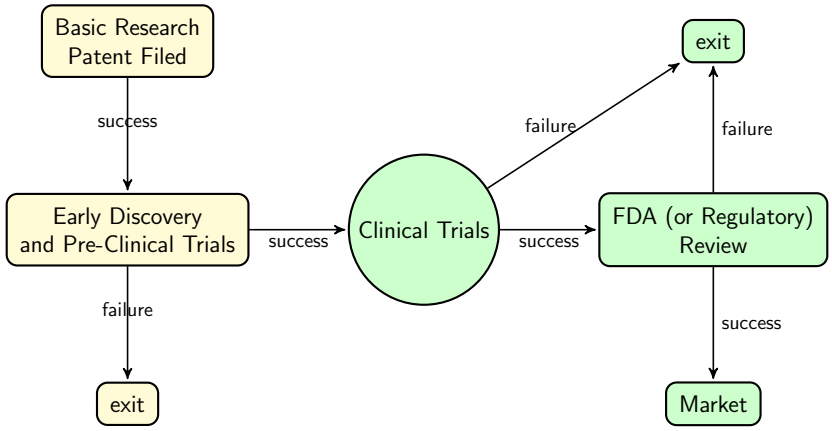
If the discovery stage is successful, the firm can start testing the drug candidate in humans through three stages of clinical trials. Phase I involves screening the drug for possible toxicity using a small sample of healthy subjects. In Phase II, firms test the drug's efficacy on a larger sample of individuals with the targeted diseases. Finally, in Phase III, the firm conducts double-blinded tests to assess the drug's effectiveness on a large sample of patients.

After successful completion of clinical trials, the firm can submit an FDA review application. The FDA has a group of internal and external experts who review the results from clinical trials and the applicant firm's manufacturing capacity before deciding 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. Table 2 presents transition probabilities based on the frequency distributions. For simplicity of presentation, we assume that the transition probabilities are homogeneous across firms and diseases. Later, we also consider heterogeneity across the size of the firms and the diseases being targeted (Tables C.1 and C.2).

Table 2 shows that, on average, only about 11% of drug candidates are successful uncon-

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.

ditionally. In contrast, 89% of drugs that apply for FDA approval are successful.[5]

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 that a drug reaches Phase II clinical trials, given that it has reached Phase I, is the ratio of drugs that have reached Phase II to those that have reached Phase I.

Three regulatory paradigms govern firms' announcements about the success and failure of drug candidates throughout the development process.[6] First, the Security and Exchange Commission (SEC) requires public companies to disclose all material information to investors via the annual 10-K, quarterly 10-Q, and current 8-K forms, with *Regulation Fair Disclosure* (2000) mandating timely disclosure of all material information. Second, the FDA controls what firms can announce about their drugs during development, requiring registration of clinical trials within 21 days of enrolling the first subject under the *Food and Drug Ad-*

[5]A drug can be developed concurrently for multiple indications (diseases) by more than one firm. We define an observation to be at the individual firm-drug-indication level. Our estimate of a 10.8% chance of success 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%. These averages do not account for differences in pre-clinical evidence about the mechanism of action or the firm's expertise in an indication.

[6]Although most announcements in our sample are made in the U.S., firms also market drugs elsewhere, such as in the E.U. and Canada. We include announcements from all "western" countries because they have similar rules governing drug development and announcements (see Ng, 2015, Chapter 7).

Table 2: **Transition Probabilities**

Stages	**Probability of Reaching a Stage** Marginal	Conditional
Phase I Clinical Trials	0.512	0.512
Phase II Clinical Trials	0.319	0.624
Phase III Clinical Trials	0.167	0.524
FDA Application	0.121	0.723
FDA Approval	0.108	0.890

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.

ministration Modernization Act (1997) and disclosure of information about clinical trials and FDA application processes whenever relevant, ensuring announcements are not *materially misleading*. Third, the *Sarbanes-Oxley Act* (2002) allows the SEC to monitor firms' announcements about their FDA review process, with the FDA referring cases of false or misleading statements to the investing public to the SEC since 2004.

These regulations incentivize firms to inform the market and the general public correctly and promptly. However, companies retain some discretion in determining what is considered "material" and "not misleading." This ambiguity is more pronounced when the results from clinical trials are *small* or when large companies are developing multiple drugs simultaneously. In such cases, firms may delay or bundle negative announcements with positive ones to soften the market reaction.

Nevertheless, we remain agnostic about the incentives for bundling, as our focus is only on days with only one announcement by a firm. Furthermore, we consider only major announcements that we are reasonably certain to be material. Specifically, we use information about drug discovery, a firm's application for FDA authorization to market the drug, the FDA's decision, and discontinuations at different stages. Thus, we exclude the dates corresponding to the start of various phases of clinical trials.

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

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.

Table 3 shows the number of different announcements in our sample. For almost 30% of the dates, firms make more than one announcement. Throughout the paper, we only use single announcement dates. Table 4 shows the distribution of announcements per day. The median number of announcements is one but with a long right tail.

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.

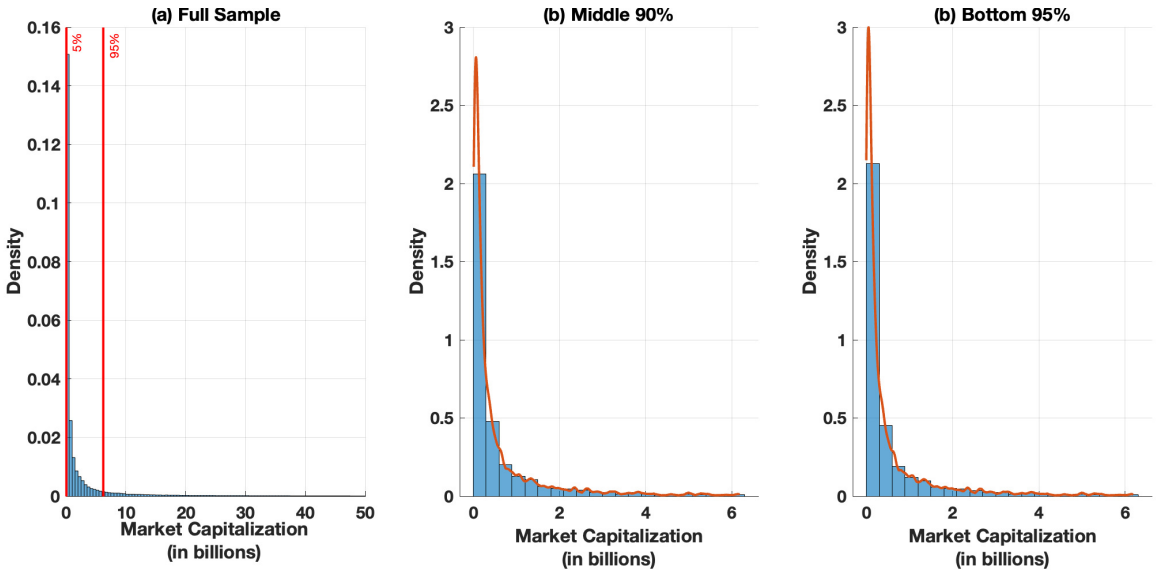
In our analysis, we do not include announcements about the start of a specific phase of clinical trials. Given that results from previous phases and announcements of plans to initiate a new phase of clinical trials are likely public, so the initiation announcement is less likely to be a "surprise" announcement. By focusing only on the announcements for which we are certain about their prominence, namely drug discovery, FDA application submission, FDA approval, and discontinuations at different stages, we maintain the validity of our conclusions. Moreover, as we demonstrate below, the nature of these specific announcements implies a straightforward relationship between the associated abnormal returns and the values and the costs of drug development projects under fewer assumptions and lower data requirements.

3.3 Market Value of a Firm

We obtain daily returns for all biomedical and pharmaceutical companies publicly listed on U.S. stock exchanges from the Center for Research in Security Prices (CRSP) database. We match firms' names between the two datasets to combine the stock market data with the drug development information from Cortellis.[7]

[7]Firms may rename, merge with, or be acquired by another company. To ensure we consistently track unique firms throughout these changes, we use CRSP-generated permanent ID numbers. We consider a match if any name the firm has had in its past (i.e., any name associated with the CRSP-generated permanent ID) matches the name in the Cortellis data, resulting in an unbalanced panel structure for our merged dataset.

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).

We define the "market return" as the return on the CRSP value-weighted portfolio, which includes dividends. As a measure of a firm's size, we use its market capitalization. To determine each firm's size, we calculate its median real market capitalization across years, using the December 2020 consumer price index to deflate all dollar amounts.

Figure 3 shows histograms of firm sizes for different subsamples. In the full sample of firms (Figure 3-(a)), market capitalization exhibits a long right tail. The vertical lines denote the 5th ($11.6 million) and 95th ($6.2 billion) percentiles, respectively.

Figures 3-(b) and (c) display the density of market capitalization for firms with market capitalization between the 5th and 95th percentiles of all firms (i.e., the "Middle 90%" sample) and for those below the 95th percentile (i.e., the "Bottom 95%" sample), respectively.

Heterogeneity between small and large firms is crucial for our approach. First, large firms may have more expertise and resources for conducting clinical development, potentially increasing their chances of success and leading to heterogeneity in transition probabilities. Second, for the same reasons, large firms may conduct their clinical trials faster, leading to

differences in the time investors wait to realize their returns. Third, large firms may select and develop certain types of drugs (Cockburn 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 percentage change in the price of a single stock, depending on the firm's size.

Therefore, including large firms in our sample with all other firms may affect our estimates. In the empirical exercises below, we focus on the middle 90% and bottom 95% samples, although we also present values estimated using the full sample for completeness.

4 Drug Valuation

4.1 Methodology

Valuation at Approval. When a firm announces that the FDA has approved a drug under review, the firm's value should immediately increase following the announcement. The size of this increase should equal the change in the *expected* profits from selling the drug. Importantly, in this context, the only relevant change in the drug's status just before the announcement is the resolution of uncertainty about the FDA's decision. Therefore, if we can estimate the level of uncertainty before the announcement and the change in the firm's market value after the announcement, we can recover the value of an approved drug.

To formalize this intuition, we introduce a few notations and define the variables. Let π be the expected (average) yearly profit from selling the drug *after* FDA approval. Additionally, let $S_k \in \{0, 1\}$ be a binary variable that is equal to one if stage k (e.g., FDA approval, application) announcement is positive and zero otherwise. So, $S_{\texttt{appr}} = 1$ if the FDA approved the drug. Then, for the discount factor $\delta \in [0, 1]$, the expected value of a drug at approval, $\mathbb{E}(V|S_{\texttt{appr}} = 1)$, is the present discounted profits, i.e., $\mathbb{E}(V|S_{\texttt{appr}} = 1) := \frac{\pi}{1-\delta}$. Let $p_{k'|k}$ denote the transition probability from stage k to stage k'.

Just before the announcement, the market expects the value of the drug to be $\mathbb{E}(V|S_{\texttt{appr}} =$

$1) \times p_{\texttt{appr}|\texttt{appl}}$, where $p_{\texttt{appr}|\texttt{appl}}$ is the probability that a drug is approved given that it has been submitted for FDA application.

Immediately after the approval announcement, the uncertainty about the FDA's decision is resolved, and the market expects the value to be $\mathbb{E}(V|S_{\texttt{appr}} = 1)$. When the only "news" that pertains to the firm is the approval, the change in the firm's market value, $\mathbb{E}\left(\texttt{CAR}_{\texttt{appr}}\right) \times \texttt{mktcap}$, is equal to the (unobserved) change in the market value of the drug, i.e.,

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

We estimate the $\mathbb{E}\left(\texttt{CAR}_{\texttt{appr}}\right)$ and the transition probabilities, $p_{\texttt{appr}|\texttt{appl}}$, and because we observe the $\texttt{mktcap}$, we can determine the left-hand-side of (1).

By equating the change in the firm's market value to the change in the market value of the drug, we can identify the value of a drug at approval. In our context, it is reasonable to assume that by the time the FDA announces its approval, the payoff-relevant information about the drug has already been disclosed to the market. So, the only change is in the uncertainty about FDA's decision, and the approval announcement resolves such uncertainty.

Valuation at Discovery. Next, determining the value of a drug at the start of the development process is crucial for understanding the potential benefits and risks involved. Besides providing information about the benefits, by comparing the expected value at discovery with the cost at the discovery stage (see Section 5), we can better understand the underlying risk facing the firms at the start.

To determine the value of a drug at discovery, $\mathbb{E}(V|S_{\texttt{disc}} = 1)$, we have to discount and risk-adjust the value at approval, $\mathbb{E}(V|S_{\texttt{appr}} = 1)$. The discounting arises because a drug generates profits only after FDA approval, a few years later than when the drug is discovered. The *time to success* is a priori uncertain, so we use the competing risk model to estimate its distribution (see Section 4.2) and use it to determine the expected discount rate.

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. Then, $\mathbb{E}(\delta^{\tau_{k\rightarrow}}) = \sum_{\tau \geq 0} \delta^\tau \times \mathbb{P}_{\texttt{appr}}(\tau|S_k = 1)$ is the expected discount rate when the drug is at stage k.

The risk adjustment arises because, at Discovery, there is additional uncertainty about the success, captured by the transition probability $p_{\texttt{appr}|\texttt{disc}}$.

Finally, under an additional assumption that the expected per-period profit is constant over time, which we did not need to estimate the value at approval, we can express the expected net present discounted value of a drug as the product of (i) the present discounted profit from the time the drug reaches the market onwards, (ii) the probability that the drug will transition from discovery to the market, and (iii) the stochastic discount rate, as follows:

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

where the first equality uses the definition of the expected value of a drug at discovery, the second uses the fact that the profit π is constant over time, and the last uses the definition of expected value at approval and expected discount rate. Once we obtain $\mathbb{E}(V|S_{\texttt{appr}} = 1)$ from (1), $\mathbb{E}(\delta^{\tau_{\texttt{disc}\rightarrow}})$ (see Section 4.2); and the transition probabilities, we can determine the value of a drug at discovery.

We estimate the market value of drugs at the discovery under two assumptions. First, as discussed earlier, the per-period profit π is constant over time. We can relax this assumption and allow the per-period profit to decline at a known parametric rate without affecting the final estimate. Additionally, we assume that development announcements affect only the transition probabilities and time to success, not the commercial value of the drug. This

assumption is limiting to the extent that unexpected news about the drug between discovery and FDA approval affects its profitability, forcing the market to update its value at approval.

4.2 Abnormal Returns and Stochastic Discounting

Equations (1) and (2) show that to identify the value of drugs we have to first estimate the predicted CAR at FDA approval, $\mathbb{E}\left(\texttt{CAR}_{\texttt{appr}}\right)$, and the expected discount rate $\mathbb{E}(\delta^{\tau_{\texttt{disc}\rightarrow}})$ from the discovery stage to FDA approval. Below, we outline the estimation of abnormal returns and present the details in Appendices A.1 and B.1.

Abnormal Returns. We use an unrestricted market model to estimate the abnormal returns and then aggregate them over one day before and two days after each announcement to determine the CAR.[8] Using the estimated regression coefficients (in Table A.2), we can determine the expected change in the firm's market value given an announcement.

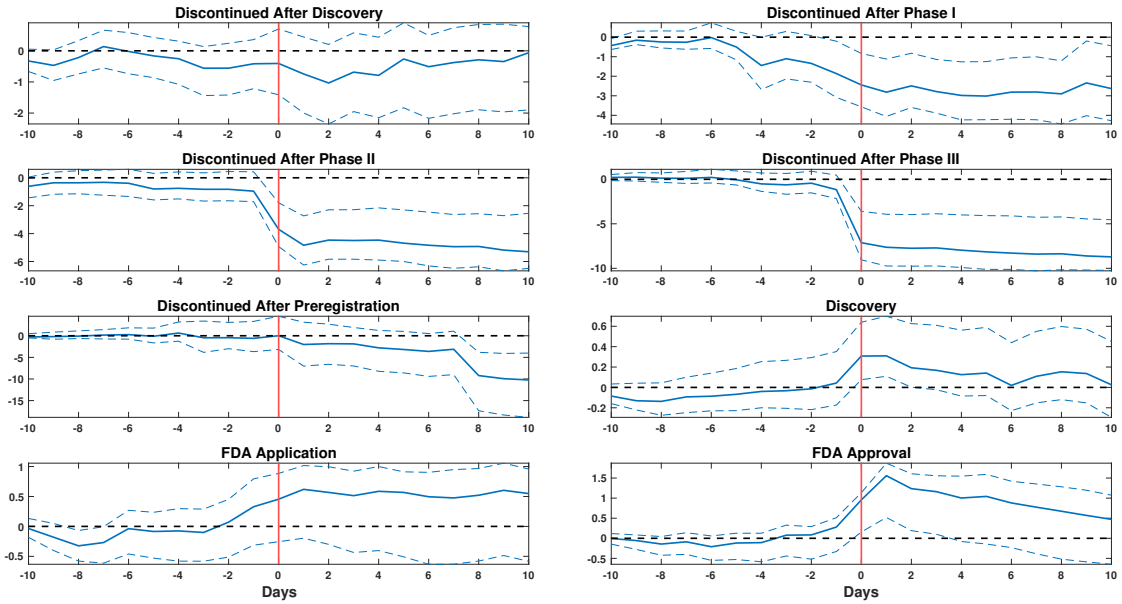
Table A.1 presents summary statistics of CAR. We find that the mean CAR is positive for positive announcements, e.g., successful transition to the next stage of development and FDA approval, and negative for negative announcements, e.g., discontinuations.

Possible information leakage before the official announcement is a concern. 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 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 expect a non-zero cumulative abnormal return (CAR) only on the announcement day, reflecting new information about the relevant drug's prospects, such as an increased probability of success. We also anticipate CAR to

[8] In practice, we use linear regression (A.3) to decompose CAR as a function of various announcements, i.e., discovery, FDA application, FDA approval, and discontinuations at different stages.

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



Note: The figures display CAR, expressed in percentage points on the y-axes, for a ten-day 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.

"jump" at time 0, indicating a sharp change in response to the announcement. Systematic increases or decreases in the CAR before an announcement would suggest information leakage. However, the figure demonstrates that CAR in the days leading up to the announcements is statistically insignificant, providing little evidence of information leakage across announcement types. The primary change in CAR consistently occurs on the announcement day, as illustrated by the panels corresponding to each announcement type.

Stochastic Discounting. To estimate the expected discount rate, we first estimate the probability distribution of time to success, using the competing risk model of Aalen (1976). In particular, we use the observed time it takes for drugs to transition to the next stage (which we call "success") to estimate $\mathbb{P}_{k'}(\tau \mid S_k = 1)$–the probability that a drug will move to stage k to stage k' by year τ. Then we can estimate the expected discount rate using Monte Carlo simulation, i.e., $\mathbb{E}(\delta^{k\to k'}) = \mathbb{E}(\delta^{\tau}) \approx \frac{1}{L}\sum_{\ell=1}^{L}\delta^{\tau_\ell}$, where $\tau_\ell \sim \mathbb{P}_{k'}(\cdot \mid S_k = 1)$,

where $L = 10,000$ is the number of simulation draws. In the homogeneous case, we pool data across all drugs to estimate the success probabilities and discounting. We consider 11 major diseases in the heterogeneous case and estimate probabilities for each separately. The estimates of the expected discount rate, for $\delta = 0.98$ are presented in Tables B.1 and B.2.

4.3 Estimated Values

Having estimated the success rates and the expected discount rates, we incorporate these estimates into equations (1) and (2) to determine the value of drugs at approval and discovery, respectively. To begin, we consider the case of homogeneous expectations across all diseases, assuming that the success rates and discount rates are consistent regardless of the therapeutic area. This approach allows us to establish a baseline understanding of drug values before delving into the more complex scenario of heterogeneity in success rates across different diseases.

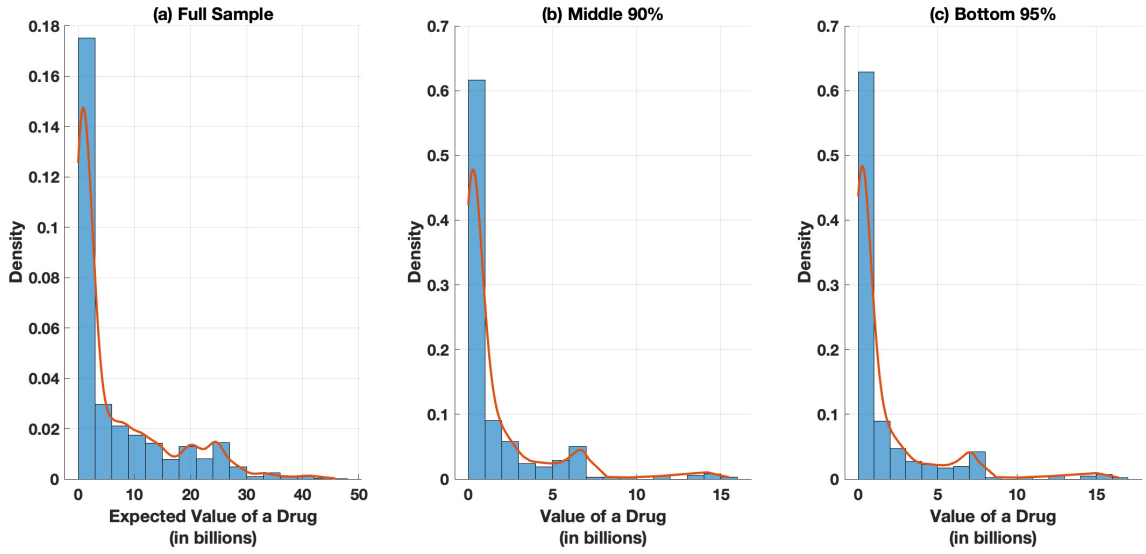
4.3.1 Homogeneity Across Diseases

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). We use the estimates of (homogeneous) transition probabilities from Table 2 and expected CAR associated with approval announcements from Table A.2.

Figure 5 shows the histograms of the expected drug values measured in billions of dollars. Figure 5-(a) considers only the drugs that received FDA approval in the full sample. There are 1,561 such drugs, and we can use Equation (1) to identify 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.

Given the pronounced right skewness in the firm size distribution, we estimate the value of drugs after removing outlier firms to ensure that extreme cases do not drive our results. This decision is motivated by our discussion of the market capitalization distribution in

Figure 5: **Expected Value of Individual Drugs**



Note: The figure shows histograms and densities of $\mathbb{E}(V|S_{\text{appr}} = 1)$ (in billions of US 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 C.1, effects of announcements on CAR from Table A.2, and the discount rates from Tables B.1.

Section 3.3, which suggests that large and small firms may have fundamentally different characteristics.

We consider two restricted samples. First, we remove firms with a market capitalization below the 5th or above the 95th percentile, creating a sample that we call the "middle 90%." This sample excludes the smallest and largest firms, focusing on firms in the middle of the size distribution.

Second, we consider a sample where we remove only the largest firms with a market capitalization above the 95th percentile, which we call the "bottom 95%." This sample allows us to examine the impact of excluding the most extreme outliers at the upper end of the size distribution while retaining a larger proportion of the data.

By considering these two restricted samples, we assess the robustness of our estimates and determine the extent to which the inclusion or exclusion of outlier firms influences our results.

We separately estimate each sample's expected CAR, transition probabilities, and dis-

count rates. These estimates are presented, respectively, in Tables A.2, B.1, and C.1. Then, substituting them in (1) gives us the desired estimates. Figures 5-(b) and (c) show the histograms of values of successful drugs for the middle 90% (378 drugs) and bottom 95% samples (401 drugs). Unlike the full sample, these samples have fewer outliers. We also present the average values in Table 5. We find that the mean of the expected values is $1.62 billion among the 378 drugs developed by mid-90% firms, which is smaller than in the full sample. For the bottom 95% sample, the average value of a drug is $1.60 billion.

Table 5: **Expected Value of Drugs**

	Full Sample	**Middle 90%**	**Bottom 95%**
At Approval			
All Drugs	$6.83 bil	$1.62 bil	$1.60 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 drugs, 29 belong to the Middle 90%. The row "Average" refers to drugs for which we observe only a few stages.

Our dataset includes a subsample of 84 drugs for which we have complete information on their development path from discovery to FDA approval. This comprehensive data allows us to determine the values and costs of these drugs without relying on estimates derived from other drugs with incomplete development path information. While our primary analysis utilizes all available drug data, we also conduct a sensitivity analysis using this subsample of drugs with complete paths to assess the robustness of our findings.

The results of this sensitivity analysis are presented in Table 5, under the row labeled "Drugs with Complete Path." When considering the middle 90% and bottom 95% samples, which exclude outlier firms based on market capitalization, we find that the estimated values for the 29 drugs with complete paths are $1.89 billion and $1.99 billion, respectively. These

estimates are similar to the values obtained when analyzing all approved drugs within the corresponding samples, suggesting that our findings are consistent even when focusing on drugs with more comprehensive data.

Value at Discovery. Building upon our estimation of the expected value of a drug at the time of approval, we can also calculate the value at the time of discovery using (2).

The second half of Table 5 presents the average discounted values of drugs at the discovery stage. When considering the entire sample, which includes all firms regardless of their market capitalization, we estimate the expected value of a drug at discovery to be $331.12 million. However, as discussed earlier, outlier firms with extremely high or low market capitalizations may influence these estimates.

We also examine the subsamples that exclude outlier firms to address this concern. When considering the middle 90% sample, which removes firms below the 5th percentile and above the 95th percentile of market capitalization, the mean expected value of a drug at discovery is $63.37 million. Similarly, when focusing on the bottom 95% sample, which removes only the largest firms above the 95th percentile, the mean value at discovery is $65.97 million.

These subsample estimates provide a more conservative valuation of drugs at the discovery stage, as they are less influenced by the extreme values associated with outlier firms. By comparing the results across different samples, we can gain a more comprehensive understanding of drugs' potential value at the beginning of the development process.

To ensure that the magnitude of our estimated values is reasonable, we compare them with observed sales data. Since the expected value represents the net present value of the cash flow accruing to the firm, our estimates should be similar to those from discount sales.

We obtain sales data from the Cortellis Competitive Intelligence database, which provides information on yearly drug-level total worldwide sales for a subset of drugs. However, a challenge arises because the sales data are at the drug level, while the announcements are at the drug-disease level. The database does not provide a breakdown of sales by disease, so we allocate the sales equally across all diseases associated with each drug. For example, if a

drug targets three conditions, each condition would be assigned one-third of the total sales.

Table 6: **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 6 presents the descriptive statistics for the sales data. One limitation of this data is that it typically covers only about ten years, which is relatively short for pharmaceutical sales, as many drugs remain on the market for over 20 years. To address this issue of short panel duration, we calculate the average sales across all available years for each drug-and-disease pair associated with a given firm.

Assuming that this average sales value is received annually, we calculate the discounted total sales for each drug-disease-firm pair.[9] Table 7 presents the results for different time horizons after aggregating the average sales. We find that the total discounted sales of a drug, assuming it had been on the market for ten years, is $2.4 billion for the full sample and $1.4 billion for both the middle 90% and bottom 95% samples. Considering that generic drugs may drive prices down over time, the observed sales presented in Table 7 support the reasonableness of our approach in estimating the expected value of a drug.

4.3.2 Heterogeneity Across Diseases

The estimates presented thus far are based on homogeneous transition probabilities and distributions of the time to success. While the transition probabilities in Table 2 are con-

[9]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 7: **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).

sistent with the averages reported in the literature (e.g., Hay et al., 2014), they likely vary across diseases or indications (e.g., Wong et al., 2019). For example, the likelihood of developing a successful drug for neoplastic illnesses, which include cancer, may differ from that of endocrine diseases, which include diabetes. If this is the case, all else being equal, an announcement (e.g., FDA approval) should affect firms' values differently depending on the relevant indications.

To account for this variation, we consider 11 major indications and estimate the valuations at FDA approval and discovery for each, following the same steps as in Table 5. An indication represents the most aggregated level of disease classification available in our sample. From the list of 34 indications, we select 11 major indications with sufficient announcements (approximately 1,000). These indications are listed in Table 8.

For each indication, we estimate expected cumulative abnormal returns for each type of announcement implied by the estimates of (A.3), the transition probabilities in Table C.2 and the expected discount rates from discovery to market in Table B.2.[10] To keep the exposition concise, we focus only on the full and middle 90% samples.

Using these indication-specific estimates in Equation (1), we determine the expected value of drugs at approval, i.e., $\mathbb{E}(V|S_{\texttt{appr}} = 1)$, for each indication. Then, using Equation (2) again, we determine the expected value of drugs at the discovery stage, i.e., $\mathbb{E}(V|S_{\texttt{disc}} = 1)$.

[10]Table C.2 in the appendix presents the estimated transition probabilities from the discovery stage to the market and from FDA application to the market, separately for each indication. As before, the ex-ante probability that a drug is successful is small and differs across indications. For instance, the likelihood of developing a successful cure for neoplastic diseases is 9% and 16% for endocrine diseases. Conditional on reaching the FDA application stage, the likelihood is high, and these probabilities differ across indications.

Table 8 presents the average values of drugs (in billions of US dollars), where the averages are taken across all drugs within each indication.

Table 8: **Expected Value of Drugs, by Indications**

Indications	**Full Sample**		**Middle 90%**	
	Approval	Discovery	Approval	Discovery
Cardiovascular diseases	\$4.96	\$0.26	\$4.01	\$0.25
Endocrine diseases	\$11.40	\$1.10	\$15.30	\$1.40
Gastrointestinal diseases	\$2.26	\$0.17	\$0.87	\$0.05
Hematological diseases	\$11.60	\$1.28	\$6.06	\$0.56
Immune disorders	\$22.60	\$1.68	\$40.60	\$2.74
Infectious diseases	\$0.91	\$0.09	\$-0.40	\$-0.02
Inflammatory diseases	\$5.45	\$0.34	\$6.84	\$0.35
Musculoskeletal diseases	\$16.00	\$1.71	\$7.85	\$0.69
Neoplastic diseases	\$3.32	\$0.18	\$2.66	\$0.09
Neurological diseases	\$8.07	\$0.41	\$2.63	\$0.19
Rare diseases	\$15.30	\$1.90	\$7.83	\$0.82

Note: The table presents the mean of the expected value of individual drugs (in billions of US dollars) at the time of approval, $\mathbb{E}(V|S_{\texttt{appr}} = 1)$, and at discovery, $\mathbb{E}(V|S_{\texttt{disc}} = 1)$, for 11 indications. The 90% sample refers to the drugs developed by firms with real market capitalization between 5% and 95% of the entire sample.

As shown in Table 8, the average values of drugs differ significantly across indications. For instance, the average value of drugs for infectious diseases is numerically zero. This estimate is consistent with the fact that infectious diseases, such as pneumonia, sinusitis, and urinary tract infections, are often treated with antibiotics, which tend to have small values (e.g., Missialos et al., 2010). In contrast, the average value of approved drugs that treat immune disorders, which include type 1 diabetes, is \$40 billion. At the discovery stage, the average value for immune disorder drugs is \$2.74 billion. The value of drugs at the discovery stage is substantially smaller than at approval, underscoring the ex-ante risk of drug development.

5 Cost of Drug Development

In this section, we leverage the estimated drug values to shed light on the average cost of drug development. This exercise offers valuable insights into the financial consequences of drug

development failures and the total investment needed to bring a drug to market successfully. For ease of presentation and without loss of generality, we use only the estimates from the homogeneous case (Section 4.3.1).

5.1 Total Cost at Discovery

To begin with, we present the estimate of the total expected cost at the discovery stage. We show that the cost can be identified directly from the expected value at discovery *without additional assumptions beyond the ones used so far.*

Again, we begin by introducing some notation. Let the random variable C denote the firm's development cost, and let $C_{k\to k'}$ denote the cost from stage k to stage k', and $C_{k\to}$ denote the cost from stage k to the end (FDA approval). For instance, $C_{\texttt{disc}\to}$ and $C_{\texttt{disc}\to\texttt{appl}}$ denote costs from discovery to approval and from discovery to FDA application, respectively. Lastly, and similar to the expected value, let $\mathbb{E}(C_{k\to k'}|S_k = 1)$ denote the expected cost of developing drug from stage k to k' conditional that reaching stage k, i.e., $S_k = 1$. Let the expected development cost from discovery to FDA approval, i.e., $\mathbb{E}(C_{\texttt{disc}\to}|S_{\texttt{disc}} = 1)$.

Let us consider the discovery announcement. 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. Following the same reasoning leading to (1), we get

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

The value at discovery, $\mathbb{E}(V|S_{\texttt{disc}} = 1)$, is identified from (2). Substituting it in (3) we can determine the cost $\mathbb{E}(C_{\texttt{disc}\to}|S_{\texttt{disc}} = 1)$.

In practice, we can calculate the average expected cost of drug development at the discovery stage by averaging both sides of (3) across all drugs and substituting the average expected value of drugs at discovery. The estimates are shown in Table 9, Panel (a). For

Table 9: **Costs Estimates**

Stages	**Middle 90%**	**Bottom 95%**
Panel (a): **Total Costs**		
All Drugs	\$58.51 mm	\$60.72 mm
Drugs with Complete Path	\$69.24 mm	\$77.01 mm
Panel (b): **FDA Review**	\$638.75 mm	\$648.04 mm
Panel (c): **Clinical Trials**		
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 at different stages of drug development. Panel (a) shows 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 row "All Drugs" refers to all the drugs in our sample, and, "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. Panel (b) shows the costs of FDA review and application at the time of discovery, and Panel (c) shows the average of the expected cost of the three phases of clinical trials. The middle 90% sample refers to the drugs developed by firms with real market capitalization between 5% and 95% of the sample.

the middle 90% sample, the expected cost is estimated to be \$58.51 million; for the bottom 95% sample, it is estimated to be \$60.72 million. This cost is risk-adjusted at the time of discovery because it incorporates the risk that the drug will fail with a high probability.

5.2 Cost of FDA Review and Application

Next, we estimate the mean costs of drug development at different stages by leveraging the valuations and examining the effects of discontinuation announcements on a firm's value. Most drugs do not succeed and are discontinued at various stages of the development process. When a firm announces that it is stopping the development of a drug, the product of the cumulative abnormal return (CAR) and the market capitalization `mktcap` should be negative, reflecting the decrease in the firm's value following this "bad news." Under some additional assumptions (made clear below), this change in the firm's market value provides information about the remaining cost of drug development.

To recover the costs associated with each stage of drug development, we use discontinua-

tion announcements at different stages (see Table 3). We use discontinuation announcements, but not the announcements about the start of clinical trials, because the discontinuation announcements are more likely to be "surprise" announcements than the initiation of clinical trials. Furthermore, discontinuations are enacted immediately, which is not the case for announcements about the start of clinical trials.

We also assume that, on average, there is no negative selection between discontinued drugs and their (ex-ante) profitability. This assumption is reasonable because the primary reason for discontinuations is negative clinical trial results; see, for example, DiMasi (2013) and Khmelnitskaya (2022).

Now, let us consider the process of identifying the costs of FDA review and application. Even though most scientific experiments are completed at the time of application, additional expenses are still involved in setting up manufacturing capacity, as well as legal and administrative fees.[11] Our estimate is, therefore, the total of these different costs.

To identify the costs, we use discontinuation announcements made just before the FDA application—that is, we use the announcements made after Phase III clinical trials. When a firm announces the discontinuation, the firm's market value *decreases* by the amount the market expected the drug to earn, $\mathbb{E}(V|S_{\texttt{phase III}}=1)$, and *increases* by the cost savings, $\mathbb{E}(C_{\texttt{appl}\rightarrow}|S_{\texttt{appl}}=1)\times p_{\texttt{appl|phase III}}$, that no longer need to be incurred.

Therefore, the change in the firm's market value, $\mathbb{E}\left(\texttt{CAR}_{\texttt{drop after phase III}}\right)\times\texttt{mktcap}$, is

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

Then, under the same assumptions that led to (2), and following the same steps but adapted

[11] The FDA has prepared a set of instructions for drugs to receive approval, which clarifies 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" (Food and Drug Administration, 2010).

to Phase III, we can recover the expected value of a drug at Phase III as

$$\mathbb{E}(V|S_{\texttt{phase III}}=1)=\mathbb{E}(V|S_{\texttt{appr}}=1)\times\mathbb{E}(\delta^{\tau_{\texttt{appl}\rightarrow}})\times p_{\texttt{appr}|\texttt{appl}}\times p_{\texttt{appl}|\texttt{phase III}}. \tag{5}$$

Then, substituting (5) in (4) we can identify $\mathbb{E}(C_{\texttt{appl}\rightarrow}|S_{\texttt{appl}}=1)$.

Table 9, Panel (b) presents the estimates of the expected cost of FDA application and review. For the middle 90% sample, we estimate the expected cost to be \$638.75 million, while for the bottom 95% sample, the estimate is \$648.04 million. Based on the earlier discussion, these cost estimates should be interpreted as the remaining costs the firm must incur between application submission and FDA approval.

5.3 Cost of Clinical Trials

Next, we estimate the costs of all three phases of clinical trials. We can use discontinuations from one step before to identify the cost of a clinical trial. For instance, we can use discontinuation announcements made after Phase II (but before Phase III) clinical trials to identify the cost of Phase III clinical trials. Similar to our reasoning in (4), the change in the firm's market value informs the cost of the Phase III clinical trial. In particular, the change in the firm's market value following the discontinuation announcement after Phase II is

$$\begin{aligned}\mathbb{E}\left(\texttt{CAR}_{\texttt{drop after}-\texttt{phase II}}\right)\times\texttt{mktcap} \;=\; &\underbrace{-\Big(\mathbb{E}(V|S_{\texttt{phase II}}=1)-\mathbb{E}(C_{\texttt{appl}\rightarrow}|S_{\texttt{phase II}}=1)\Big)}_{\texttt{net value lost after discontinuation}}\\ &+\underbrace{\mathbb{E}(C_{\texttt{phase III}\rightarrow\texttt{appl}}|S_{\texttt{phase III}}=1)\times p_{\texttt{phase III}|\texttt{phase II}}}_{\texttt{cost savings}}.\end{aligned} \tag{6}$$

As before, under the same assumptions that led to (2), the expected value at Phase II clinical trials is given by $\mathbb{E}(V|S_{\texttt{phase II}}=1)=\mathbb{E}(V|S_{\texttt{appr}}=1)\times\mathbb{E}(\delta^{\tau_{\texttt{phase III}}\rightarrow})\times p_{\texttt{appr}|\texttt{appl}}\times p_{\texttt{appl}|\texttt{phase III}}\times p_{\texttt{phase III}|\texttt{phase II}}$ and the expected cost from application to FDA approval is given by $\mathbb{E}(C_{\texttt{appl}\rightarrow}|S_{\texttt{phase II}}=1)=\mathbb{E}(C_{\texttt{appl}\rightarrow}|S_{\texttt{appl}}=1)\times\mathbb{E}(\delta^{\tau_{\texttt{phase III}}\rightarrow\texttt{appl}})\times p_{\texttt{appl}|\texttt{phase III}}\times p_{\texttt{phase III}|\texttt{phase II}}$. Substituting these into (4) identifies $\mathbb{E}(C_{\texttt{phase III}\rightarrow\texttt{appl}}|S_{\texttt{phase III}}=1)$.

Although we do not show them here, following similar steps, we can use discontinuations at earlier stages to identify the expected costs of running Phase I and II clinical trials. For instance, we can use discontinuation announcements made after discovery (but before Phase I clinical trials) to identify the costs of Phase I clinical trials. Likewise, we can use discontinuation announcements after Phase I clinical trials to identify the costs of Phase II.

It is important to note that to identify these costs, we rely on our estimates of the values, transition probabilities, discount rates, and the estimated costs of later stages, including Phase III clinical trials, FDA application, and review. So, implicitly, we are maintaining all the previous assumptions sufficient to estimate those parameters.

Table 9, Panel (c) presents the estimation results. Using discontinuation announcements after discovery, we estimate the expected cost of Phase I clinical trials for the two samples to be \$620,510 and \$219,240, 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 our estimates of the costs of clinical trials, the transition probabilities (Table C.1), and expected discount rates (Table B.1), 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 for the bottom 95% sample.

5.4 Cost Estimates in the Literature

Cross-checking the validity of our proposed methodology is essential, and comparing our cost estimates with prior estimates from the literature serves this purpose. For instance, if we use the estimates of transition probabilities, average durations for each phase, and average costs for each stage from DiMasi et al. (2016), the discounted expected cost of

clinical development would be approximately $51.2 million (in December 2020 US dollars). While this estimate is larger than the comparable estimate of $11.8 million obtained using our preferred specification, it is important to note that estimates from DiMasi et al. (2016) are known to be higher than other estimates in the literature (Wouters et al., 2020).

In contrast, Sertkaya et al. (2014) suggest that the cost of bringing a new drug to market ranges between $161 million and $2 billion, providing a broader spectrum of estimates.

Similarly, Sertkaya et al. (2016) use proprietary cost data on a sample of drugs developed from 2004 to 2012 to estimate the average non-risk-adjusted (i.e., accounting) costs of the three phases of clinical trials. They find that these costs range from $1.82 million to $8.58 million for Phase I, $9.1 million to $25.48 million for Phase II, and $14.95 million to $68.77 million for Phase III (in December 2020 US dollars). Finally, Adams and Brantner (2006) estimate the cost per new drug to be $868 million for drugs entering human clinical trials for the first time between 1989 and 2002.

6 Designing Schemes to Support Drug Development

In this section, we consider the potential applications of our findings for policymakers seeking to address the declining productivity in drug development (Munos, 2009; Pammolli et al., 2011; Scannell et al., 2012). The methodology and estimation results presented in this study can serve as a valuable tool for designing effective support systems and interventions to foster pharmaceutical innovation.

We propose two potential avenues for policy intervention. First, we consider the government using our estimates to promote "open innovation" through *drug buyouts,* which involve purchasing manufacturing rights of drugs and placing them in the public domain. We consider two stages at which drug buyouts could be implemented: (i) after FDA approval and (ii) at the discovery stage. Although the current system is not designed for drug buyouts, our analysis is motivated by considerations similar to those discussed in patent buyouts (Kre-

mer, 1998) and transferable patents Dubois et al. (2022), among others. Second, we examine the possibility of the government adopting a cost-sharing strategy to support innovation by covering some development costs. For ease of presentation, we only use estimates from the homogeneous case (Section 4.3.1), but the lessons apply to any disease indication.

6.1 Drug Buyout After FDA Approval

Let us first consider the case of the government buying out a drug after FDA approval. All uncertainties associated with the drug's development have been resolved at this stage, and all major R&D costs have been incurred. All that remains is to determine the buyout price.

As mentioned in the introduction, a key concern for a policy intervention based on our estimates is the *Lucas critique*: when the regulatory environment changes, agents re-optimize and change their decisions, potentially nullifying the intervention. To make progress and understand how this market feature may constrain the drug buyout policy, we begin by setting aside the Lucas critique and assuming that the government offers to buy the drug at the value we estimated using Equation (1).

To illustrate this, consider an example where CAR associated with FDA approval is 1%, the market capitalization is \$1 billion, and the transition probability from FDA application to FDA approval is 0.9. Then the value of the drug is $\mathbb{E}(V|S_{\texttt{appr}} = 1) = (1\% \text{ of } \$1 \text{ billion})/(1 - 0.9) = \$100$ million. In this scenario, the government would offer to pay \$100 million for the drug, and the firm would agree to sell it at this price.

Now, suppose the social value of the drug is $W = \$500$ million, and it is common knowledge. The market values of drugs tend to be substantially smaller than their social values (e.g., Chabot et al., 2004; Schrag, 2004; Howard et al., 2015; Conti et al., 2020). If the government announces that it intends to implement the drug buyout policy, *before* the FDA approval, then after the announcement of FDA approval, the market will react as if the value of the drug is \$500 million and not \$100 million because it expects the government to pay up to its social value. In other words, if the market anticipates that the government

will buy the drug, investors will "bid" the drug's value up to the amount they expect the government to pay, which is likely to be the social value.

Alternatively, the government could wait until after FDA approval to announce its intention to buy the drug. In this case, the unobserved market value will be revealed when the drug is approved. At this point, the government can make a take-it-or-leave-it offer to purchase the drug at its market value.

However, this approach has a significant drawback: it can only be a one-off intervention. If the market anticipates that the government might buy the drug, investors will bid the drug's value up to the social value, W, and hedge their positions. For example, they could use *put options* that can be exercised if the drug receives FDA approval and the government does not purchase it. More generally, the government's actions will influence market behavior and reactions, ultimately affecting our estimates of the drugs' market value.

Furthermore, paying a markup $\alpha > 0$ over the estimated market value, as suggested by Kremer (1998) for patents, i.e., paying $(1+\alpha) \times \mathbb{E}(V|S_{\texttt{appr}} = 1)$, does not solve the problem either. In this case, the market can still "manipulate" the final payment by reacting to the total payout rather than the "true" value. Consequently, when considering these drug buyout interventions, we cannot rely on market reactions to FDA approval announcements to determine the market value of a newly approved drug.

To address Lucas' critique, drug buyouts must be designed to prevent the market from anticipating the government's decision *and* bidding up the drug's value. We propose a solution that does not require the government to commit to an offer, making it unprofitable for investors to hedge their positions.

We envision that the government announces its intention to implement a drug buyout with a small *exogenous* probability $\varphi \in (0, 1)$ and pay the social value of the drug W. Then, the expected value of the drug is

$$\text{Expected Value} = (1-\varphi) \times \mathbb{E}(V|S_{\texttt{appr}} = 1) + \varphi \times W.$$

Next, note that Equation (1) identifies the expected value of the drug after FDA approval, which is the left-hand side of the above equation. Then, simplifying the above equation, we can identify the market value of the drug as

$$\mathbb{E}(V|S_{\texttt{appr}}=1)=\frac{\text{Expected Value}-\varphi\times W}{1-\varphi}.$$

Thus, by committing to implementing the drug buyout program with probability $0<\varphi<1$, the government can run the program and estimate the market value of drugs. We can use the estimated market value of drugs to design rules that can lower the cost of drug buyouts.

Suppose the government announces the probability φ and the price $\Pi_0 \leq W$. Let $F_v(\cdot)$ denote the distribution of the market value of drugs, corresponding to the density in Figure 5. The expected savings from offering Π_0 instead of W is

$$\text{Expected Savings}=\varphi\times(W-\Pi_0)\times F_v(\Pi_0),$$

where $F_v(\Pi_0)$ is the probability that the drug's market value is smaller than the offered Π_0 and the firm accepts the offer. So, the government's problem is to choose Π_0 that maximizes the savings. Under the assumption that $F_v(\cdot)$ satisfies the regularity conditions in Myerson (1981), the optimal price Π_0^* is determined by: $\Pi_0^*+\frac{F_v(\Pi_0^*)}{f_v(\Pi_0^*)}=W$, where the left-hand side is the marginal cost of offering Π_0^*, and the right-hand side is the marginal benefit.

As $\Pi_0^*<W$, the government can use our estimates to lower costs. We envision that as the government continues the buyout program, we would use the estimated market value $\mathbb{E}(V|S_{\texttt{appr}}=1)$ from the previous "round" to update the values to be used the next time. To see that, note that because the value and Π_0^* are known to the market before the approval announcement, there will be two cases: (1) $\mathbb{E}(V|S_{\texttt{appr}}=1)<\Pi_0^*$ and the market knows that the drugs will be sold at Π_0^* with probability φ, and (2) $\mathbb{E}(V|S_{\texttt{appr}}=1)>\Pi_0^*$ and the market

knows that the drug will not be sold to the government. Then from (1), we get

$$\mathbb{E}\left(\texttt{CAR}_{\texttt{appr}}\right) \times \texttt{mktcap} = \begin{cases} \Big((1-\varphi) \times \mathbb{E}(V|S_{\texttt{appr}}=1) + \varphi \times \Pi_0^*\Big) \times (1 - p_{\texttt{appr}|\texttt{appl}}), & \text{if } \mathbb{E}(V|S_{\texttt{appr}}=1) \leq \Pi_0^* \\ \mathbb{E}(V|S_{\texttt{appr}}=1) \times (1 - p_{\texttt{appr}|\texttt{appl}}), & \text{if } \mathbb{E}(V|S_{\texttt{appr}}=1) > \Pi_0^* \end{cases},$$

where everything is identified, except the value, $\mathbb{E}(V|S_{\texttt{appr}}=1)$, for the current drug. Thus, we can identify the market value of the drug from the above equation and then update the distribution $F_v(\cdot)$ to be used for the next drug.

6.2 Drug Buyout At the Discovery Stage

Let us consider implementing the drug buyout scheme at the start of the discovery stage. This policy intervention faces different tradeoffs compared to the intervention after FDA approval. The main difference is that, at the discovery stage, the uncertainty associated with drug development has yet to be resolved, and the development costs are still ahead. In contrast, after FDA approval, all uncertainties are resolved, and R&D costs are sunk.

At the time of discovery, the government and the market only know the expected value based on previously successful drugs. So, they have symmetric information about the value of the drug. Our estimates (Table 5) suggest that the average value of drugs at the discovery stage is $63.37 million. Neither party knows the ex-post market value of the drug in question.

So, if the government can pay $63.37 million for a drug, the market cannot affect the payment like it could earlier. Thus, unlike the post-approval scenario, the government will not be "forced" to pay the (discounted) social value of the drug because there is still uncertainty associated with drug development. This observation is crucial, as it highlights that the drug buyout policy at the early stage of development is not subject to the Lucas critique insofar as the drug's value at discovery is concerned.

Next, suppose the government buys a drug at the discovery stage, pays the present discounted expected value, and places it in the public domain, allowing anyone to develop it

further. This approach raises another fundamental question: Who pays for the subsequent development costs? Once the drug is in the public domain, the monopoly profit may no longer accrue to the first firm that obtains approval. Generic equivalents of the drug can enter the market relatively quickly through the abbreviated new drug application process, which is the express objective of the drug buyout program.

To address the issue of who pays for the development costs, the government could complement the early-stage drug buyout program with another program that incentivizes drug development.

One possible approach is to cover ex-ante R&D costs for each successful completion of a stage. If we are willing to make assumptions in Section 5 to estimate the drug development costs, then based on our estimates in Table 5, the average expected value of a drug at discovery is $63.37 million, and the development cost is $58.51 million. 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 completes the discovery stage. Further payments, determined below, can be made after each successful step.

Next, consider a scenario where the drug completes pre-clinical trial research and is about to enter Phase I clinical trials. The firm is expected to incur $0.62 million in costs for these trials (Table 9, Panel (c)), which the government can cover. Similarly, the government can promise to pay $30.48 million (Table 9, Panel (c)) if a firm completes Phase II clinical trials and $41.09 million for the successful completion of Phase III clinical trials.

While this solution is straightforward, it is potentially subject to moral hazard because firms' research efforts are non-contractible. After receiving the payment, the firm may not exert the necessary effort to develop the drug.

To overcome this challenge, we consider two approaches. The first approach involves all interested firms reporting their R&D costs as the drug progresses through the development process. With the cost information, the government could design an incentive (research) contract (e.g., Laffont and Tirole, 1986), that balances R&D costs and (unobserved) research

effort. The second approach involves using advanced market commitment (Kremer and Glennerster, 2004; Kremer et al., 2020, 2022). Under this approach, the government commits to buying a certain number of FDA-approved drugs from the firm(s) at a pre-fixed price. If the commitment is sufficiently large, it will incentivize R&D because it can be viewed as a prize for the early developers who can "pull" the R&D efforts.

Although determining the exact (incentive or advanced market) mechanism is beyond the scope of our paper, having the value and cost estimates allows policymakers to design appropriate mechanisms to boost pharmaceutical drug innovations.

6.3 Cost-Sharing Agreements

Our cost estimates can be useful for the drug buyout program and cost-sharing agreements between the government and pharmaceutical companies. Under these agreements, the government provides funding to cover development costs, either fully or partially, in exchange for the firm agreeing to sell its drugs to the government at a pre-fixed price after FDA approval. A notable example of a research cost-sharing agreement is the funding provided to develop the COVID-19 vaccines under *Operation Warp Speed*.[12]

We can envision two types of cost-sharing agreements between the government and pharmaceutical companies. In the first type of agreement, the government pays $58.51 million of R&D costs at the discovery stage, which is the expected cost of drug development when evaluated at discovery (see Table 9, Panel (a), All Drugs). For simplicity, let's assume that the fixed cost of operation is either zero or, if necessary, it is included in the average yearly revenue as assumed in Equation (2). In return, the firm would commit to selling $q = \$58.51/\$p$ units for free, for pre-negotiated price $\$p$.

The second type of agreement is a development stage contingent cost-sharing agreement. Using the estimates in Table 9 Panel (c), the government can offer to cover $620,000, $30.48 million, and $41.09 million of Phase I, II, and III clinical trial expenses, respectively. Upon

[12]For COVID-19, the U.S. government used a combination of cost-sharing and advanced market purchase agreements; see Zimmer (2020); Congressional Budget Office (2021) and Restuccia and Hopkins (2022).

successful completion of clinical trials, the government can offer to pay \$638.75 million towards FDA review and application costs based on the estimates in Table 9 Panel (b). This sequential payment structure can lower the government's overall costs. If the FDA approves the drug, the firm delivers $q = \$(0.62 + 30.48 + 41.09 + 638.75)\text{million}/\$p$ units for free.

7 Conclusion

In this paper, we propose a tractable framework to value pharmaceutical innovations by combining an event-study approach with a model of discounted cash flows. Our approach relies on market responses to drug development announcements. We demonstrate that, under additional assumptions, our method can be extended to estimate drug development costs at different stages, from pre-clinical research to FDA approval.

Using data from public firms, we estimate that, on average, the market value of a drug at the time of FDA approval is around \$1.62 billion. In contrast, at the time of discovery, the value is \$63.37 million. We also find that the values of drugs differ substantially across indications. Furthermore, we estimate the development cost to be \$58.51 million and the costs of the three phases of clinical trials to be approximately \$0.62, \$30.48, and \$41.09 million, respectively.

We explore a broader context for these estimates by investigating how policymakers might use them as inputs in designing policies that support drug development. We consider the use of our estimates by a government that wants to "open innovation" via drug buyouts (either after FDA approval or at the discovery stage) and place them in the public domain. We also examine the possibility of the government adopting cost-sharing agreements to support innovation by covering development costs.

Our estimates suggest several important areas for future research. First, our approach excludes drugs developed by large pharmaceutical firms with a market valuation above the 95th percentile of the firm size distribution. To the extent that these firms develop different

types of drugs, our approach fails to capture those drugs. To relax this assumption, we could consider large firms separately and keep track of announcements about acquired drugs.

Second, a competitor's announcements can affect a firm's drug valuation. For instance, if two firms are developing competitive drugs, the impact of one firm's announcement on the market value of the other firms would provide information about the expected effect of competition. We can adapt our approach to include "competitive announcements" in our estimation method and plan to work on this in the future.

Third, in this paper, we focused only on single announcements and ignored days with more than one announcement. The difficulty lies in determining the change in a firm's market value in response to each announcement when there are multiple announcements on any given day. One approach to address this issue would be to use a binomial pricing tree (Shin, 2003) to determine the change in the firm's value following multiple announcements.

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.

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.

Azoulay, Pierre, Joshua S. Graff-Zivin, Danielle Li, and Bhaven N. Sampat, "Public R&D Investments and Private-sector Patenting: Evidence from NIH Funding Rules," *Review of Economic Studies*, 2018, *86* (1), 117–152.

Chabot, Isabelle, Mireille M. Goetghebeur, and Jean-Pierre Grégoire, "The Societal Value of Universal Childhood Vaccination," *Vaccine*, 2004, *22*, 1992–2005.

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.

Cockburn, Iain M. 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.

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

Conti, Rena M., Jonathan Gruber, Daniel A. Ollendorf, and Peter J. Neumann, "Valuing Rare Pediatric Drugs: An Economic Perspective," *NBER Working Paper 27978*, 2020.

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.

Dubois, Pierre and Margaret K. Kyle, "Are Cancer Drugs Worth the Price? The Effects of Pharmaceutical Innovation on Cancer Mortality Rates," *Working Paper*, 2016.

_ **, Paul-Henri Moisson, and Jean Tirole**, "The Economics of Transferable Patent Extensions," *Working Paper*, 2022.

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.

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.

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

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

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.

Howard, David H., Peter B. Bach, Ernst R. Berndt, and Rena M. Conti, "Pricing in the Market for Anticancer Drugs," *Journal of Economic Perspectives*, 2015, *29* (1), 139–162.

Hwang, Thomas J., "Stock Market Returns and Clinical Trial Results of Investigational Compounds: An Event Study Analysis of Large Biopharmaceutical Companies," *PLOS ONE*, 2013, *8* (8), e71966.

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.

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.

Laffont, Jean-Jacques and Jean Tirole, "Using Cost Observation to Regulate Firms," *Journal of Political Economy*, 1986, *94* (3), 614–641.

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

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

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

Missialos, Elias, Chantal M Morel, Suzanne Edwards, Julia Berenson, Marin Gemmill-Toyama, and David Brogan, "Policies and Incentives for Promoting Innovative Antibiotic Research," Technical Report, European Observatory on Health Systems and Policies 2010.

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.

Myerson, Roger B., "Optimal Auction Design," *Mathematics of Operations Research*, 1981, *6* (1), 58–73.

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

Pakes, Ariel, "Patents as Options: Some Estimates of Value of Holding European Patent Stocks," *Econometrica*, 1986, *54* (4).

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.

Restuccia, Andrw and Jared S. Hopkins, "U.S. Agrees to pay $3.2 Billions for More Pfizer Covid Vaccines," The Wall Street Journal June 2022.

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.

Schrag, Deborah, "The Price Tag on Progress –Chemotherapy for Colorectal Cancer," *New England Journal of Medicine*, 2004, *351*, 317–319.

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.

Singh, Manish, Roland Rocafort, Cathy Cai, Kien Wei Siah, and Andrew W. Lo, "The Reaction of Sponsor Stock Prices to Clinical Trial Outcomes: An Event Study Analysis," *PLoS ONE*, 2022, *17* (9), e0272851.

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

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.

Zimmer, Carl, "Is Moderna in Operation Warp Speed?," The New York Times November 2020.

Appendix

A.1 Cumulative Abnormal Returns

Here, we estimate the cumulative abnormal return (CAR) associated with each drug-firm-announcement combination. 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.,

$$r_{i,t} = \alpha_i + \beta_i r_t + \varepsilon_{i,t}. \tag{A.1}$$

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 (A.1) using linear regression.[13]

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 (A.1), 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

[13] 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

$$\widehat{\texttt{CAR}}_{i,j,t_{(i,j)}} = \sum_{t=t_{(i,j)}-w_l}^{t_{(i,j)}+w_u} \hat{\varepsilon}_{i,j,t}, \tag{A.2}$$

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

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

Table A.1: **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 (A.2), by each type of announcement.

We observe that the magnitude of positive and negative announcements increases as we progress in the development process. For example, the average CAR associated with 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 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. Overall, these estimates suggest that the release of clinical trial results is an economically significant event and has meaningful effects on market value. However, negative news (discontinuation) has a larger impact than positive news.

Effects of Announcements on CAR

From Table A.1, 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

$$\begin{aligned}\widehat{\texttt{CAR}}_{i,j,t_{(i,j)}} &= \beta_{\texttt{disc}} \times \texttt{disc}_{i,j,t_{(i,j)}} + \beta_{\texttt{appl}} \times \texttt{appl}_{i,j,t_{(i,j)}} + \beta_{\texttt{appr}} \times \texttt{appr}_{i,j,t_{(i,j)}} \\ &\quad + \beta_{\texttt{drop}} \times \texttt{discontinuations}_{i,j,t_{(i,j)}} + \omega_{i,j,t_{(i,j)}}, \end{aligned} \tag{A.3}$$

where the dependent variable is from (A.2), and $\texttt{discontinuations}$ 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, $\texttt{disc}_{i,j,t_{(i,j)}}$ is equal to 1 and the other right-hand-side variables in (A.3) for $t_{(i,j)}$ are 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 (A.3) are in the first column (full sample) of Table A.2. 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

Table A.2: **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 (A.3) using only single announcements. Each coefficient is followed by a 90% bootstrap confidence interval estimated using 1,000 bootstrap samples.

have the expected signs in line with our results for average CAR in Table A.1.

Consider the following example to help interpret the coefficients from the regression. 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 A.2, 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 negative disclosed outcome is higher than a positive outcome.

B.1 Expected Discount Rates

In this section, we present the estimates of discount rates for different stages. 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 as a product of those probabilities, which we denote 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. Similarly, we can define the probability distribution for all other stages of drug development.

Therefore, to estimate the expected discount rate, we 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"). 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 we partition $[0, t]$ into small intervals with 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

Table B.1: **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 are 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.

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 the transition to Phase I clinical trial; (ii) drugs in Phase I clinical trials where a failure is a discontinuation and the success is the transition to Phase II clinical trial; (iii) drugs in Phase II clinical trials, where the failure is discontinuation, and the success is the 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. Using these estimates, we determine the expected discount rates using $\delta = 0.98$, which are in Table B.1.

Table B.2: **Expected Discount Rates, by Indications**

Indications	**Full Sample**	**Middle 90%**
Cardiovascular diseases	0.5895	0.6313
Endocrine diseases	0.5898	0.6306
Gastrointestinal diseases	0.5902	0.6300
Hematological diseases	0.5913	0.6302
Immune disorders	0.5900	0.6310
Infectious diseases	0.5895	0.6313
Inflammatory diseases	0.5902	0.6319
Musculoskeletal diseases	0.5893	0.6310
Neoplastic diseases	0.5903	0.6301
Neurological diseases	0.5908	0.6306
Rare diseases	0.5893	0.6304

Note: Expected discount rates for different stages, estimated using competing risk models, with a yearly discount rate of 0.98. For each indication, we first estimate the distribution of time-to-success from discovery to the market, $\mathbb{P}(\cdot)$. Then using $\mathbb{P}(\cdot)$, determine the expected discount rate $\mathbb{E}(\delta^{\texttt{disc}\rightarrow}) = \mathbb{E}(\delta^{\tau}) \approx \frac{1}{L}\sum_{\ell=1}^{L}\delta^{\tau_\ell}, L = 10,000$, where $\tau_\ell \sim \mathbb{P}(\cdot)$.

C.1 Additional Tables

Table C.1: **Transition Probabilities**

		Probability of Reaching a Stage	
Sample	**Stages**	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. The bottom 95% refers to the drugs developed by firms whose real market capitalization is below 95% of the entire sample.

Table C.2: **Transition Probabilities, by Indications**

Indications	**Full Sample**		**Middle 90%**	
	Discovery to Market	FDA Application to Market	Discovery to Market	FDA Application to Market
Cardiovascular diseases	0.09	0.83	0.10	0.85
Endocrine diseases	0.16	0.92	0.14	0.98
Gastrointestinal diseases	0.13	0.88	0.10	0.89
Hematological diseases	0.19	0.93	0.15	0.99
Immune disorders	0.13	0.93	0.11	0.98
Infectious diseases	0.16	0.87	0.06	0.70
Inflammatory diseases	0.11	0.91	0.08	0.88
Musculoskeletal diseases	0.18	0.92	0.14	0.88
Neoplastic diseases	0.09	0.92	0.06	0.91
Neurological diseases	0.09	0.84	0.12	0.87
Rare diseases	0.21	0.93	0.17	0.95

Note: The unit of observation is a development project, i.e., a specific firm-drug-disease combination, associated with at least one announcement. Here, each indication is associated with almost 1,000 announcements (across all development stages). These indications are not exclusive because some diseases belong to multiple indications. For example, lung cancer is an example of cardiovascular disease and neoplastic disease. In such cases, we use the announcements for drugs developed for lung cancer for both neoplastic diseases and cardiovascular diseases.