Missing Markets for Innovation: Evidence from New Uses for Existing Drugs*

Eric Budish Maya Durvasula Benjamin Roin Heidi Williams

July 2025

Abstract

For large classes of potential inventions, intellectual property rights that are available "on paper" are either not possible or not profitable for firms to enforce. In this paper, we show that these missing incentives yield quantitatively significant underinvestment in research and development. We develop a simple model that formalizes the conditions under which such missing markets for innovation arise. We identify an empirical setting—research into new uses for existing drugs—in which there is sharp variation in the enforceability of intellectual property rights on otherwise comparable inventions over time. We show that when intellectual property rights become unenforceable, research investment and commercialization nearly cease. In doing so, we test two claims central both to our model and the innovation literature more generally—that stronger intellectual property protection does, in fact, induce investment, and that heterogeneity in the availability of these rights distorts investment. The welfare consequences of inadequate incentives in our empirical context are large. Our estimates suggest that 200-800 new uses for existing drugs would have been developed under counterfactual policies. Measures of the value of these uses drawn from existing literature suggest that the social cost of this particular missing market is on the order of several trillion dollars.

*_Budish_: University of Chicago & NBER (eric.budish@chicagobooth.edu); _Durvasula_: Stanford University & Stanford Law School (maya.durvasula@stanford.edu); _Roin_: Independent (bnroin@gmail.com); _Williams_: Dartmouth College & NBER (heidi.lie.williams@dartmouth.edu). We are grateful to Daron Acemoglu, Daniel Fetter, Amy Finkelstein, Matthew Gentzkow, Paul Milgrom, Abhishek Nagaraj, Matt Notowidigdo, Al Roth, Carolyn Stein, Chad Syverson, and seminar participants at MIT Sloan, NBER Productivity lunch, the Radcliffe Institute, Stanford, UC-Berkeley Haas, Northwestern Kellogg, and TNIT for comments, to Yali Friedman, Ben Berger, Amitabh Chandra, and Craig Garthwaite for data, and to Joey Anderson, Ryan Broll, Jeremy Brown, AnneMarie Bryson, Henry Manley, Tamri Matiashvili, Gideon Moore, Lia Petrose, Maya Roy, Mahnum Shahzad, Ralph Skinner, Liyang Sun, and Zahra Thabet for excellent research assistance. Research reported in this publication was supported by the National Institute on Aging (NIA) and the NIH Common Fund, Office of the NIH Director, through Grant U01-AG046708 to the National Bureau of Economic Research (NBER); the content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or NBER. This research was also funded in part by NIA grant P01AG005842, the Alfred P. Sloan Foundation, NSF Grant Number 1151497, the Polsky Center for Entrepreneurship and Innovation at the University of Chicago Booth School of Business, the Hultquist Faculty Research Endowment, the Knight Hennessy Scholars Program at Stanford University, the John M. Olin Program in Law and Economics at Stanford Law School, and the National Science Foundation through grant DGE-1656518 (Durvasula) to Stanford University. While writing this paper, Williams was a visiting scholar and (was subsequently and is currently) a part-time consultant at the Congressional Budget Office (CBO); the views expressed in this paper should not be attributed to CBO.

1 Introduction

Most inventions are far more costly to discover and develop than they are to imitate and reproduce. Economists have long recognized that, absent some mechanism that allows inventors to recoup their investments in these processes of research and development, competitive markets will offer insufficient incentives for innovation (Nelson, 1959; Arrow, 1962; Nordhaus, 1969). Policies that embed this intuition are much older. "To promote the Progress of Science and useful Arts," Article I Section 8 of the United States Constitution gives Congress the power to award intellectual property rights—legal grants that provide inventors with a time-limited ability to exclude their competitors from the market.

In this paper, we investigate the idea that, for large classes of socially valuable inventions, intellectual property rights are unenforceable. The implication is that, in many settings, a key policy lever used to induce research and development provides no meaningful incentive for innovation.

Consider three hypothetical inventions: a protocol of diet and exercise that, if properly calibrated, would substantially reduce the risk of developing Alzheimer's disease and increase overall healthy lifespan; a process that would increase the precision of pharmaceutical manufacturing and reduce the risk of large-scale medicine recalls; and a motorcycle helmet that outperforms existing alternatives on metrics of safety and comfort, available for sale on an online platform.[1] For each, an inventor could file for some form of intellectual property protection—likely, a patent from an entity like the U.S. Patent and Trademark Office.[2] Here, the problem of interest to this paper arises.

Most forms of intellectual property rights, including patents, are not self-enforcing. An inventor must be able both to identify instances of infringement and to take effective legal action to stop it. But when it is difficult to determine, say, that someone is combining food and physical activity in a way that infringes on a patent, or that a bottle of pills was manufactured using a protected method, inventors will be unable to enforce their patents simply because they cannot observe infringement. And even when the inventor can observe, for instance, that dozens of small, overseas sellers are offering motorcycle helmets that infringe her rights, a patent is of little use if she cannot track down

[1] We choose, for the sake of illustration, three types of socially valuable inventions—methods primarily used by consumers, methods primarily used by firms, and easily imitable products. For a discussion of the value of evidence-based protocols for diet and exercise, see Mozaffarian et al. (2011); for a discussion of the difficulty of establishing such an evidence base, see Oster (2024). For a discussion of the social costs of low-quality methods of pharmaceutical manufacturing, see Price (2014), who notes that M&M chocolate candies are made with more sophisticated manufacturing techniques than many prescription drugs sold in the United States; for an analysis of the welfare consequences of low pharmaceutical manufacturing quality—which can trigger contamination events and large-scale recalls—see Yurukoglu et al. (2017). For an example of the consequences of easy imitation of critical protective products like motorcycle helmets, see https://www.congress.gov/117/meeting/house/114092/documents/HMKP-117-JU00-20210929-SD217.pdf.

None of these products are amenable to protection with trade secrecy, as each must be publicly disclosed to end users and regulators as a condition of sale. In Section 2, we develop a framework that identifies instances in which trade secrecy is likely to be an effective incentive for innovation. Mansfield (1986) and Moser (2005) provide empirical evidence on the practical importance of secrecy.

[2] We use the term "invention" in the same manner as the literature on economic growth, to refer to any instance of knowledge production, whether in the form of physical products, sets of instructions, or pieces of information (Arrow, 1962; Romer, 1994; Jones, 2005). It may surprise some readers that all three inventions are likely eligible for patent protection. In Section 2, we provide a primer on the standards for patentability in the United States and clarify that, in general, inventions that take the form of physical products or sets of instructions are patent-eligible.

a party against whom to file suit.

This paper argues that these legal and practical gaps in intellectual property protection are widespread and shows that their social costs are significant. We develop a simple theoretical framework that uses the characteristics of inventions to identify settings in which gaps in intellectual property protection translate into missing private incentives for investment. We identify one such setting—research into new therapeutic uses for existing drugs—where the enforceability of intellectual property rights on otherwise comparable inventions changes sharply over time and hence allows us to test the link between intellectual property rights and private incentives central to this paper. When intellectual property rights become unenforceable, we show that research and development essentially cease.

We begin by formalizing the heterogeneity that generates these differences in enforceability across inventions. We introduce a model in the style of Nordhaus (1969), in which intellectual property rights confer some period of time under monopoly. To the standard framework, we add three complications. First, some inventions can be profitably commercialized while remaining secret. In these cases, firms can recoup their investments in research and development without formal grants of intellectual property—so long as they can maintain trade secrecy. Second, only some inventions are eligible for formal intellectual property protection. By law, others are not. Third, of those inventions that are eligible for intellectual property rights, only some can be protected by rights that are enforceable in practice. Enforcement requires a firm to identify each instance of imitation and successfully bring suit against every infringer. When imitation is challenging or impossible to observe, or when the parties on whom liability falls are not attractive defendants, intellectual property rights that exist "on paper" offer no meaningful protection.

We say that a *missing market* arises when formal intellectual property rights are necessary to incentivize research and development, but such rights are either unavailable or unenforceable. Note that this characterization rules out a standard policy solution to the problem of inadequate incentives for innovation: canonical models linking intellectual property rights to incentives embed the idea that sufficiently strong property rights—that is, longer or broader periods of market exclusivity—can make any invention profitable to develop (e.g., Nordhaus, 1969; Gilbert and Shapiro, 1990; Klemperer, 1990). When gaps in incentives arise because intellectual property rights are unavailable or unenforceable, there are no "Nordhaus parameters" that can induce private investment.

This model allows us to identify large classes of (potential) inventions, across sectors of the economy, where missing markets are likely to arise. Our empirical analysis focuses, however, on a specific context—investments into new uses for existing drugs—where sharp variation in the enforceability of intellectual property rights over time allows us to test for evidence of a missing market.

To fix ideas about the institutional features that enable this test, an example is helpful. In 1995, the U.S. Food and Drug Administration (FDA) approved Glucophage (metformin hydrochloride) as a treatment for diabetes.[3] Generic versions of the drug entered the market in 2002. Around

[3]NDA #020357, approved on 03 March 1995.

this time, oncologists began to document tentative evidence that metformin may also be effective at treating or preventing cancer. Though these researchers proceeded with various "pre-clinical" studies, commercial interest in funding a clinical trial validating metformin as a treatment for human cancer was nonexistent (Gale and Gerlin, 2012). The problem? There is no mechanism that would allow a private firm to *sell* the knowledge that this existing drug could be used in a new way.[4]

To be sure, the USPTO would, likely, have granted patents on methods of using metformin to treat cancer, and the FDA would have barred other firms from marketing generics for these uses.[5] But once the information about metformin's value as a treatment for cancer was known, any physician could make use of it freely—by, for example, prescribing the generic version to their cancer patient—without compensating the firm that undertook the expensive clinical trial. Even if the innovating firm could detect unauthorized use, enforcement would be practically impossible: the only parties on whom liability would definitively fall would be the patients who would use the drug to treat their own cancer.[6] Decades after preliminary evidence first emerged, whether metformin has potential as a treatment or even potentially as a preventative for cancer is an open question that "still warrant[s] well-designed clinical investigation" (Lord and Harris, 2023).

In most missing markets identified by our model, intellectual property rights are unenforceable from the moment that they are granted. Observe, however, that patents on new uses for metformin become unenforceable many years after commercialization, only when generic versions of the drug are approved for sale. This delay—a feature of the regulatory structure for pharmaceuticals in the United States—provides our research design. For a period in each new drug's life, the FDA bars the entry of generic competitors. During this time, a single firm controls the market and can set a price for its drug that reflects all commercialized uses. So long as the firm retains market exclusivity, it is compensated for any "use" of these new uses. Patents on new uses are, thus, perfectly enforceable. But after generic entry, as the metformin case illustrates, the likelihood that the firm can enforce such a patent drops essentially to zero.

Mapping this variation to our framework yields three predictions. First, firms will have the strongest incentives to invest in new uses for existing drugs soon after initial approval. Second, as market exclusivity expiration approaches, the number of years in which investments in new uses will yield profits declines—and thus, incentives should fall. Third, exclusivity expiration creates a missing market for new uses, and private research and development should cease.

Testing these predictions is the focus of our empirical work. To do so, we construct new data tracing out research and development investments associated with all new drugs approved by the FDA between 1985 and 2014. We collect records of re-approvals for new uses, as well as indicators

[4] Gale and Gerlin (2012) includes portions of an interview with one oncologist, who indicated that she was ready to begin work on this lead around 2002, but was unable to garner support from any private firm, noting that "*When they realized the results wouldn't be available until they lost their patent, they pulled out.*"

[5] Section 3.2 provides details on the availability of, both, patents on methods of using existing drugs for new therapeutic indications and FDA-administered regulatory exclusivities that serve the same purpose.

[6] Section 3.2 elaborates on the difficulties of enforcement in this context. In a small number of cases, hospitals, physicians, and pharmacists may share in the liability—though suits against them for patent infringement are rare: it is rarely advantageous for firms to sue their own (repeat) customers.

of upstream investments in basic research and early-stage commercialization. To characterize trends in investment around market exclusivity expiration, however, we must first define some measure that captures the point in time when the firm expects its intellectual property rights to lapse and, hence, its drug to be vulnerable to generic competition. We draw on legal and regulatory analyses, as well as newly-digitized records of the intellectual property rights associated with each drug in our data, to propose and validate such a measure of expected exclusivity expiration.[7]

Using these data, we show that investments in new uses for existing drugs closely track our predictions. Ten years prior to market exclusivity expiration, roughly ten percent of drugs in our sample are re-approved for at least one new use. The likelihood of re-approval declines monotonically in each subsequent year and falls to roughly zero following market exclusivity expiration. Nearly 95 percent of new uses ever approved by the FDA for our sample of drugs were approved prior to exclusivity expiration. We document comparable trends in clinical trials, which decline sharply in frequency over a drug's life cycle. Together, these pieces of evidence are consistent with—though not dispositive of—the existence of a missing market for new uses following generic entry.

There are two natural alternative explanations. If the set of "scientific opportunities" for new uses is gradually exhausted, research investments may naturally decline over a drug's lifecycle. Alternatively, firms may recognize that generic entry is imminent and strategically sequence research and development to bring all viable new uses to market during this initial profitable period. On the basis of two additional empirical exercises—one documenting that longer periods of market exclusivity induce additional commercialization of new uses, and one showing that commercialization of related products on which intellectual property rights remain enforceable does not drop off at the same time—we argue that both alternative explanations are unlikely.

First, we provide causal evidence that increasing the duration of market exclusivity increases the number of new uses that are commercialized. This test relies on the observation that firms in the pharmaceutical sector have strong incentives to file key patents on new drugs at the time of invention, often many years before drugs are approved for sale. Sizable delays between invention and approval—due to required clinical trials and regulatory delays—mean that certain drugs have many more years of effective market exclusivity than do others, for reasons unrelated to their potential for new uses (Roin, 2013; Budish et al., 2015; Gilchrist, 2016). We use this quasi-random variation in the market exclusivity provided to drugs in our sample to show that drugs with longer periods of exclusivity have more commercialized new uses, consistent with the idea that when opportunities to develop new uses arise stochastically, only those firms with sufficient remaining market exclusivity can pursue them.

Second, we establish that the enforceability of intellectual property rights is the mechanism driving our results. Not all re-approvals of existing drugs face the same enforceability problem as new uses once generic entry occurs. We show that although the likelihood that a drug in our sample is approved for a new use falls essentially to zero when market exclusivity expires, the

[7] A separate paper, Durvasula et al. (2023), serves as a "user's guide" to the datasets and measurement strategies employed in this paper's empirical analyses; in particular, it introduces and validates the contents of our primary dataset reporting patent protection and assesses alternative measures of market exclusivity.

likelihood that it is re-approved for other types of uses—such as a new patient population, dosage, or strength—declines far less sharply. Moreover, the likelihood that it is re-approved as part of a "fixed dose combination" actually increases. Fixed dose combinations bundle multiple drugs into a single "dosage form"—for example, metformin combined with glyburide in the branded drug Glucovance. Such combinations are regulated as distinct new drugs and, as such, receive their own regulatory protections that block generic entry. Continued investment into fixed-dose combinations after the innovator drug's exclusivity expires provides evidence in support of two conclusions. First, it demonstrates that scientific opportunities for innovation do not coincidentally vanish at this point in a drug's lifecycle. Second, it confirms that our measure of market exclusivity accurately reflects firms' expectations about enforceability.[8] Taken together, this body of evidence provides support for the idea that there are scientifically feasible new uses for existing drugs that are not profitable to develop given the structure of existing incentives.

Our empirical setting allows us not only to establish the existence of this missing market, but also to quantify its social costs. We estimate that counterfactual periods of market exclusivity—selected to mirror policies that would provide fully enforceable intellectual property rights to new uses—would yield between 0.1 and 0.4 additional new uses per drug.[9] As roughly 2,000 drugs that satisfy our sample criteria have been approved since the modern FDA was created in 1962, this implies between 200 and 800 missing new uses. Valuing each missing new use at $1 billion per year—a rough estimate of social value drawn from existing economic analyses of the social value of new drugs—indicates that the net present value of missing research and development in this market is on the order of $2 to $10 trillion.[10]

Although our back-of-the-envelope estimates of the social cost of missing new uses are rough, they are consistent with evidence offered by the scientific literature. Scientists characterize the relatively small number of chemicals that are ever approved as drugs by the FDA as "promiscuous," by which they mean that a drug that is sufficiently effective at treating *some* human disease to clear the threshold for FDA approval is very likely to be effective in the treatment of *many* diseases (Wermuth, 2006). Dozens of studies have validated this prediction. Gelijns et al. (1998), for example, document that within a sample of "blockbuster" drugs sold in 1993, nearly 90 percent had important uses beyond those for which they were originally approved. Against these estimates, our estimated range of 0.1 to 0.4 new uses per drug appears, if anything, conservative.

[8] Our framework and this empirical test hint at one perverse consequence of gaps in the enforceability of certain intellectual property rights. When patents are not enforceable, firms have strong incentives to seek out alternative mechanisms that allow them to exclude their competitors from the market. The example of fixed-dose combinations highlights one such option: firms might respond to the effective expiration of intellectual property rights on one drug by developing related products on which intellectual property rights remain enforceable—a practice pejoratively referred to as product hopping (Carrier and Shadowen, 2016). We do not undertake a complete welfare analysis and thus do not weigh in on the social costs of these practices, but emphasize that one cost of incomplete or unenforceable patent rights not captured by our analysis is any deadweight loss associated with "self help" solutions.

[9] Sections 6 and 7 clarify that while this exercise is useful to estimate the volume of missing innovation in this setting, this style of blunt patent term extension is unattractive as a policy solution to the missing markets problem.

[10] In Section 6, we collect various pieces of evidence on the social value of new uses relative to the initial approvals used to form this estimate and construct a set of scaling factors that might be applied to this estimate. Even if new uses are worth a fraction of initial approvals, our estimates imply that the social cost of this particular missing market is large.

The idea that insufficient private incentives exist to develop new uses for existing drugs has been a long-standing concern for legal scholars, scientists, and policymakers (Gelijns et al., 1998; Mossinghoff, 1999; Eisenberg, 2005; US Institute of Medicine, 2008; Rai, 2012; Roin, 2014). This paper is, to the best of our knowledge, the first to formalize the source of missing incentives in this setting theoretically and to document underinvestment empirically. We confirm that the scale of underinvestment is extraordinarily large. As advances in health data infrastructure and screening technology reduce the cost of discovering potential new uses, the problem may only be growing (Collins, 2011b,a, 2012; Collins and Varmus, 2015).

Evidence of a quantitatively significant missing market for new uses provides, as well, clear empirical support for the model of intellectual property and innovation incentives that we introduce. Our framework embeds and formalizes two ideas that have appeared in various forms and under many names across much of the legal and economic scholarship on incentives for innovation—that formal intellectual property rights are not necessary to incentivize the development of certain inventions (e.g., Mansfield et al., 1981; Teece, 1986; Mansfield, 1986; Levin et al., 1987; Cohen et al., 2000; Moser, 2005), and that formal intellectual property rights cannot incentivize the development of others (e.g., Nelson, 1959; Arrow, 1962; Teece, 1986; Lemley and Shapiro, 2005; Eisenberg, 2005; Abramowicz, 2007; Kapczynski and Syed, 2012; Rai, 2012; Budish et al., 2015). In theoretical and empirical studies that aim to inform the design of optimal patent policy, however, every invention is assumed patentable (and patent-protected), and every patent perfectly enforceable (e.g., Machlup, 1958; Nordhaus, 1969; Scherer, 1972; Nordhaus, 1972; Klemperer, 1990; Gilbert and Shapiro, 1990; Scotchmer, 1991; Budish et al., 2015). Our modifications to the standard Nordhaus (1969)-style framework, though simple, capture the importance of legally-protected secrecy as a complement to patent protection. In particular, we formalize and test an idea central to the legal analysis in Kapczynski and Syed (2012)—that there are "domain[s] of innovation, that patents, whatever their scope, cannot adequately address."

The theoretical predictions generated by this model hinge on a causal link between intellectual property protection and incentives for innovation. Although decades of theory and centuries of policymaking have assumed such a relationship—that longer and broader intellectual property rights induce investment—there is remarkably little direct empirical support, owing to the difficulty of identifying meaningful variation in the protection provided to otherwise comparable inventions. A small number of papers—notably, Lerner (2002) and Sakakibara and Branstetter (2001)—have leveraged changes in national patent laws to test for such a relationship, but have found no effect, likely because country-level changes have relatively small impacts on global markets for innovation. Budish et al. (2015) document evidence consistent with the idea that longer patent terms induce additional research investments, but are unable to rule out alternative explanations. The model we introduce in this paper implies an alternative research design, which allows us to test for direct evidence that intellectual property rights affect the level of research and development by leveraging variation in the *enforceability* of intellectual property rights across inventions. The missing markets problem, here, creates variation in the effective patent terms (years before generic entry) provided to

otherwise comparable investments (new uses). We find that increases in patent term cause increases in the number of new uses brought to market and—with the caveat that this estimate is specific to our empirical context—estimate the elasticity of research and development with respect to patent term.

We contribute more broadly to closely related literatures that study how property rights affect incentives for investment (e.g., Moser, 2005; Budish et al., 2015; Hodgson, 2024; Dix and Lensman, 2024) and how under-provision of incentives yields missing research and development (Budish et al., 2015; Krieger et al., 2022; Michelman and Msall, 2024; Dix and Lensman, 2024). Our model includes, as a special case, the problem of interest in Budish et al. (2015), who study incentives for the development of cancer treatments. Our model also includes the problem of interest in Dix and Lensman (2024), who investigate incentives for combining existing cancer drugs into new treatment regimens. We view the two papers as highly complementary. While our paper uses a theoretical framework to clarify the missing markets problem and uses reduced form empirical analysis, including quasi-random variation in market exclusivity, to quantify missing innovation, their paper builds a dynamic structural model that enables this quantification. Their structural model, importantly, allows for the evaluation of stylized counterfactual policies. Our paper also offers a clarification on the problem studied in Dix and Lensman: missing incentives for combinations of complementary products, in that setting, are not the result of property rights that are *unavailable* ($p_{\text{eligible}} = 0$ in our model) but rather property rights that are *unenforceable* because of generic drug availability ($p_{\text{enforce}} = 0$). Concretely, therapies combining existing generic drugs are patentable,[11] but any rights would offer little incentive for innovation. Finally, our data and research design allow us to provide the first empirical evidence in support of an assumption central to Dix and Lensman's structural model, and to the structural literature on innovation more broadly: there is a causal link between intellectual property rights and research investment. While such a link is often assumed, causal evidence has been surprisingly elusive.

This paper proceeds as follows. Section 2 introduces our conceptual framework. Section 3 provides context on the problem of new uses. Section 4 describes our data and measurement. Section 5 presents three tests that, together, establish the existence of a missing market for new uses for existing drugs. In Section 6, we present evidence on the magnitude of missing investment—in terms of missing "uses" and dollars of social value. In Section 7, we discuss the open question of how missing markets ought to be "fixed" if standard levers of intellectual property policy are unavailable, before concluding.

2 Characterizing Missing Markets for Innovation

Not all inventions can be protected from imitation with grants of intellectual property rights—either because these rights are not available by law or because available rights are not enforceable in

[11] See Section 2.2 for details.

practice.[12] For certain inventions, these gaps in protection have no impact on incentives for research and development. For others, incomplete intellectual property rights yield *missing markets* for innovation.

This Section develops, in two parts, a simple framework that formalizes this heterogeneity. The first part introduces a set of complications to a standard Nordhaus (1969)–style model and, in doing so, allows us to define the conditions that give rise to a missing market. The second part transforms this model into a classification exercise, which uses characteristics of inventions to identify contexts where missing markets are likely to arise.

2.1 Defining Missing Markets

2.1.1 Preliminaries

A representative firm must decide whether to commercialize a potential invention, the characteristics of which it regards as exogenous.[13] Commercialization refers to the (costly) process of bringing a product or idea to market—running clinical trials to satisfy regulatory requirements, expanding manufacturing capacity to allow for production at scale, etc.

The starting point for this model is a Nordhaus (1969)–style framework, as adapted by Budish et al. (2015). A potential invention is characterized by the following sets of parameters, the first six of which are common to prior work.

- **Timing parameters**: The firm must decide whether to commercialize the potential invention in year t_{invent}, which we normalize to zero. Commercialization involves a deterministic delay t_{comm}.
- **Cost of commercialization**: To commercialize the potential invention, the firm incurs—in net present value terms at t_{invent}—a cost c.
- **Likelihood of successful commercialization**: Commercialization may not be successful.[14] We denote the probability of success by p_{comm}. We restrict consideration to potential inventions for which p_{comm} is greater than zero.
- **Obsolescence risk**: If an invention is successfully commercialized, it may lose usefulness over time as new, substitutable alternatives are brought to market. This is captured with an exogenous risk of obsolescence $1-\gamma$ in each year following t_{comm}.

[12] As the introduction suggests, there are two categories of intellectual property rights of interest in this paper: grants that confer monopoly or monopoly-like rights—patents, copyrights, contract-law forms of intellectual property—which must be actively enforced; and trade secrets, which provide a default protection to inventions if the inventor took active steps to avoid disclosure. When we refer to formal grants of intellectual property rights that must be enforced, we are referring to the first category. There are other forms of intellectual property (e.g., trademarks) that neither provide monopoly rights nor have as their goal the provision of incentives for innovation.

[13] We use the term "invention" to refer to any instance of knowledge production—including in the form of physical products, sets of instructions, and pieces of information (Arrow, 1962; Romer, 1994; Jones, 2005).

[14] In our empirical context, for example, clinical trials may fail to enroll enough patients or may indicate that a drug is ineffective as a treatment for a particular disease.

- **Discount factor**: The firm discounts future profits by a factor δ.
- **Monopoly profits and social value**: A commercialized invention that is not obsolete and produced by the firm as a monopolist yields profits π per year. The invention also yields social value v^{monop} when priced by the monopolist and v when priced by the social planner, where $v > v^{\mathsf{monop}}$.

To this standard framework, we add three sets of parameters.

- **Secrecy**: Some inventions can be profitably commercialized while remaining secret. That is, even without a formal grant of intellectual property, the risk of imitation is low. The parameter $p_{\mathsf{secret}} \in [0,1]$ captures the likelihood that the invention remains secret—and, thus, profitable—upon commercialization.[15] After commercialization, the firm risks losing secrecy. We capture this by assuming the firm faces an exogenous risk $1-\eta$ of losing secrecy each year following commercialization.
- **Eligibility**: Not all inventions are eligible for protection with grants of intellectual property rights. The probability that a commercialized invention satisfies these eligibility requirements is given by $p_{\mathsf{eligible}} \in [0,1]$. The associated protection lasts for t_{IP} years.
- **Enforceability**: Grants of intellectual property rights are not self-enforcing. To enforce these rights, the firm must seek out instances of imitation and successfully bring suit against each infringer. The probability that the firm can successfully enforce its intellectual property rights is $p_{\mathsf{enforce}} \in [0,1]$.

In Section 2.2, we discuss factors that affect the value of each parameter.

2.1.2 Monopoly Duration

For each potential invention, there is some period in which the firm will, in expectation, earn monopoly profits. This *expected time under monopoly* T depends on the mechanism used to deter imitation: secrecy or formal intellectual property rights.

If the invention is protected by secrecy, the expected time under monopoly is given by

$$T_{\mathsf{secrecy}} = p_{\mathsf{comm}} \cdot p_{\mathsf{secret}} \cdot \sum_{t_{\mathsf{comm}}}^{t_{\mathsf{secret}}} (\delta\gamma\eta)^t, \tag{1}$$

where t_{secret} goes to infinity. In words, time under secrecy T_{secrecy} is determined by (i) the likelihood of successful commercialization p_{comm}, (ii) the likelihood that secrecy effectively deters imitation p_{secret}, and (iii) the number of years in which secrecy is maintained, adjusted for discounting (by δ), the risk of obsolescence (by γ), and the risk of losing secrecy (by η).

[15] One could, instead, interpret $1 - p_{\mathsf{secret}}$ as the fraction of monopoly profits lost to imitation in the absence of a formal intellectual property grant.

If, instead, the invention is protected by a formal grant of intellectual property (e.g., a patent), the expected time under monopoly is given by

$$T_{\mathsf{IP}} = p_{\mathsf{comm}} \cdot p_{\mathsf{eligible}} \cdot p_{\mathsf{enforce}} \cdot \sum_{t_{\mathsf{comm}}}^{t_{\mathsf{IP}}} (\delta\gamma)^t. \tag{2}$$

Here, the factors p_{eligible} and p_{enforce} capture the likelihood that the invention is effectively protected by a patent, and the term t_{patent} captures the duration of protection.

2.1.3 Incentives to Invest

A profit-maximizing firm will commercialize a potential invention if and only if the expected profits exceed the costs of commercialization. For time under monopoly

$$T^* = \max(T_{\mathsf{secrecy}}, T_{\mathsf{IP}}), \tag{3}$$

a firm will proceed if and only if

$$T^* \cdot \pi - c \geq 0. \tag{4}$$

If the invention can be more profitably commercialized under trade secrecy than under with an intellectual property grant —that is, if

$$p_{\mathsf{comm}} \cdot p_{\mathsf{secret}} \cdot \sum_{t_{\mathsf{comm}}}^{t_{\mathsf{secret}}} (\delta\gamma\eta)^t \geq p_{\mathsf{comm}} \cdot p_{\mathsf{eligible}} \cdot p_{\mathsf{enforce}} \cdot \sum_{t_{\mathsf{comm}}}^{t_{\mathsf{IP}}} (\delta\gamma)^t, \tag{5}$$

then the firm chooses secrecy. Otherwise, the firm chooses the intellectual property grant. T is, then, determined by the availability and enforceability of intellectual property rights, as in (2).

2.1.4 Missing Markets

Definition 1 (Missing Market)**.** A missing market arises when an invention's expected time under monopoly T^* is zero.

With (1)–(5), observe that we can write the expected time under monopoly T^* for all potential inventions as

$$T^* = \max\left(p_{\mathsf{comm}} \cdot p_{\mathsf{secret}} \cdot \sum_{t_{\mathsf{comm}}}^{t_{\mathsf{secret}}} (\delta\gamma\eta)^t, p_{\mathsf{comm}} \cdot p_{\mathsf{eligible}} \cdot p_{\mathsf{enforce}} \cdot \sum_{t_{\mathsf{comm}}}^{t_{\mathsf{patent}}} (\delta\gamma)^t \right). \tag{6}$$

Under what conditions is $T^* = 0$?

First, $T^* = 0$ when the invention's patent term is shorter than the time required to commercialize the invention—if $p_{\mathsf{secret}} = 0$ and $t_{\mathsf{patent}} < t_{\mathsf{comm}}$.[16]

[16]This case is the focus of Budish et al. (2015), who collect evidence consistent with the idea that these missing incentives generate a quantitatively important distortion in incentives for research and development, using the context of cancer treatments and preventives as an example.

In this paper, our focus is on a second case, which arises when $p_{\text{secret}} = 0$ and *either* $p_{\text{eligible}} = 0$ *or* $p_{\text{enforce}} = 0$. In words, if an invention cannot be protected via secrecy ($p_{\text{secret}} = 0$) and is either not eligible for intellectual property protection ($p_{\text{eligible}} = 0$) or available intellectual property rights are unenforceable ($p_{\text{enforce}} = 0$), then the firm's expected time under monopoly is zero.

Observe that the first case reflects, in effect, a misalignment of timing parameters. The gap in incentives that generates a missing market there can be corrected by policy changes that restructure intellectual property rights—for example, by altering the time required for commercialization t_{comm} (e.g., by offering accelerated approval paths), the time provided by a grant of intellectual property t_{IP} (e.g., by extending patent terms), or the relationship between the invention's timing parameters (t_{invent}, t_{comm}, t_{IP}) (e.g., by starting patent terms at commercialization instead of invention). These shifts—a focus of Budish et al. (2015)—can yield a non-zero T^* and, thus, restore incentives for innovation.

But in the second case, our focus, these types of policy solutions are unavailable. Because intellectual property rights are either unavailable or unenforceable, standard innovation policy tools that rely on tweaks to the structure of property rights are inapplicable.

Proposition 2. *If $p_{\text{secret}} = 0$ and either $p_{\text{eligible}} = 0$ or $p_{\text{enforce}} = 0$, then neither longer ($t'_{\text{IP}} > t_{\text{IP}}$) nor broader ($\pi' > \pi$) patent terms would make commercialization profitable.*

Theoretical studies of innovation and intellectual property rights often characterize optimal policy in terms of two "Nordhaus parameters"—t_{IP} (capturing patent duration) and π (capturing patent breadth).[17] Implicit in these analyses is the idea that careful calibration of these dimensions of intellectual property rights can, in fact, strike an optimal balance between incentives for innovation and deadweight loss. But when intellectual property rights are unavailable or unenforceable changes to these "Nordhaus parameters"—t_{IP} and π—have no impact on the firm's investment decision in (4). When an invention can be profitably commercialized under secrecy, as in (5), neither of these parameters enters the firm's decision problem.

2.2 Identifying Missing Markets

In Section 2.1, we introduce three parameters—$p_{\text{secret}}, p_{\text{eligible}}, p_{\text{enforce}}$—that affect both whether an invention is commercially viable and whether standard levers of innovation policy can be used to calibrate incentives for innovation. This section provides context on each parameter.

2.2.1 p_{secret} : Can the invention be profitably commercialized while remaining secret?

For many inventions, both private and social value come entirely from public disclosure. Consider inventions that are, in effect, "recipes" (Jones, 2023) or sets of directions—such as regimens of diet

[17] Although patent breadth is typically included in these optimal patent term analyses, it is not obvious how it ought to be modeled in a Nordhaus (1969)-style framework. Gilbert and Shapiro (1990), for one, define the breadth of a patent as the flow rate of profit available to the firm so long as the patent is in force—that is, as the profit parameter π. We do the same.

and exercise (Section 1), instructions for the use of an existing drug in the treatment of a new disease (Section 3), or protocols for the use of checklists to reduce the risk of medical error (Gawande, 2009; Kapczynski and Syed, 2012). Each of these inventions consists entirely of the knowledge that some series of steps will yield a particular outcome. Under some circumstances, firms can implement restrictions that limit how widely this information will diffuse—for example by binding recipients contractually with non-disclosure and non-compete agreements (Singh and Marx, 2011; Johnson et al., 2023; Reinmuth and Rockall, 2023). Without these types of contractual restraints, however, secrecy is not viable.

For other types of inventions, secrecy is impractical. If an invention must be placed on display to facilitate sale, it cannot be kept secret. Even if disclosure of an invention is not immediate upon sale, it may be necessary to persuade a potential customer to make a purchase. If, say, a physician determining whether to adopt a new diagnostic test requires detailed assurances about the nature of the technology, secrecy may be destroyed in the process of sale (Arrow, 1962).[18] When there are no circumstances under which the invention can be commercialized while remaining secret, p_{secret} takes a value of zero.

Existing work has identified some settings in which p_{secret} is clearly positive (e.g., Levin et al., 1987; Cohen et al., 2000; Moser, 2005, 2007).[19] Internal manufacturing processes—including proprietary recipes, like the formula for Coca-Cola (Carr and Choi, 2023)—can be held as secrets for as long as the firm can limit their public disclosure.[20] Other inventions may be amenable to secrecy simply because they are complex or otherwise difficult to reverse engineer. Vaccines are a classic example. Even when the chemical makeup of a vaccine is known, the technical know-how required to produce a product that is stable and safe for use in human patients poses a substantial barrier to replication (Price et al., 2020).[21] Outside of markets for health technologies, engineered

[18] This observation in Arrow (1962) is often described as Arrow's Information Paradox.

[19] In a now classic survey, Mansfield (1986) posed questions to a randomly drawn set of 100 firms from twelve industries in the United States, intended to understand the extent to which patent protection—as opposed to secrecy—was seen as vital to the development of new inventions. In particular, the survey asked respondents to estimate the share of its recent inventions that would have been developed had patent protection not been available. For firms in the textile, rubber, motor vehicle, and office equipment industries, where key processes and products could be held as secrets, respondents indicated that the absence of patent protection would have made no difference. Consistent with our emphasis in this paper, however, firms developing products with a close nexus to human health—in the pharmaceutical industry—suggested that the majority of their inventions would not have been developed. Moser (2005) provides complementary evidence, that inventors in the 19th century in countries without patent laws disproportionately concentrated in industries where secrecy was thought to be effective. Moser (2007) provides a test of this idea by leveraging a breakthrough—the publication of the periodic table in 1896—that made it easier for competitors to reverse engineer chemical innovations. When the cost of reverse engineering dropped—in our framework, when p_{secret} fell close to zero—inventors' propensity to patent increased.

[20] In the case of the Coca-Cola formula, efforts to preserve the secrecy are wide-ranging. The formula itself is locked in a vault, which visitors are invited to see from the outside: https://www.worldofcoca-cola.com/explore-inside/explore-vault-secret-formula. Only a small number of employees have access to the formula, each of whom is bound by a set of non-compete and non-disclosure agreements. For all other employees, the company tracks their movements within its facilities and monitors their use of internal servers: https://www.bloomberg.com/news/features/2023-05-11/the-plot-to-steal-the-secret-coke-can-liner-formula.

[21] In 2021, it was widely reported that a team of scientists at Stanford had reverse engineered the Moderna COVID-19 vaccine, using a sample of vaccine collected from a garbage can. As the scientists themselves clarified, however, they had not actually reverse engineered the vaccine; they had determined, only, its chemical sequence. This information alone was not sufficient to enable imitation or reproduction without additional reverse-engineering of the methods used to

secrecy is quite common. Every aspect of the "fuzzy logic" used to modulate cooking time and temperature within the award-winning Zojirushi rice cooker has been held as a trade secret since its introduction in the 1980s: the key piece of technology is embedded within the product in a way that ensures that disassembly will destroy the relevant piece (Santo, 2023).

When secrecy is a commercially viable option, state and federal laws in the United States provide formal trade secret protection.[22] As far as legally enforceable intellectual property rights go, trade secrecy is both expansive and fragile. Protection lasts until secrecy is destroyed, which may mean that protection is indefinite. But any instance of disclosure, reverse-engineering, or independent discovery is secrecy-destroying. So long as the owner takes "reasonable efforts" to maintain secrecy, these laws provide recourse in instances of misappropriation—but cannot restore an already-disclosed secret.

2.2.2 p_{eligible} : Is the invention eligible for a legally-enforceable intellectual property grant?

In the United States, government-granted intellectual property rights take many forms—including patents (utility, design, plant), copyrights, and regulatory exclusivities. Whether any invention can be protected by intellectual property rights depends both on the subject matter of the invention and on the invention's specific characteristics. Certain inventions fall within categories of subject matter excluded from specific forms of protection (e.g., most forms of DNA are ineligible for patent protection), and some inventions may be ineligible for all forms of government-administered intellectual property rights (e.g., most databases, which cannot be protected by copyrights or patents).

Patent law, in particular, draws strict boundaries around eligible subject matter. Utility patents may not be granted on inventions that claim abstract ideas, natural laws, or physical phenomena. Judicial decisions in recent years—in particular, *Mayo v. Prometheus*, *Association for Molecular Pathology v. Myriad Genetics*, and *Alice Corp. v. CLS Bank*—have clarified these restrictions and, in doing so, substantially narrowed the set of patentable potential inventions.[23] At present, inventions determined (by the United States Patent and Trademark Office or a court) to claim abstract ideas or natural phenomena without adding an additional "inventive step"—such as sequences of isolated genomic DNA,[24] methods for processing loan information through a clearinghouse,[25] and computer programs that automatically categorize digital images[26]—are ineligible for patent protection.

---

produce the commercially-available version of the product. For a longer discussion, see https://www.nytimes.com/interactive/2021/11/03/health/mrna-vaccine-production.html.

[22] For example, Uniform Trade Secrets Act (with 1985 Amendments), 14 U.L.A. 437 (2005); Cal. Civ. Code §§ 3426–3426.11; Defend Trade Secrets Act of 2016, 18 U.S.C. §§ 1836–1839.

[23] See Mayo v. Prometheus, 566 U.S. 66 (2012) (holding that a claimed method optimizing drug dosages based on natural correlations is not patent eligible because it recites a law of nature without adding an inventive concept beyond routine practices); Association for Molecular Pathology v. Myriad Genetics, 569 U.S. 576 (2013) (holding that naturally occurring DNA sequences—even when isolated—are not patent eligible, though cDNA may be eligible because it is not naturally occurring); and Alice Corp. v. CLS Bank, 573 U.S. 208 (2014) (holding that claims directed to abstract ideas implemented on generic computer technology are not patent eligible absent an inventive concept).

[24] Association for Molecular Pathology v. Myriad Genetics, 569 U.S. 576 (2013).

[25] Buysafe, Inc. v. Google, Inc., 765 F.3d 1350 (Fed. Cir. 2014).

[26] Content Extraction and Transmission, LLC v. Wells Fargo Bank, 776 F.3d 1343 (Fed. Cir. 2014).

Even if an invention's subject matter is eligible for protection, it may fail to meet other statutory requirements. Copyright protection, for example, attaches only to works that meet a minimum threshold of originality (U.S. Copyright Act, Title 17), and patents are granted only to inventions that are novel, non-obvious, and useful (35 U.S.C. §§ 101, 102, 103). Roin (2016) observes that these statutory requirements have the effect of generating large classes of potentially valuable health technologies that cannot be protected.

As the introduction suggests, the boundaries of patent-eligible subject matter may not be wholly intuitive. The two hypothetical "method" inventions described there—a program of diet and exercise and a new process for pharmaceutical manufacturing—are both likely to satisfy existing requirements, provided they include some "inventive step" (and, of course, satisfy other requirements of novelty, non-obviousness, and usefulness). Although clearing these hurdles may be challenging for a program of diet and exercise, the U.S. Patent and Trademark Office has granted patents on weight-loss diet regimens that restrict certain food groups and introduce specific supplements,[27] protect cells against chemotherapy,[28] and that employ specific methods of treating carbohydrate-rich foods to reduce their glycemic indices.[29] In the manufacturing example, new processes might be protected by process patents that claim a novel process or by product patents that claim a new piece of equipment.[30]

In our model, p_{eligible} takes a value of zero when an invention falls squarely within a category excluded from protection by law. While there remain many categories of inventions for which patent eligibility is clear—where $p_{\text{eligible}} = 1$—recent shifts in legal standards have introduced considerable uncertainty in both the tests used to determine eligibility and in the validity of previously-issued patents (Hoyt, 2022). For many inventions, then, p_{eligible} lies between zero and one.

2.2.3 p_{enforce} : Can the firm profitably enforce any intellectual property grant?

Even when an invention is eligible for formal grants of intellectual property, not all intellectual property rights are enforceable. Though many factors determine this enforceability, we highlight three: whether the firm can observe instances in which its rights are infringed, whether enforcement against particular defendants is practical and profitable, and whether the underlying intellectual property grant is strong enough to permit litigation.

If a firm cannot observe instances in which its intellectual property rights are infringed, those rights are unenforceable, and p_{enforce} takes a value of zero. Of course, for many inventions, unauthorized uses—like "knockoffs" of luxury handbags and counterfeit consumer electronics (Robbins, 2018)—are readily apparent. Advances in technology, including reverse-image searching, digital watermarking, and algorithmic plagiarism detection, have continued to expand the types of infringement that are easy for inventors to detect (Eckhause, 2022).

[27] U.S. Patent No. 5,855,949 – Dietary system high in oil intake for the treatment of obesity.

[28] U.S. Patent No. 8,865,646 B2 – Dietary compositions and methods for protection against chemotherapy, radiotherapy, oxidative stress, and aging.

[29] U.S. Patent No. 8,568,820 B2 – Method of treating carbohydrate rich foods for reducing their glycemic indices.

[30] See, e.g., U.S. Patent No. 10,106,974 B2 – Flexible manufacturing system.

In many contexts, however, infringement can be extremely difficult to observe. Manufacturing processes crucial to the success of an industry can easily be copied and used internally by competing firms, with no external indication of infringement (Price, 2014). When firms are unable to protect these processes with secrecy, they may be able to find alternative ways of creating observability. Teece (1986) highlights one strategy used by firms licensing methods of petroleum refining to other parties within the petrochemical industry: the sellers included, in their patent-protected processes, a specialized chemical catalyst produced only by the original seller. To use the method, other firms were required to purchase this input controlled by the patent-holding firm, thus deterring unauthorized imitation. Other mechanisms, including intermediaries (e.g., regulators and large institutions), can have a similar effect (Kapczynski and Syed, 2012).

Even if infringement is observable, enforcing intellectual property rights may still not be practical or profitable. One difficulty arises from legal rules that specify who, technically, infringes certain patents. In the United States, patents on methods—specific ways of using a product or carrying out a set of instructions—are directly infringed only by the individual or entity performing each step of the method. To understand what this means in practice, consider the problem at the center of *Ormco Corp. v. Align Technology*.[31] Ormco held a patent on a method for using orthodontic retainers to straighten teeth. Align sold competing retainers, to be used in the same way. Although Align was, technically, copying Ormco, its action did not directly infringe Ormco's patent. Rather, infringement occurred each time a patient used the retainer to straighten her teeth—thus *performing* the method. To enforce its patents, Ormco would have been required to file suit against each retainer-wearing defendant. Unlike in our empirical context, Ormco would not have necessarily been suing its *own* customers. In general, though, consumers are rarely attractive defendants in cases like this, both because these suits are likely to engender backlash and because they have comparatively shallow pockets that limit the extent of any recovery. Many infringers are likely to be essentially "judgment proof," either because they lack financial resources or because they are beyond the practical reaches of litigation.

Finally, even when infringement is observable and a suitable defendant can be identified, firms may be deterred from enforcing intellectual property rights if they are concerned about risks associated with litigation. Not all grants of intellectual property rights are legally valid, which the scrutiny of litigation is likely to uncover. Lemley and Shapiro (2005) document that roughly half of all litigated patents are invalidated. If the firm anticipates a high probability of invalidation, it may opt not to enforce its rights at all.

2.3 Summary

Figure 1 summarizes the conditions under which gaps in the availability or enforceability of intellectual property rights give rise to a missing market, in the sense of Definition 1.

[31] Ormco Corp. v. Align Technology, Inc., 498 F.3d 1307 (Fed. Cir. 2007).

3 The New Uses Problem

3.1 Background

Effective therapies exist for only a fraction of known diseases. As of 2025, treatments have been developed for approximately 20 percent of recognized medical conditions (EveryCure, 2025). Among the subset of diseases with well-characterized molecular causes—those linked to specific genetic mutations, protein dysfunctions, or other biochemical abnormalities—this figure is lower; there, effective therapies are available for roughly five percent (Collins, 2011b).

Scientists have long argued that efforts to repurpose existing drugs are among the most cost-effective ways of expanding access to high-quality medical treatments. Some have suggested that systematic efforts to develop new uses for existing drugs could "convert cancer into a treatable chronic disease" (Telleria, 2012), unlock treatments for many or even all of the 8,000 rare diseases (Muthyala, 2011), and yield breakthroughs in diseases that have long resisted therapeutic intervention, such as Alzheimer's (Corbett et al., 2012; Butcher, 2013). Recent efforts to leverage artificial intelligence to identify potential new uses are premised on the same idea (EveryCure, 2025).

From a practical perspective, drug repurposing has clear advantages over de novo drug discovery. Any drug that is repurposed has necessarily already undergone long, costly clinical trials that establish its safety. Repurposing, then, requires only producing proof of efficacy in a new therapeutic context. Years—or, in some cases, decades—of clinical experience provide natural leads for these new contexts, streamlining the standard processes of trial-and-error to identify disease targets. As a result, efforts to commercialize new uses are faster, cheaper, and more likely to succeed than de novo drug development programs. While de novo drug development efforts typically take 12 to 16 years to bring products to market, new uses can be commercialized on a 3 to 12 year horizon (Dudley et al., 2011). And while the cost of developing a drug "from scratch" can exceed $1 billion, new uses can be commercialized for roughly $300 million (Sahoo, 2007). A final advantage: the Institute of Medicine estimates that 30 percent of repurposed drug trials succeed, compared to just 10 percent for new drugs (Collins, 2011a).

Yet, investment in repurposing is quite rare (EveryCure, 2025). Part of the challenge is thought to be timing. Evidence on new uses often emerges only with substantial delay, typically long after the original drug's patent has expired. Gelijns et al. (1998) observe that "it took half a century for the cardiovascular benefits of aspirin, the most widely used drug in the world, to be recognized, and nearly 40 more years before it was widely used for cardiovascular indications." By this time, generic use is widespread, and any investor will likely struggle to recoup the investment required to satisfy regulatory requirements.

Even when drugs are not formally repurposed, patients may receive prescriptions for new uses "off-label." In certain fields—notably, oncology—off-label prescribing is common, especially for patients who lack other treatment options (Saiyed et al., 2017). But many instances of off-label prescribing are, at best, ineffective and, at worst, actively harmful (Eguale et al., 2012). Even when substantial evidence supports off-label use, various barriers—such as insurer reluctance to cover

off-label prescriptions and limited physician knowledge of treatment options—can limit adoption. Berger et al. (2021) provide evidence consistent with the idea that large-scale diffusion of new uses occurs only following formal regulatory approval. The authors compare changes in the size of a drug's market after a clinical trial documents a link between the drug and a new use to changes after that new use is approved, formally, by a regulator. They find significant increases in market size only after the latter event. That is, formal regulatory approval appears key to facilitating market adoption.

3.2 A Missing Market for New Uses

In this section, we show that the conditions that give rise to a missing market, in the sense of Definition 1, are satisfied in the case of new uses for existing drugs. To illustrate, we walk through each step of the classification exercise introduced in Section 2.

p_{secret} : **Can the invention be profitably commercialized while remaining secret? No.**

It is difficult to imagine how the information that a drug has a new therapeutic use could be profitably commercialized while remaining secret. In the United States, federal law requires drugs to carry labels specifying each FDA-approved therapeutic indication.[32] This legal requirement aside, secrecy is impractical. A new use has economic value only if it is disclosed: physicians must know when to prescribe the drug, pharmacists know when to dispense it, patients must know how and why to take it, and insurers must know how to process reimbursements. Like other inventions that are essentially "recipes," a new use must be disclosed. Thus, $p_{\text{secret}} = 0$.

p_{eligible} : **Is the invention eligible for a legally-enforceable intellectual property grant? Yes.**

In the United States, new uses for existing drugs are eligible for two distinct forms of intellectual property protection. First, they may be protected by method-of-use patents. Even if a drug has already been commercialized (and its active ingredient is already protected by a patent), firms may seek patent protection for methods of using a drug to treat a new disease or condition.[33]

Second, new uses are eligible for regulatory exclusivities administered by the FDA. The FDA grants a three-year period of "new clinical investigation exclusivity" (NCI) to incentivize firms to develop new indications, dosage regimens, patient populations, or formulations of previously approved drugs. During this three-year exclusivity period, the FDA will not approve applications from

[32] See 21 U.S.C. § 352(f); 21 C.F.R. § 201.57(c).

[33] In re Marshall, 578 F.2d 301, 304 (C.C.P.A. 1978).

Recall, as introduced in Section 2.2, that the United States Patent and Trademark Office wil not grant patents on claims that are "directed to" natural laws, which do not add an "inventive concept" that "amounts to significantly more" than a patent on the ineligible concept itself. In one of the cases where this restriction was imposed, the Supreme Court invalidated a patent on a method of administering a drug. See Mayo Collaborative Services v. Prometheus Laboratories, 566 U.S. 66 (2012). Nonetheless, federal courts continue to grant patents on new uses for existing drugs. See Illumina, Inc. v. Ariosa Diagnostics, Inc., 967 F.3d 1319, 1325 (Fed. Cir. 2020) (holding that "method of treatment claims are patent-eligible"). See also Vanda Pharmaceuticals Inc. v. West-Ward Pharmaceuticals International Ltd., 887 F.3d 1117, 1134–36 (Fed. Cir. 2018) (holding that claims directed toward particular methods of treatment are patent eligible); Nat. Alternatives Int'l, Inc. v. Creative Compounds, LLC, 918 F.3d 1338, 1344 (Fed. Cir. 2019) (determining that "[t]hese are treatment claims and as such they are patent eligible").

generic competitors for the specific use protected by NCI exclusivity.[34] The limits of this restriction are worth emphasizing: when NCI exclusivity is in force, the FDA may still approve generic competitors for *other* indications. That is, a brand-name firm with NCI exclusivity protecting one approved use may still face generic competition for other, unprotected uses.

Thus, new uses may be protected by both patents and regulatory exclusivities. Hence, $p_{\text{eligible}} = 1$.

p_{enforce} : **Can the firm profitably enforce any intellectual property grant? Only for a fixed period of time.**

Patents on new uses for existing drugs are fully enforceable so long as the drug itself faces no generic competition. If there are no generic versions of a drug, a single firm controls the market and can set a price for its drug that reflects all of the uses that it has developed (even if it may not be able to perfectly price discriminate).

The structure of the pharmaceutical regulatory system in the United States ensures that, for most drugs, there is some guaranteed period without generic competition. The Hatch-Waxman Act of 1984 created a streamlined system of generic entry, which eased approval of generic products while guaranteeing periods of market exclusivity for brand-name drugs.[35] Under Hatch-Waxman, manufacturers of generic drugs need only demonstrate that their products are equivalent, in a chemical sense, to an innovator drug; they are not required to complete independent (costly, time-intensive) clinical trials to demonstrate the safety and efficacy of their products. Before the FDA will approve a generic competitor, however, the generic firm must challenge every patent associated with the brand-name drug. These challenges require the generic firm to review a list of patents designated by the brand-name firm and indicate that each is invalid, not infringed, or expired.[36] Until these challenges are resolved, the FDA will not approve a generic competitor; patents on the brand-name drug itself *and* patents on uses of the drug are perfectly enforceable as a result.[37]

Once generic entry occurs, enforceability of intellectual property rights on new uses collapses. The problem, for brand-name firms, is two-fold. First, they must identify instances in which a generic drug approved for the treatment of disease *a* is, instead, improperly used to treat disease *b*.

[34] To qualify for new clinical investigation exclusivity, the innovator firm must perform clinical studies defined by the FDA as "an investigation in humans, the results of which (1) have not been relied upon by FDA to demonstrate substantial evidence of effectiveness of a previously approved drug product for any indication or of safety in a new patient population, and (2) do not duplicate the results of another investigation relied upon by FDA to demonstrate a previously approved drug's effectiveness or safety in a new patient population." 21 CFR §314.108.

[35] The Hatch-Waxman Act is formally titled the Drug Price Competition and Patent Term Restoration Act of 1984. Pub. L. No. 98-417, 98 Stat. 1585 (codified in sections of 15, 21, and 35 U.S.C.).

[36] The list of relevant patents is provided by the brand-name drug and listed in a regulatory document maintained by the FDA, colloquially referred to as the "Orange Book." We provide details on this patent list, which is central to our empirical analysis, in Section 4.4.

[37] If a generic firm markets its product without FDA approval, it faces civil liability, criminal charges, and regulatory penalties under, for example, 21 U.S.C. § 355(a), which prohibits introduction into commerce of a "new drug" without approval; 21 U.S.C. § 331(d), which specifically prohibits introducing or delivering for introduction into interstate commerce any unapproved new drug; and 21 U.S.C. § 352, which allows the FDA to intervene in instances of drug "misbranding." The FDA can immediately seek injunctions to halt sales, issue warning letters, seize the offending products, or recommend criminal prosecution.

Second, they must devise a legal strategy that allows them to put an end to infringement. Neither is practical.

It is nearly impossible for brand-name firms to determine which form—brand-name or generic—of a medicine is dispensed to treat a particular disease.[38] Physicians will not hesitate to prescribe a generic "off-label" for uses associated with the brand-name drug; pharmacists, in some cases by law, may elect to dispense generic versions.

Even if brand-name firms could observe infringement of patents on new uses, the nature of liability in this setting makes enforcement unattractive. Courts have consistently shielded generic manufacturers from patent infringement liability unless they explicitly induce infringing use. This means that a generic manufacturer faces no legal repercussions even if its drug is used widely for unapproved uses, so long as it does not actively encourage or facilitate that use.[39] Instead, liability is likely to fall on those parties who actually "practice" the method of using a drug in violation of a patent—physicians who dispense drugs in in-patient settings and patients who comply with medical instructions at home. Technically, each risks a lawsuit. But even setting aside the impracticability of filing expensive lawsuits against entities with limited capacity to pay damages, Eisenberg (2005) notes that it is rarely a winning strategy for firms sue their (sympathetic, ill) customers en masse.

Thus, before generic entry, p_{eligible} takes a value of one: the brand-name firm can fully enforce its full portfolio of intellectual property rights. After generic entry, p_{eligible} falls to zero. It is generic entry, then, that creates a missing market in the sense of Definition 1.

3.3 Empirical Signatures

The conditions described in Section 3.2 generate testable empirical predictions. Consider an innovator drug that is approved by the FDA at time $t = 0$. For some fixed period, the innovator faces no generic competition, thanks to FDA oversight of generic competition. So long as this market exclusivity is in force, patents on new uses are fully enforceable. Since each new use that is commercialized yields monopoly profits until this period of exclusivity expires, incentives for investing in new uses are strongest at $t = 0$. As the date of exclusivity expiration approaches, the length of time during which any newly-identified use can be protected from competition drops: a new use identified at year $t + 1$ has one fewer year of enforceability than a use identified at year t.

This framework, thus, suggests three empirical signatures of a missing market for new uses:

1. Firms will have the strongest incentives to invest in discovering and commercializing new uses soon after initial approval, when the potential duration of enforceable monopoly rights is longest.

2. As the date of market exclusivity expiration approaches, incentives to invest in new uses

[38] To our knowledge, there have been no attempts to introduce a system of "indication-specific pricing" in the United States that would allow for this type of differentiation. Italy's health data infrastructure has allowed for pilot versions of such pricing regimes. There, the drug Avastin is priced differently for distinct types of cancer, as is the drug Erbitux. For details, see Pearson et al. (2016).

[39] Warner-Lambert v. Apotex Corp., 316 F.3d 1348 (Fed. Cir. 2003).

decline monotonically, as the expected time under monopoly for a new use decreases each year.

3. Upon generic entry and market exclusivity expiration, p_{enforce} drops to zero. Investment in new uses should, thus, fall to zero.

Section 4 introduces data that allow us to test these predictions empirically. Tests are introduced in Section 5.

4 Data and Measurement

4.1 Example: Lilly's Gemzar

The data we construct are best introduced with an example. In 1996, the FDA approved Lilly's Gemzar for the treatment of pancreatic cancer.[40] The molecule in Gemzar responsible for its pharmacological action—its *active moiety*—is called gemcitabine.[41]

Early scientific efforts to develop gemcitabine focused on its potential as an antiviral. When preclinical studies in the early 1990s documented that it was effective at killing leukemia cells *in vitro*, research effort shifted toward its potential as a treatment for cancer (Hertel et al., 1990). At the time of Gemzar's initial approval, Lilly was also investigating other potential uses, including gemcitabine as a treatment for smallpox.[42] Lilly then obtained re-approvals of Gemzar for several additional uses: in 1998 for the treatment of non-small cell lung cancer, in 2004 for the treatment of breast cancer, and in 2006 for the treatment of ovarian cancer.[43] By 2010, at the time of first generic entry, sales of Gemzar generated approximately $800 million per year.[44]

To test the predictions in Section 3.3, we would like to collect data on the timing of these research investments into new uses, relative to the end of Gemzar's period of market exclusivity. In light of the discussion in Section 3, patent expiration itself may be an imperfect proxy for this exclusivity period, as certain portions of a patent's term may be unenforceable in practice. The approval

[40] NDA #020509. See https://www.accessdata.fda.gov/drugsatfda_docs/nda/pre96/020509Orig1s000rev.pdf.

[41] Though the distinction between a drug's "active moiety" (the key pharmacological agent) and its "active ingredient" (the pharmacological agent, plus any stabilizing chemicals) is technical, it is significant for regulatory determinations that hinge on a drug's novelty. For our purposes, use of active moieties instead of active ingredients simplifies a series of medication name matching procedures. Concretely, Gemzar's "generic name" or "active ingredient" is often listed across databases as *gemcitabine*, *gemcitabine hydrochloride*, or *gemcitabine hci*. Each is chemically and therapeutically equivalent, but this type of variability in the extra inactive compounds (salts and esters) attached to the drug makes the string matching central to our empirical work extremely challenging. For each, the active moiety is *gemcitabine*. Appendix A.1.3 provides more details on the distinction between these objects, on the value of using active moieties to implement a series of string matches across datasets, and on our crosswalk from various forms of drug names (as are listed in nearly all input datasets) to active moieties.

[42] In our data, active clinical trials after Gemzar's initial approval investigated its effectiveness as a treatment for various types of cancer, hepatitis, and smallpox.

[43] Non-small cell lung cancer: NDA #020509, supplement #005; Breast cancer: NDA #020509, supplement #029; Ovarian cancer: NDA #020509, supplement #039

[44] See https://www.sec.gov/Archives/edgar/data/818686/000081868611000014/gemzar260111.htm ("Annual sales of Eli Lilly and Company's Gemzar® were approximately $814 million in the United States in 2010, based on IMS sales data").

of generic gemcitabine validates this concern. According to FDA's "Orange Book" database—an administrative record that collects all patents on approved drugs for which a claim of patent infringement could "reasonably be asserted"—Gemzar was protected by two patents.[45] The first-filed (and first-expiring) patent covered gemcitabine's chemical structure, while the second covered a method of using chemicals like gemcitabine to treat tumors in mammals.[46] The first generic version of Gemzar was approved on November 15, 2010, the day the chemical patent expired. The method-of-use patent, which remained in force until May 7, 2013, did not fully block generic entry.[47]

Section 4.2 details our approach to identifying drugs, like Gemzar, for which we track research investments into new uses. Section 4.3 introduces our measures of research investment—scientific publications, clinical trials, and re-approvals for new uses. In Section 4.4, we return to the challenge of determining a drug's period of market exclusivity using data on generic entry and patent expiration. Section 4.5 summarizes our data construction and presents basic summary statistics. Appendix A provides additional details.

4.2 Measuring Drug Approvals

We construct a census of new drugs approved by FDA between 1985 and 2014 using an administrative database, drugs@FDA.[48] We classify a drug an *innovator* and include it in our primary sample if it satisfies two criteria:

1. It was approved under a New Drug Application (NDA).

2. It represents the first approval of a unique combination of chemical components.

The first restriction limits our analysis to small-molecule drugs, which are relatively simple, stable chemical compounds synthesized through chemical processes. It excludes biologic drugs—large,

[45] The formal name of the "Orange Book" administrative record is *Approved Drug Products with Therapeutic Equivalence Evaluations*. Details on its history and contents are available in Durvasula et al. (2023). The main text of the Orange Book collects records of "therapeutic equivalence evaluations" for approved drugs—determinations of the equivalence of generic products to approved brand-name drugs.

When we refer to the Orange Book, we are referring to its appendices, which collect records of all patents and "regulatory exclusivities" associated with each brand-name drug. These records were made available as part of the 1984 Hatch Waxman Act, and coverage extends only to small molecule (i.e., non-biologic) drugs. The Hatch Waxman Act is formally called the Drug Price Competition and Patent Term Restoration Act. It is codified at 21 U.S.C. ch. 9 § 301. 21 CFR 314.54 lays out the listing requirements for patents and regulatory exclusivities.

[46] The first patent is U.S. Patent No. 4,808,614 (granted February 29, 1989), and the second patent is U.S. Patent No. 5,464,826 (granted November 7, 1995). Originally, both were set to expire 17 years after issuance: February 28, 2006 and November 7, 2012, respectively. In Durvasula et al. (2023), we provide additional details on each.

[47] This simplified description of patent protection, litigation, and generic approval for Gemzar captures the details that are central to our empirical analysis. In Durvasula et al. (2023), we offer a much more extensive case study of Gemzar.

[48] Our sample begins in 1985, one year after the 1984 Hatch Waxman Act introduced the modern system of pharmaceutical patent protection and generic drug regulation. The Hatch Waxman Act increased terms of patent protection for certain new drugs and introduced a new regulatory pathway for generic entry. From a practical perspective, it led to the creation of the Orange Book database that is central to our measurement of market exclusivity, introduced in Section 4.4. We select 2014 as an end point for two reasons. First, a key table in the drugs@FDA database that we use to identify re-approvals of drugs for new indications is unavailable in later years, and we have been unable to obtain analogous records via Freedom of Information Act requests. Second, for drugs approved in later cohorts, measures of market exclusivity may be censored (e.g., if a firm is still applying for new patents or regulatory exclusivities).

complex molecules derived from living organisms—which are subject to distinct legal and regulatory frameworks in the United States.[49]

The second restriction ensures that we focus on research incentives in the period following a drug's first approval.[50] Many drugs are approved under multiple brand names over time, usually after generic entry; these are often cases in which brand-name products can capture some substantial premium over generic alternatives (Bronnenberg et al., 2015). Acetaminophen, for example, is available both as a generic and under the brand names "Tylenol" and "Panadol." Ibuprofen, too, is available both as "Advil" and "Motrin." We keep only the first of these brand-name approvals.

For each approval and re-approval event in our sample, we construct data on the drug's uses—the specific indications for which it is approved—building on a method described in Berger et al. (2021).[51] We collect text descriptions of these uses from multiple sources, including FDA approval packets, company press releases, scientific articles, and other publicly available records. Whenever possible, we use the FDA's full-text drug labeling service (FDALabel) to retrieve official drug labels. We supplement data on labels—which are primarily available from 2005 forward—with data from drug approval packets, dated company press releases, scientific articles, and other internet searches. To standardize uses across drugs, we assign each approval or re-approval one or more International Classification of Diseases codes, based on its Tenth Revision (ICD-10). These alphanumeric codes systematically classify diseases, symptoms, and related health conditions.

4.3 Measuring Research Investments

For each drug in our sample, we construct three measures of research and development into new uses: scientific publications, clinical trials, and FDA re-approvals for new uses.

Scientific Publications We collect records of scientific publications from the National Library of Medicine's PubMed database. PubMed indexes nearly every biomedical publication and, importantly for this analysis, curates links to chemicals mentioned in the text of each paper. For each innovator drug, we identify all publications indexed with its chemical components. We track, also, whether each scientific publication received any funding from a public agency.

To match drug names in FDA records to standardized PubMed terms, we develop a crosswalk procedure, detailed in Appendix A.2. In short, we clean and standardize drug names from FDA administrative records, then iteratively link them to standardized forms of chemical names in PubMed.

[49] Biologic drugs include gene therapies, monoclonal antibodies, and vaccines. A different set of statutes govern federal research incentives for biologics, and FDA applies a different set of standards for generic ("biosimilar") approval. One important distinction, from the perspective of our empirical analysis, is that our data construction relies on the existence of a mapping between small-molecule drug patents and approved drug products, the Orange Book. No analogous regulatory document exists for biologic drugs. Frakes and Wasserman (2025) hand-review drug patent filings to construct what is, to our knowledge, the most complete listing of biologics patents.

[50] This restriction is similar to, but not the same as, the regulatory designation of "new molecular entity" (NME). NMEs are a strict subset of our sample. See Appendix A.1.2 for details.

[51] We are grateful to Ben Berger, Amitabh Chandra, and Craig Garthwaite for sharing a crosswalk between drugs and indications, which covers a subset of our period of interest. We build on Berger et al. (2021) to produce a crosswalk between drugs and uses that covers our full sample period.

This process successfully maps approximately 75 percent of drugs in our final sample to a PubMed term.

Clinical Trials We use clinical trial activity as a proxy for drug commercialization. To maximize coverage over our sample period, we compile trial records from four sources: NDA Pipeline (1982-2001), Pharmaprojects (1995-2010), Cortellis (1995-2016), and ClinicalTrials.gov (2010-2016). As with scientific publication records, the challenge in using these data is that drug names in FDA administrative records often do not match those in these external databases. Unlike PubMed, however, where drug names are indexed with standardized terms, clinical trial databases often contain inconsistent formatting and typographical errors. We develop a distinct linkage procedure for each database, detailed in Appendix A.3. This approach successfully matches nearly 70 percent of drugs in our final sample to at least one clinical trial record.

Our final dataset includes trial records for all four sources. To avoid concerns about inconsistencies in coverage, double-counting, and other errors in measurement across datasets, our primary measure of trial activity is a binary indicator that takes a value of one if an innovator drug has any active clinical trial in a given year and zero otherwise.

New Use Re-Approvals FDA administrative records classify applications into three types: new applications, supplemental applications, and generic drug applications. We collect records of all approvals within the supplemental applications category that were tagged as "new and modified indications." As a point of comparison, we also collect records of all other re-approvals associated with supplemental applications.

4.4 Measuring Market Exclusivity

The framework introduced in Section 2 offers a simple prediction: when intellectual property rights on new uses become unenforceable, private firms will lose any incentive to pursue further research and development. It is the entry of substitutable generic competitors that creates the problem of enforceability. Empirically, then, we are interested in measuring trends in investment and commercialization in the periods around expected market exclusivity expiration.

A natural proxy for a firm's expectations regarding timing is the date on which generic approval actually occurs, insofar as this captures market participants' rational expectations. The difficulty in using date of generic approval as a proxy for expected market exclusivity expiration is that more than half of innovator drugs in our sample (548 / 990, 55%) never experience generic entry. This is not a peculiar feature of our data. Although a series of legal reforms in the 1980s—including the 1984 Hatch Waxman Act itself—are credited with increasing generic entry for small-molecule drugs, FDA (2025) lists more than 350 ever-approved small-molecule drugs with no active patents (or regulatory exclusivities) for which there is no associated generic version.[52]

[52] Scott Morton (1999) documents similar patterns and analyzes the determinants of these generic entry decisions. For a recent analysis of challenges that limit generic entry in the "patent afterlife," see Hemel and Ouellette (2023a).

An alternative proxy, then, might be the date on which the "main" patent on a drug expires. By construction, we observe at least one form of intellectual property for every drug in our sample. Using the date of patent expiration as a proxy is, nonetheless, challenging for three reasons. First, new drugs approved for sale in the United States are, often, protected both by patents and by non-patent "regulatory exclusivities." The FDA will deny approval to a competing generic product so long as certain exclusivities are in force, and these exclusivities may run sequentially or concurrently with patents.[53] Second, new drugs are, increasingly, protected by many patents.[54] The median innovator drug approved in 1990 was protected by a single patent grant. In 2014, the median drug had 5 associated patents. These patent portfolios can confer periods of *nominal* market exclusivity that far exceed that provided by a single patent—though as the Gemzar case highlighted, not all of these patents will limit generic entry. Third, and relatedly, would-be generic competitors are incentivized—as part of the regulatory design of the 1984 Hatch Waxman Act—to "challenge" patents on brand name drugs by asserting that any patents in force are invalid, unenforceable, or not infringed by the generic.[55] Successful challenges allow generic entry before the expiration of all patents. Any potential measure based on the expiration of patents, then, must account for the likelihood that each patent and regulatory exclusivity will successfully block generic entry.

With these practical limitations in mind, our primary measure of market exclusivity is constructed as follows. For all drugs, we begin by assuming that firms expect all patents and regulatory exclusivities to bind.[56] If a generic competitor is approved prior to this date, however, we revise our measure as,

$$\text{market exclusivity} = \min\{\text{generic approval}, \text{expiration of all patents/exclusivities}\}.$$

This approach leverages the strengths of both proxies—capturing both the market's rational expectations and the legally defined barriers to competition—while correcting for the limitations of each. In effect, we are selecting an upper bound on the firm's expected period of market exclusivity.[57] In Section 5.2, we define an additional measure of market exclusivity, the *minimum patent exclusivity*, that uses only the first expiring patent, in order to address concerns that our primary measure captures strategic secondary patent accumulation.

[53] Regulatory exclusivities are often called quasi-patents, in recognition of the fact that they also allow a firm to exclude its competitors from the market for some fixed period of time. They differ from patents, however, in two key regards. First, regulatory exclusivities in the pharmaceutical sector are awarded by the FDA on the basis of its own assessments of a product. Second, exclusivities are self-enforcing.

[54] For more details on "patent portfolios" in pharmaceutical markets, see Hemphill and Sampat (2012); Durvasula and Ouellette (2026).

[55] 21 U.S.C. § 355(j)(5)(B)(iv). Under the Hatch Waxman Act, the first (successful) generic challenger receives a bounty in the form of exclusive rights to the generic market for a 180-day period. See Hemphill (2006) for one analysis of the incentives that this system creates for agreements between generic and brand-name firms.

[56] We restrict consideration to the four regulatory exclusivities that we determined, in Durvasula et al. (2023), are "generic blocking" in the sense that the FDA will not approve a generic form of the active moiety while these exclusivities are in place: new chemical entity exclusivity, orphan drug exclusivity, pediatric exclusivity, and generic antibiotic incentives now exclusivity. Technically, there are sixteen types of FDA-granted regulatory exclusivities. Table 5 in Durvasula et al. (2023) provides a list.

[57] In Durvasula et al. (2023), we consider alternative measures in detail.

To construct our primary exclusivity measure for each innovator drug in our sample, we collect records of generic competitor approval and records of all patents and exclusivities. Specifically, we identify the first approval of a generic competitor for each drug in our sample in Drugs@FDA, as the first approval of the same combination of active moieties via the FDA's generic approval pathway.[58] We also collect records of patents and regulatory exclusivities in the Orange Book, as digitized in the NBER Orange Book Dataset, for each of our sample years. Appendix Figure A1 plots the distribution of this market exclusivity measure for our sample of innovator drugs.

4.5 Sample Construction & Summary Statistics

There are 1,102 drugs that meet our criteria for innovator drugs. In most analyses, we restrict consideration to drugs with at least one patent or regulatory exclusivity. We also exclude a small number of observations that predate the modern FDA, for which our measure of market exclusivity takes a negative value.[59] This yields a final sample of 990 drugs. Table 1 presents summary statistics.

5 Empirical Tests

This Section develops empirical tests of the three central elements of the framework introduced in Section 2. Section 5.1 documents evidence consistent with the predictions introduced in Section 3.3—that firms will have the strongest incentives to invest in new uses soon after initial approval; that as market exclusivity expiration approaches, the number of years in which new uses will be profitable declines and hence so does investment in their commercialization; and that private research and development will cease after exclusivity expiration. Section 5.2 provides causal evidence that longer durations of market exclusivity induce the commercialization of additional new uses. Section 5.3 documents that the enforceability of intellectual property rights is the key mechanism determining the availability of incentives for investment.

[58] This definition of generic approval introduces an important simplification. Generics are typically approved at the level of a particular product (e.g., gemcitabine in a one gram injectable form) and may not be approved for every dosage or form that is available for the brand-name product. In some cases, these limited approvals reflect the existence of patents and exclusivities that protect some forms of a drug but not others. The full details of the Gemzar case study are reported in Durvasula et al. (2023), which provides additional context on this type of product-by-product entry. In other cases, limited entry is part of a negotiated settlement. A second feature of this definition is neither an assumption nor simplification, but a piece of institutional context worth noting: we focus on the timing of generic approval, rather than generic entry. Hemphill and Sampat (2011) establish that these two dates are typically close in time. In a small number of cases, however, there may be lengthy delays between generic approval and entry—if, for example, there is some agreement between the parties to delay generic entry. For an example of this type of delay, see FTC v. Actavis, Inc., 570 U.S. 136 (2013).

[59] There are nine drugs that are excluded because market exclusivity duration is negative: NDA #19008 (bretylium); NDA #19931 (sulfacetamide); NDA #21794 (dapsone); NDA #22041 (hydroxocobalamin); NDA #22556 (carbinoxamine); NDA #50606 (vancomycin); NDA #50640 (oxacillin); NDA #50763 (mitomycin); NDA #204031 (acetaminophen; oxycodone). We reviewed FDA records and drug approval histories. In each case, although our data construction method properly captures the first NDA approved by the modern FDA—that is, since the 1962 passage of the Kefauver–Harris Amendment to the Federal Food, Drug, and Cosmetic Act created the modern regulatory system—we did not identify an earlier NDA approved by the pre-1962 FDA. Generic versions of these much older drugs were approved before the drugs we flagged as innovators. There are a small number of other difficulties that arise when drugs in our sample were grandfathered into new regulatory schemes introduced in the 1960s and 1980s. We offer details in Durvasula et al. (2023).

5.1 Testing for Empirical Signatures of a Missing Market

Figure 2(A) presents the simplest test of the predictions introduced in Section 3.3, a comparison of the two objects of interest: the probability that a drug is re-approved for any new use in each year (y-axis) against the time relative to the drug's market exclusivity expiration (x-axis). As predicted, Figure 2(A) indicates that the probability that a drug is re-approved for a new use peaks early in its lifecycle. The average innovator drug in our data receives approximately thirteen years of market exclusivity, and the probability that a drug is re-approved for a new use is highest in year $t = -10$.[60]

While Figure 2(A) focuses on the likelihood that *any* new use is re-approved in each year Figure 2(B) presents trends in the *quantity* of new uses graphically, plotting the cumulative share of new uses approved in each year relative to market exclusivity expiration. Of the 433 new uses approved by the FDA in our data, nearly ninety-five percent were approved before market exclusivity expiration. Put differently, only five percent of new use approvals associated with the nearly 1,000 innovator drugs in our sample —twenty-three new uses total—occurred after intellectual property rights became unenforceable.

Figure 3 replicates Figure 2 using time since initial approval as the x-axis. This reorientation makes clear patterns in our data also support the two other central predictions in Section 3.3. The probability of re-approval for a new use is nearly zero at the time of the drug's initial approval, reaches its peak two years after initial approval, then declines monotonically. At this peak ($t = 2$), approximately five percent of drugs are re-approved for a new use. Fifteen years after initial approval, the likelihood of re-approval for a new use falls to—and remains at—zero.[61]

Figures 2 and 3 pool trends for all 990 innovator drugs in our sample. Figure 4 confirms that these trends are robust to the use, instead, of balanced panels, in which we have data on each drug in each year relative to market exclusivity expiration. Of course, the consequence of a balanced panel approach is selection: there are 124 innovator drugs in the ten-year balanced panel (which requires twenty years of data), and 415 in the five-year balanced panel (which requires ten years of data). The sample size is, however, the only feature of these plots that is meaningfully shifting. Figure 4(A) and Figure 4(B) match the predictions in Section 3.3: new use approvals are most likely shortly after approval and decline sharply over time, before falling to zero five years after exclusivity expiration.

There are two natural concerns with the interpretation of these patterns as consistent with a missing market for new uses. First, the set of scientifically viable opportunities for new uses might be naturally exhausted over a drug's lifecycle, such that the drop off in research investment at the time

[60] Our estimate of average market exclusivity is not atypical. Average and median measures of market exclusivity have, perhaps surprisingly, remained essentially constant across our period of interest, despite large increases in the total number of patents protecting each innovator drug. See Hemphill and Sampat (2012) and Durvasula and Ouellette (2026) for extended discussions of these trends.

[61] As Section 4.4 notes, our market exclusivity measure abstracts from some complications in pharmaceutical markets. Recall that we define market exclusivity as the earlier of (i) generic approval and (ii) the expiration of all patents and regulatory exclusivities. Although the timing of generic approval is tightly linked to the timing of generic entry, these dates need not coincide. In complex cases, the timing of generic entry following regulator approval may be negotiated to include some period of delay (Hemphill, 2006) or a restriction on sales volume and pricing.

of exclusivity expiration reflects only that there are no potential new uses left to develop. Second, and relatedly, firms may sequence their investments strategically, as the timing of market exclusivity expiration is known to all market participants in advance. If firms choose to commercialize all potential new uses before exclusivity expiration, then the patterns in Figure 2 would not suggest that innovation is "missing."

The remainder of this Section constructs tests to address both concerns. First, however, we ask whether the trends in Figure 2 are consistent with trends in two additional measures of research effort: clinical trials and scientific publications. Clinical trials are a useful point of comparison as they capture both successful and unsuccessful investments in new uses, whereas FDA approvals capture only the former; scientific publications provide a rough measure of the set of available scientific opportunities over time. If the likelihood of an active clinical trial increases, or even remains stable, across a drug's lifecycle, it would suggest at least three interpretations of Figure 2 inconsistent with our hypothesis: (i) that firms do continue to invest in new uses after exclusivity expiration; (ii) that the most successful investments, by strategy or coincidence, occur early in a drug's lifecycle; (iii) that firms continue to produce useful information about drugs after exclusivity expiration but are less likely to seek regulatory approval. If the quantity of scientific publications dropped off sharply around the time of exclusivity expiration, we might be concerned that the trends in Figure 2 reflect the exhaustion, also by strategy or by coincidence, of scientific opportunities.

Figure 5(A) plots the probability that any clinical trial studying an innovator drug is active in each year relative to exclusivity expiration. The likelihood of an active clinical trial peaks eighteen years prior to exclusivity expiration, when nearly half of innovator drugs are studied in at least one trial. The likelihood of any ongoing investment declines sharply and monotonically thereafter. Ten years after exclusivity expiration, less than twenty percent of innovator drugs are part of any active clinical trial; this fall to less than ten percent after year fifteen. Recall, as discussed in Section 4.3, that our clinical trial data, constructed to span nearly forty years, do not allow us to distinguish between trials of different phases. This declining trend then also includes phase four "post-marketing surveillance" trials, academic studies that aim to refine the use of a drug in clinical practice, and foreign studies intended to satisfy the requirements of other regulatory agencies.[62] Put differently, this pattern of progressively declining commercial interest in innovator drugs over time, if anything, may overstate the amount of investment in potential new uses later in a drug's lifecycle. With this note about the level of investment in Figure 5(A) in mind, the trend is consistent with the three predictions in Section 3.3, as well as the pattern of approvals in Figure 2.

Figure 5(B) replicates Figure 2(A) and Figure 5(A) using, instead, counts of scientific publications that study the innovator drug in each relative year. To the extent that scientific publications capture "scientific opportunities" rather than commercial interest, we do not expect the number of

[62] As Section 4 noted, constructing clinical trial data that covers our time period of interest requires drawing on four distinct sources of data, with different amounts of detail for each study. Gaps in trial sponsor and phase information limit our ability to refine this sample to include a narrower set of trials. Even if we had sponsor information, the fact that trials may be run by entities other than the firm that eventually markets a drug make this type of restriction conceptually challenging to implement.

scientific publications studying a drug to fall to zero over its lifecycle. Indeed, although we observe comparable declining trends in both publications and clinical trials—both measures drop by a factor of two over the thirty year period around exclusivity expiration—the level of the publication measure remains comparatively high. Even twenty years after exclusivity expiration, the average drug in our sample is the subject of twenty scientific publications each year.

Our publication data allow us to shed some light on the source of the declining trend, by differentiating between publications with public versus private funding. Appendix Figure A5 reproduces Figure 5(B), with publicly-funded publications plotted separately. The total number of publicly-funded publications (red) is essentially stable, around ten, year-on-year across our full period of interest. The decline in scientific publications over time in Figure 5(B) appears to be driven entirely by a decline in publications with exclusively private funding (blue). The stability of publicly-funded scientific publications is suggestive evidence that research opportunities continue to emerge over time, even as private incentives fall. And scientific opportunities for new uses continue to arise across a drug's useful life—potentially decades after market exclusivity expiration, as in the case of aspirin highlighted in Gelijns et al. (1998)—then firms are, necessarily, not commercializing all potential uses in the period immediately following initial approval.

5.2 Estimating the Causal Effect of Exclusivity on Research Investment

Section 5.1 documents trends in research investment and commercialization consistent with the three empirical predictions introduced in Section 3.3. A concern with these descriptive analyses, however, is that they may capture only the exhaustion of opportunities for investment in new uses, rather than the effect of market exclusivity on innovation. This Section leverages sources of variation in the periods of market exclusivity provided to innovator drugs that are plausibly unrelated to the drugs' underlying scientific potential to provide this causal evidence.

5.2.1 Investigating the effect of market exclusivity on the quantity of new uses

If gaps in private incentives—rather than scientific opportunities—explain the trends in Figures 2 through 5, then drugs with longer periods of market exclusivity should have a larger number of realized new uses. The intuition for this relationship is similar to that underlying our explanation of the gradual declines in research investment documented in Section 5.1: increases in the duration of market exclusivity affect the number of new uses that are profitable to commercialize on the margin. Here, we show that drugs with longer periods of market exclusivity have more commercialized new uses. That is, there are some scientifically feasible new uses that firms pursue only when there is sufficient market exclusivity remaining.

Our empirical analyses use a measure of market exclusivity for which there is substantial variation across drugs in our sample: the average innovator drug receives 12.9 years of exclusivity, but drugs at the 10th and 90th percentiles of this measure receive, respectively, 5.5 and 42.5 years.[63]

[63] Appendix Figure A1 plots the distribution of this measure.

These differences arise in two ways. First, recall that our primary measure of market exclusivity is intended to account for various complicating aspects of patents on pharmaceuticals—in particular, that most innovator drugs in our sample are protected by multiple overlapping patents and regulatory exclusivities, only some of which are likely to bind. These secondary patents and regulatory exclusivities can, in some cases, push a drug's market exclusivity period beyond the statutory twenty-year patent term. Second, various delays in the drug development process can erode a drug's effective patent life. Firms in the pharmaceutical sector have strong incentives to file primary patents on an innovator drug as soon as possible—both to protect their investments from competitors and to avoid creating "prior art" that would make the drug unpatentable for lack of novelty.[64] At least three types of delays between patent filing and initial drug approval can decrease effective patent terms: commercialization lags due to differences in clinical trial lengths (Budish et al., 2015), delays in clinical trial initiation (Gilchrist, 2016), and delays in patent prosecution (USPTO, 2025).

Figure 6(A) plots the relationship for drugs in our sample between this measure of market exclusivity and the number of realized new uses as a binned scatterplot with an overlaid line of best fit. Table 2(A) formalizes this relationship in a regression framework. In particular, for innovator drug i, we estimate the regression

$$y_i = \alpha + \beta \cdot \text{exclusivity}_i + \mu_i + \varepsilon_i \tag{7}$$

where the dependent variable y_i captures the number of re-approvals of innovator drug i for a new or modified indication and μ_i is an approval year fixed effect. The estimate of interest is β, the coefficient on our measure of each drug's market exclusivity exclusivity_i. Columns (1) and (2) report results using the measure of market exclusivity introduced in Section 4.4—the earlier of generic approval and the expiration of all patents and exclusivities. The relationship between this measure of exclusivity and the quantity of realized new uses is positive and significant in all specifications. with and without the approval year fixed effect μ_i. We return to an economic interpretation of these estimates below and in Section 6.

Of course, market exclusivity terms are not randomly assigned across drugs. In fact, to the extent that firms apply for secondary patents with the goal of delaying generic entry, one may be concerned that firms can, in effect, endogenously determine the length of this period. Figure 6(A) and Table 2(A), then, may capture several related effects, including the causal effect of longer market exclusivity on innovation, as well as the strategic behavior of firms that choose to extend periods of exclusivity on certain drugs.

To address this concern, we construct an alternative measure of market exclusivity that captures sources of variation that decrease effective patent term—clinical trial durations, clinical trial delays, and patent prosecution delays—but excludes all secondary patents. For each drug, we define its *minimum patent exclusivity* as the time between its initial approval and the expiration of its earliest patent and all FDA-administered regulatory exclusivities. By focusing on the minimum period

[64] Budish et al. (2015) and Durvasula and Ouellette (2026) provide additional context on firms' incentives to file patents early in the research and development process.

of exclusivity—rather than the maximum period that firms can secure by accumulating secondary patents—we retain variation across drugs that is plausibly exogenous. On this measure, the average drug in our sample has 9.4 years of (minimum patent) exclusivity; the 10th and 90th percentiles of this measure are 5 and 35 years.[65]

Figure 6(B) plots the relationship between this minimum patent measure of exclusivity and the number of realized new uses. In Table 2(A), Columns (3) and (4) reproduce the regression specifications introduced above. Again, the relationship between these measures is positive and significant.

Our preferred estimates are reported in Columns (2) and (4) of Table 2(A) and include year of approval fixed effects, which account for the concern that earlier-approved drugs have had more years in which to realize additional new uses. In both specifications, the estimated coefficient is roughly 0.02.

There is one additional econometric challenge for this exercise, which arises because nearly three quarters of innovator drugs in our sample ($N = 725, 73\%$) have zero re-approvals for new uses: the existence of a mass point at zero may bias estimates in Table 2(A) (e.g., Mullahy, 1998). To address this concern, in Table 2(B), we report estimates from a two-part model that considers the extensive and intensive margins separately. On the extensive margin, we recode the outcome variable y_i to take a value of one if there are any re-approvals of innovator drug i and zero otherwise. We report estimates from a logit model. On the intensive margin, we restrict consideration to those drugs with at least one re-approval and estimate, via ordinary least squares, the effect of market exclusivity on the number of re-approvals.

On the extensive margin, we find a significant, positive relationship across specifications between the duration of market exclusivity and the approval of any new indication. Estimated coefficients—0.26 in Column (1) and 0.42 in Column (2)—suggest that an additional year of market exclusivity increases the odds of observing any re-approval for a new uses by a factor on the order of 2 to 8 percent.[66] The intensive margin restriction, however, leaves us underpowered to estimate an effect. Estimated coefficients from each specification are positive and comparable in magnitude to those reported in Panel A, but none are significantly different from zero.

In addition to providing support for the existence of missing incentives for investment in new uses, this empirical exercise speaks directly to a long-standing open question in the law and economics of innovation: do longer periods of intellectual property protection *cause* additional invest-

[65] Appendix Figure A2 plots the distribution of this measure. In 52 cases (5%), this minimum patent measure of exclusivity is longer than a statutory twenty-year patent term. This arises, in nearly all cases, because the earliest-expiring patent was filed after the drug's initial approval. While it is generally true, as we note above, that firms have strong incentives to file key patents before approval, this is not always the case. On inspection, many of these later-filed patents appear to be method-of-use patents on drugs for which there is no Orange Book patent protecting the active ingredient. We are unaware of data that would allow us to examine Orange Book patents in our data that are method-of-use versus active ingredient patents; the only effort to implement this type of categorization, to our knowledge, is Kapczynski et al. (2012), who do so for a subset of the Orange Book by reading the text of each patent and assigning labels by hand.

[66] We obtain these estimates by exponentiating the logit coefficients to obtain the corresponding odds ratios: $e^{\beta} - 1$. The coefficient of interest β in Column (1) is 0.026, which corresponds to an odds ratio of $e^{0.026} - 1 \approx 0.026$, or a 2.6% increase in odds per year.

ments in research and development? That more new uses are commercialized when drugs quasi-randomly receive longer periods of market exclusivity provides, to our knowledge, the first direct evidence that the answer to this question is yes.

5.2.2 Interpreting these estimates as an elasticity

The structure of this empirical exercise allows us to estimate directly, as well, the elasticity of research investments with respect to market exclusivity, a key parameter for the design of optimal patent policy that has been similarly challenging to estimate (Nordhaus, 1969; Budish et al., 2016).[67] This elasticity is, of course, specific to our empirical context. Nevertheless, it is directly relevant to any efforts to alter the structure and availability of intellectual property rights. In Section 6, we use this elasticity to estimate the number of uses that would have been developed had intellectual property rights on new uses been enforceable for a longer period of time the social cost of missing incentives for new uses.

Returning to Table 2(A), these regression estimates lend themselves to a straightforward interpretation: an additional year of market exclusivity would yield, on average, 0.02 new uses. To realize one additional new use per drug would, thus, require providing an additional 50 years of exclusivity on average.

To contextualize this estimate relative to prior work, we calculate the implied elasticity in percentage terms. We benchmark both the dependent variable (the number of new use approvals) and our measures of exclusivity at their sample means. Let $\bar{y} = 0.437$ be the average number of new use approvals per innovator drug and $\bar{x}$ be the average years of exclusivity (13.35 for our primary measure and 10.24 for the "minimum patent" measure). Given our coefficient estimate of 0.02, the point elasticity at the mean is $\beta \cdot \frac{\bar{x}}{\bar{y}}$. For our primary exclusivity measure and the minimum patent exclusivity measure, respectively, this implies an elasticities of approximately 0.61 and 0.47.[68] Interpreted directly, these estimates imply that doubling the average market exclusivity would increase the expected number of re-approvals by approximately 47 to 61%.

Two points of comparison for our estimate appear in Budish et al. (2015) and Abrams (2009). Budish et al. (2015) estimate the semi-elasticity of research investment with respect to a one-year change in the commercialization lag—t_{comm} in Section 2—for new cancer therapies as 7-23%. Abrams (2009) uses changes in the duration of patent term introduced by the adoption of the 1995 Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) to estimate the responsiveness of patent filings to changes in patent term length—t_{IP} in Section 2. In that context, a one year increase in patent term is estimated to generate a 66% increase in patent filings.

[67] In particular, Budish et al. (2016) write: "A key parameter needed to inform optimal patent policy design is the elasticity of research investments with respect to the patent term. Estimating this elasticity is conceptually difficult because it requires constructing a counterfactual in which we can infer that some scientifically feasible inventions would have been brought to market under an alternative patent policy design. Despite a near half-century of research effort, we have essentially no credible empirical evidence on this elasticity."

[68] Primary exclusivity measure: $= 0.02 \times \frac{13.35}{0.437} \approx 0.61$.

Minimum patent measure: $= 0.02 \times \frac{10.24}{0.437} \approx 0.47$.

Our estimates imply that adding one year of market exclusivity increases the average number of approvals for new uses by 0.02—roughly a 4.6% increase relative to the average of 0.437. That our estimate is smaller than that in either Budish et al. (2015) or Abrams (2009) is, in several ways, unsurprising. First, note that our context may provide the cleanest setting in which to isolate the effect of intellectual property duration on research investments. Budish et al. (2015) capture an effect that reflects both patent duration and corporate short-termism and, thus, may be an over-estimate. We expect Abrams (2009) to offer an over-estimate of the object of interest, given the context of the policy change used to generate the estimate: TRIPS changed both the expected patent term for certain inventions and the structure of intellectual property rights more generally. For some inventors, holding patent rights under the new system was preferable, for reasons unrelated to the duration of the term. Though we can set aside the strategic timing dimension and consider only the change in duration to calculate an elasticity, it is unsurprising that our estimate, drawn from a more precise research design, is smaller in magnitude. That being said, the second explanation is that, in many ways, our context is still quite narrow. Inventions in our setting are clinically validated links between drugs and new diseases, which in effect function as improvements to existing products. Third, Eisenberg (2005) provides an additional, context-specific explanation: qualitative evidence suggests that firms may be hesitant to invest in new uses for commercially successful drugs if there is some risk that additional clinical trials will yield "bad news," such as large numbers of adverse events in a particular patient population, that would threaten the profitability of the entire market.

5.3 Establishing Enforceability as a Mechanism

Section 2 introduces a framework in which the enforceability of intellectual property rights is the key determinant of the availability of incentives. Section 5.1 documents evidence of a sharp decline in investment and commercialization at the time of exclusivity expiration—when intellectual property rights become unenforceable—that is consistent with this hypothesized relationship. This Section introduces an empirical test that provides direct evidence that the mechanism creating a missing market for new uses is the unenforceability of intellectual property rights.

Recall that a missing market for new uses arises at the time of generic entry because a generic drug approved only to treat disease a can nearly always be used "off-label" to treat diseases b, c, d etc., even if a brand-name firm holds patents on those uses. Patents on new uses are entirely unenforceable because the generic drug is perfectly substitutable for the brand-name drug.

Not all re-approvals of existing drugs face the same enforceability problem at the time of generic entry. Generic competitors must individually challenge patents and demonstrate bioequivalence for each distinct formulation. Stacked patents and exclusivities protecting certain dosages, forms, and patient populations of a brand-name drug may mean that generic products are approved narrowly. For example, in the case of Gemzar (see Section 4.1), the first generic competitor approved was allowed to market only one strength of the drug, its 2-gram version. Patents and regulatory exclusivities blocked generic entry on the 1-gram and 200-milligram versions. In many cases, first generic competitors are approved only for one patient population (e.g., adult, not pediatric) or one

form (e.g., tablet, not liquid).

Thus, incentives to invest in any formulation of an innovator drug should decline as market exclusivity expiration approaches, as the expected profitability of an entire drug development program drops over time. But intellectual property rights on other forms of a drug that are not perfect substitutes for the generic may remain at least partially enforceable ($p_{\text{enforce}} \in (0,1)$), even as intellectual property rights on new uses become entirely unenforceable ($p_{\text{enforce}} = 0$). For forms of innovator drugs that will not face immediate generic competition, we thus do not expect research investment to fall as sharply at the time of market exclusivity expiration.

To assess this prediction in our data, we first consider the likelihood that an innovator drug is re-approved in each relative year for any reason—including for addition of a new strength, dosage, or patient population. This trend is depicted graphically in Figure 7(A). Visually, we find a much more gradual decline in re-approvals around the time of market exclusivity expiration than that documented in Figure 2, consistent with the idea that firms may retain some incentive to invest in new forms of existing drugs that are not direct substitutes for early generic competitors. Recall that 95 percent of successfully commercialized new uses in our data were approved prior to market exclusivity expiration. When we consider, instead, re-approvals of an innovator drug for any reason, this estimate falls to 71 percent.

We test this idea more directly by considering the case of fixed-dose combinations—drugs dispensed as single "dosage forms" (e.g., a single tablet) that contain two or more distinct drugs.[69] Fixed-dose combinations often represent substantial improvements in the quality of care. For example, the widely-used HIV treatment and preventive Truvada is a single tablet that contains two distinct drugs, emtriciabine and tenofovir disoproxil fumarate, neither of which on its own offers the same therapeutic benefits.[70] Fixed-dose combinations are also often preferred by patients and providers as they simplify complex treatment regimens and increase prescription adherence.[71] For our purposes, however, the relevant feature of a fixed-dose combination is that even if its components have generic competitors, these components are not perfect substitutes for the fixed-dose combination.

Thus, we examine whether firms shift investments toward fixed-dose combinations as exclusivity expiration approaches. Specifically, we restrict attention to innovator drugs composed of a single active moiety and track combination products containing the same moiety. Figure 7(B) documents that, in the year of exclusivity expiration, the likelihood that a new indication for a combination drug

[69] Dix and Lensman (2024) study incentives for the development of combination therapies that are distinct from both clinical and regulatory perspectives. Cancer combination drugs have components that are dispensed separately, and the components are regulated as distinct drugs.

[70] Truvada (emtriciabine; tenofovir disoproxil fumarate), NDA#021752, was approved by the FDA on August 2, 2004. A generic version, manufactured by Teva Pharmaceuticals, was approved on June 8, 2017. Emtriciabine was approved on July 2, 2003, and a generic approved on April 17, 2017. Tenofovir disoproxil fumarate was first approved on October 26, 2001, and a generic was approved on January 17, 2012.

[71] Combination drugs have many advantages for patients, chief among which is simplicity. The development of the first single-tablet regimen for HIV, Atripla (Efavirenz/emtricitabine/tenofovir), was lauded as a significant improvement for patients, many of whom had previously been taking up to 36 distinct pills per day. See Pollack (2006). Analyses of claims data suggest that treatment adherence is higher in single-tablet regimens than in multi-tablet regimens. See Silverman (2018).

is approved more than doubles. The likelihood of a combination approval remains at this elevated level in the following years, as the likelihood of new use re-approvals on the innovator drug itself falls to zero.

This fact has two implications worth emphasizing. First, that the number of approvals of fixed-dose combinations *increases* at the time of market exclusivity expiration counsels against the idea that scientific opportunities coincidentally drop off at this point in time. Second, that the change in approval patterns arises in the year that our measure assigns as the expected point of market exclusivity expiration suggests that we are accurately capturing firm expectations.[72]

6 Quantification

Section 5 collects evidence from three empirical tests that, together, provide evidence in support of the existence of a missing market for new uses. This section leverages the elasticity of research investment with respect to market exclusivity calculated in Section 5.2 to develop a rough quantification of the social value of this missing innovation. To develop a back-of-the-envelope estimate of the number of missing new uses, we proceed in two steps. First, we estimate how many uses would have been realized, had innovator drugs received longer periods of market exclusivity. Second, we scale each missing use by an estimate of its social value drawn from the existing literature.

It is worth noting that this exercise is designed to take advantage of the elasticity recovered in Section 5.2 and, as a result, should be interpreted carefully. The thought experiment we present, in effect, asks how many new uses would have been commercialized had generic entry, *on all uses of a drug*, been delayed for some additional period. We do not undertake a complete welfare analysis, but note that the delays in generic entry on the scale considered below—5 to 20 years—are very costly and are not presented here as a practical policy solution. In the final section of the paper, we return to the question of how this problem of missing markets might be solved, in our empirical context and more generally.

6.1 How Many New Uses are Missing?

Figure 6 and Table 2 indicate that more new uses are commercialized when innovator drugs have longer periods of market exclusivity. While this establishes a causal relationship between exclusivity and investments in research and development, it leaves open an important question of magnitude: how many new uses would have been developed had patents on these inventions been fully enforceable?[73] As discussed in Section 3, the existing scientific literature provides little guidance on the

[72] The pattern we are documenting is consistent with practices that are often labeled as "product hopping" (Carrier and Shadowen, 2016), wherein firms respond to the loss of market exclusivity on one product by directing innovation toward products that are still protected by enforceable patents.This practice can impose real social costs. We do not undertake a complete welfare analysis of innovation decisions around exclusivity expiration, but note that any such analysis must take care to account for both incentives for innovation and the deadweight loss associated with delayed generic entry.

[73] It also leaves open a perhaps more important question, beyond the scope of this paper: how many new uses would have been developed had (enforceable) patent term been set optimally?

absolute number of potential new uses for each drug—other than suggesting that most existing drugs likely have many therapeutic targets.

To estimate the number of missing new uses, we will consider how many new uses would have been developed under counterfactual periods of market exclusivity.[74] As a lower estimate, suppose that the average innovator drug received a five-year exclusivity extension. Five-year extensions are relatively common: new chemical entities have received five-year market exclusivity grants since the 1984 Hatch Waxman Act, and certain new antibiotics now receive five-year grants as part of a package of 2012 reforms (Durvasula et al., 2023). As an upper estimate, suppose that a new use identified at the end of the innovator drug's period of exclusivity received its own fully enforceable patent term of twenty years. This counterfactual, which adds twenty years of exclusivity for the average innovator, approximates a scenario in which patents on later-discovered new uses are fully enforceable.

Estimates from Columns (2) and (4) of Table 2(A), imply that, counterfactually, 0.1 new indications would have been developed in the lower estimate case and 0.4 in the upper estimate case. Since 1962, when the modern FDA was established, the agency has approved roughly 2,000 drugs that satisfy our criteria for an innovator drug. Thus, our estimates indicate that between 200 and 800 new uses would have been developed had market exclusivity been extended along these lines.

6.2 How Valuable are New Uses?

The welfare implications of missing new uses depend on whether these re-approvals would have translated into real improvements in patient health. Though previous studies have concluded that, in general, new drugs significantly increase life expectancy, there is enormous heterogeneity in estimates of social value.[75] Philipson and Jena (2006) estimate, for example, that the social value of survival gains from breakthrough HIV/AIDS therapies in the late 1980s was $1.4 trillion over a twenty-year period in the United States alone, while other studies suggest that many "me-too" drugs—drugs that provide little clinical differentiation from existing therapies—contribute relatively little to social welfare (Dranove et al., 2014).

We survey the existing literature and select two complementary approaches to select an estimate for the average social value of a new drug. Noting the difficulty of this task, our aims are, both, to draw on credible sources and to select estimates that are relatively transparent. We then return to the question of whether, and how, this estimate ought to be scaled to apply to the case of new uses for existing drugs.

[74] By limiting ourselves to fixed market exclusivity extensions, we also need not make the assumption that innovator firms in our data could—with indefinite periods of exclusivity—ever identify every new indication for a drug. If ideas for new uses are "scarce" in the sense of Scotchmer (2004), then it is unlikely that the original firm would identify every possible lead.

[75] This reflects the inherent difficulty of this exercise. Trusheim et al. (2010) write, "[i]t is difficult if not impossible to quantify reliably, objectively and unambiguously the extent to which new biopharmaceuticals embody significant innovation and address unmet medical needs." This difficulty arises, in part, because drugs may be both scientific advancements—which offer little therapeutic impact for current cohorts of patients but enable follow-on innovation—and therapeutic advancements—which immediately improve patient health.

Our first approach makes use of drug quality ratings that value the health benefits of new therapeutics in terms of "quality adjusted life years" (QALYs). There are a variety of frameworks that are used to calculate QALYs but, in short, these measures aim to capture the value of both wellbeing improvements and increases in life expectancy in standard units. We use estimates from Shafrin et al. (2023) who consider all drugs approved by the FDA between 2011 and 2021 for which a published cost-effectiveness analysis was available—approximately 50% of approvals. Shafrin et al. (2023) estimate that the average drug added 1.04 QALYs per patient relative to the standard of care, though many drugs provided less than 0.5 QALYs per patient and a small number provided multiple QALYs.

We assume, on the basis of the Shafrin et al. (2023) analysis, that the average new drug adds one QALY relative to the standard of care. In the United States, cost-effectiveness analyses typically value each QALY at $100,000.[76] To generate an estimate of the average social value of a new drug, then, we need only scale this average QALY estimate, valued at $100,000, by a measure of average market size. For this, we select a value from industry reports: IQVIA Institute for Human Data Science (2021) indicates that a typical specialty drug is used to treat roughly 10,000 to 50,000 patients per year, with an average market size of 40,000. If we assume an incremental benefit of 0.5 to 1 QALY per patient per year, on average, the annual social value of an initial approval is between $2 billion (0.5 QALY $\times$ 40,000 $\times$ $100,000) and $4 billion (1.0 QALY $\times$ 40,000 $\times$ $100,000).

Our second approach draws, instead, on dollar-denominated estimates of the social value of new drugs, as reported in existing economic analyses. As with estimates of QALYs, there is enormous heterogeneity in these estimates. For example, Philipson and Jena (2006) estimate that the first class of initial antiretroviral drugs—the first treatments that prolonged the lives of those living with HIV and limited mother-to-child viral transmission—generated annual social value on the order of tens of billions of dollars; Yin et al. (2012) estimate that early tyrosine kinase inhibitors, which transformed the treatment of leukemia and other cancers, produced roughly $6 billion per year in the years following initial approval. On the low end of estimates, Sertkaya et al. (2014) report that new antibiotics yield annual social value between $24 million and $600 million. Unsurprisingly, these figures vary considerably across therapeutic areas. Generally, however, they suggest that innovative drugs generate annual social benefits in the range of hundreds of millions to several billion dollars per year.

The final question we must address is whether new uses of existing drugs are, on average, more or less valuable than initial approvals. Estimates from health services research typically find that new uses offer lower incremental therapeutic benefits than the use associated with an initial approval. A recent analysis of drug approvals in the United States and European Union from 2011 to 2020 found that roughly one-third of new uses had high therapeutic value, compared to nearly half of first indications (Vokinger et al., 2023). In oncology, first indications typically deliver nearly double the QALY gain per patient—around 0.99 QALYs—relative to second or third indications,

[76] Cutler (2004) also values life-years at $100,000, though Hemel and Ouellette (2023b) note that this estimate is "extraordinarily low relative to the values that federal agencies use in other contexts."

which average approximately 0.5–0.6 QALYs. Nonetheless, new uses can still generate substantial social value. Berger et al. (2021) show that approval of a new use increases, on average, the share of patients taking a drug by approximately 4.1 percentage points—a nearly 40% increase relative to baseline.[77] And Cohen et al. (2006) document that roughly 15 percent of drugs on the WHO Essential Medicines List are included because of a use approved after the drug's initial approval.

These measures suggest that a credible, if rough, range for the average social value of a new drug is between $500 million and $4 billion. To the extent that new uses provide less therapeutic benefit than initial approvals, we will select as our preferred estimate—$1 billion—a value on the lower end of this range.

6.3 Valuing the Missing Market for New Uses

To quantify the social cost of a missing market for new uses, we calculate the present value of a single missing new use, as

$$PV = \frac{\text{annual social value}}{r - g}$$

where r and g are the social discount rate and long-run population growth rate, respectively. We choose standard, conservative values for each and let $r = 0.05$ and $g = 0.01$. Using the estimate that an initial approval generates $1 billion in social value per year, on average, the present value of one missing approval is thus $25 billion.

Scaling this by our estimate of the number of missing new uses (200 to 800) implies that the social cost of a missing market for new uses is on the order of $5 to $20 trillion from 1962 to 2014. If new uses are roughly one-third to one-half as valuable as initial drug approvals, this range shifts accordingly—suggesting social costs instead on the order of approximately $2 to $10 trillion. Although these estimates are necessarily imprecise, the qualitative conclusion is clear: in just one market, the social costs of the missing markets problem appear to be extraordinarily large.

7 Discussion

In this paper, we investigate the idea that there are large classes of potential inventions for which private incentives to invest in research and development are missing. We develop a theoretical model that clarifies conditions under which these incentive gaps arise—when inventions are at risk of immediate imitation in the absence of intellectual property protection, but intellectual property rights are either unavailable or unenforceable—and predicts that *missing markets* will arise across sectors of the economy. In doing so, we formalize a set of ideas that have long been central to theoretical studies of innovation and intellectual property—that stronger and broader intellectual property

[77] As complementary evidence, we assess one measure of value—market size—in our own data, using data drawn from the Medical Expenditure Panel Survey on prescriptions written for diseases associated with initial approvals and new uses. n roughly one quarter (26%) of cases, new uses are approved for indications in different disease areas than the initial approval (where disease areas are defined broadly, e.g., ICD-10 Category G: Diseases of the Nervous System, Category J: Diseases of the Respiratory System). Although the exercise is necessarily rough, we find that, on average, the market size for a new use is roughly 70% that of each drug's initial indication.

rights induce additional investment in research and development, and that not all inventions are amenable to protection with any existing form of intellectual property. Our empirical work offers, to our knowledge, the first direct tests of both. We identify a setting, investments into new uses for existing drugs, in which there is sharp variation in the enforceability of intellectual property rights on otherwise comparable inventions over time. In this setting, intellectual property rights shift from perfectly enforceable to wholly unenforceable at a measurable point in time. We show that when enforceability drops to zero, so too does private investment. Plausibly exogenous variation in these periods of enforceability allows us to provide causal evidence that longer periods of (enforceable) intellectual property protection yield higher levels of research and development.

We use this elasticity of research and development with respect to intellectual property duration to estimate counterfactual changes in the development of new uses that would have occurred had intellectual property rights been fully enforceable. That is, we quantify the size and value of a missing market for new uses. We estimate that, since 1962, between 200 and 800 new uses would have been commercialized but for these gaps in intellectual property protection. Estimates from the scientific literature on the therapeutic potential of existing drugs suggest that this range is, if anything, conservative (Gelijns et al., 1998; Wermuth, 2006). Using measures of the value of a new use drawn from the literature, we estimate that the social cost of this *particular* missing market is on the order of $2 to $10 trillion in present value terms. Although these estimates are necessarily rough, they suggest that the social costs of the missing markets problem, in general, may be extraordinarily large.

The potential scale of this problem makes more salient the question of solutions. Is it possible to "fix" a missing market? Our theoretical model rules out one standard tool, patent policy, deployed by policymakers to correct inadequate incentives for innovation. Necessarily, neither longer nor broader periods of formal intellectual property protection—what we term "Nordhaus parameters"—affect incentives to develop potential inventions in missing markets. The model does, however, highlight two alternatives. First, a policymaker could attempt to change the parameters that give rise to a missing market: p_{eligible} and p_{enforce}. Second, missing markets might be natural targets for funding mechanisms that detach innovation incentives from private-sector profitability, which include both traditional forms of public subsidies and innovative mechanisms like prizes and advanced market commitments.

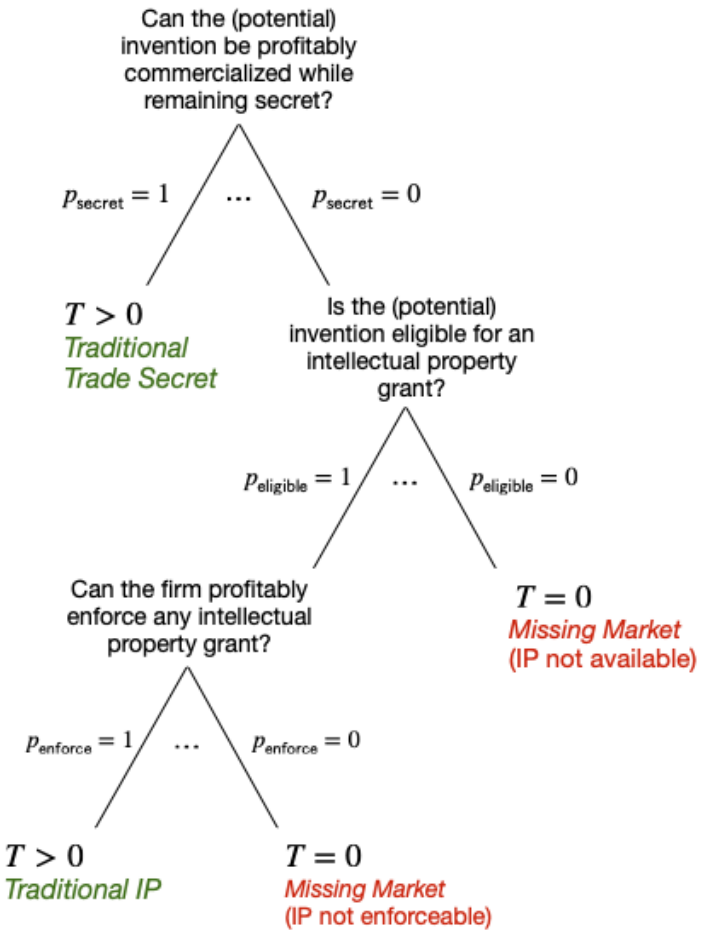
With respect to the first alternative, changes to p_{eligible} are an ongoing topic of public debate (Critharis, 2022). This paper provides one piece of empirical evidence that informs these debates by documenting a link between intellectual property protection and the level of investment in innovation, at least in the context of pharmaceutical innovation. But any change to the expansiveness of intellectual property rights would affect all inventions, across all sectors; the dearth of empirical evidence on other key questions about the relationships between intellectual property rights, investments in innovation, and consumer access makes it challenging to confidently assess large-scale changes at present. One can imagine, however, a variety of context-specific solutions to the problem of enforceability ($p_{\mathsf{enforce}} = 0$). Advances in technology may allow firms to observe instances

in which their intellectual property rights are infringed. In the context of new uses, integrated electronic health record systems can solve at least part of the enforceability problem; Roin (2014) sketches how this data infrastructure might be constructed, and Pearson et al. (2016) describes one attempt to implement such a system for cancer therapies in Italy.

More generally, this paper provides empirical support for the importance of innovation policy levers beyond intellectual property. Theoretical advances in the design of these incentives over the past two decades (see Barder et al. (2005) and Kremer et al. (2020) for two summaries) provide templates for immediate policy interventions in the markets identified in Section 2. In particular, this literature makes clear that non-patent approaches to research incentives are viable, whether missing investments are "technologically distant"—like vaccines for diseases where even the basic science is poorly understood—or "technologically close"—like new uses for existing drugs where the missing piece is financing for a large-scale clinical trial. This paper estimates that the value of better designed incentives in just a single market, new uses for existing drugs, is in the trillions. One struggles to fully comprehend the scale of potential welfare gains from extending meaningful incentives across the full range of missing markets for innovation.

Figures and Tables

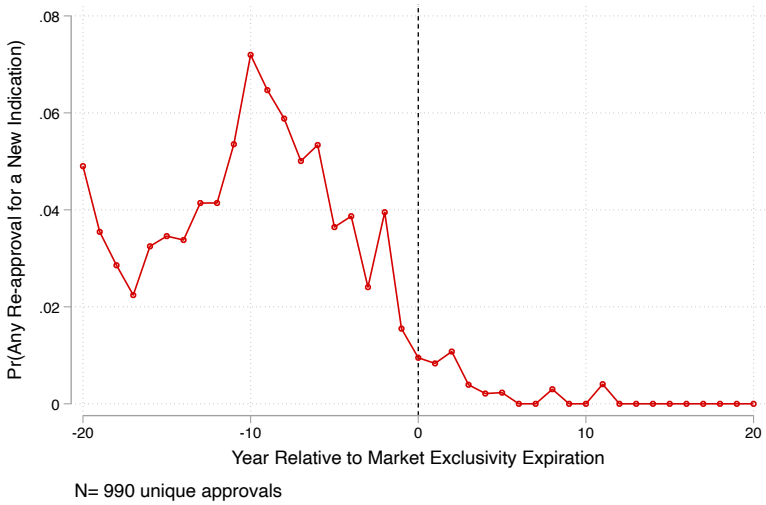
Figure 1: Missing Markets for Innovation



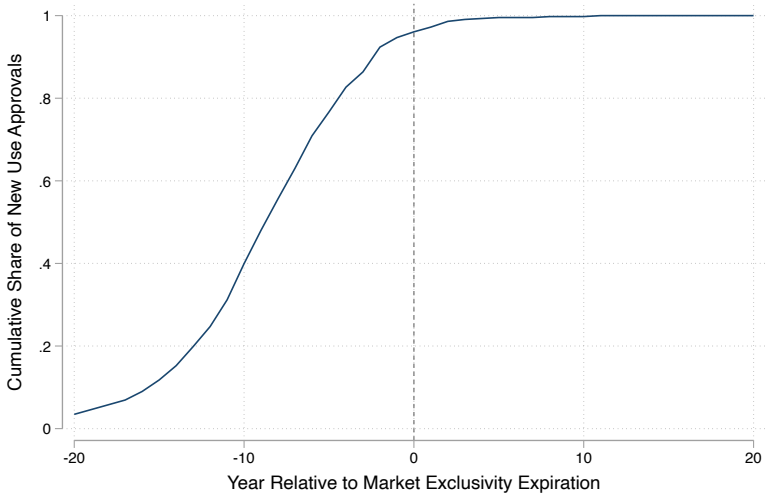
Notes: Figure summarizes the framework introduced in Section 2.1.

Figure 2: New Use Approvals Relative to Exclusivity Expiration

A. Likelihood of New Use Approvals



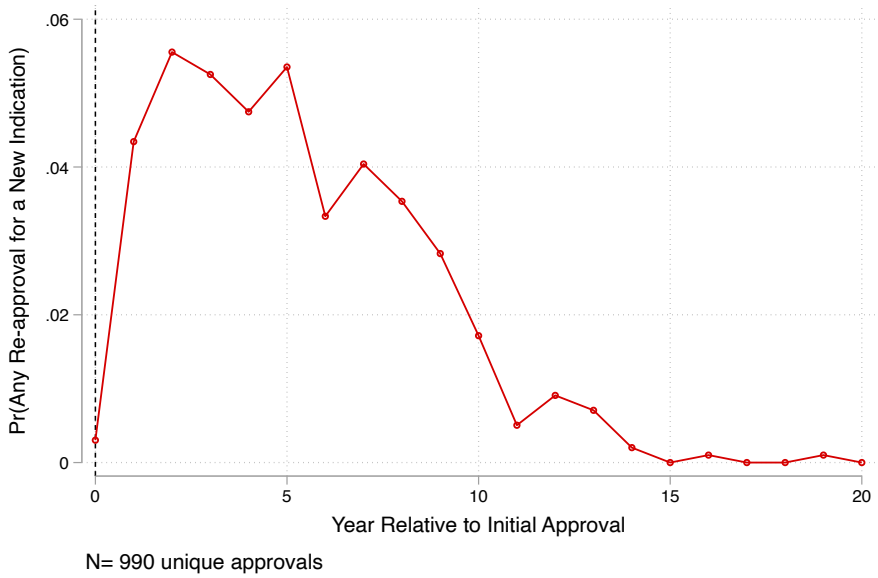
B. Cumulative Share of New Use Approvals



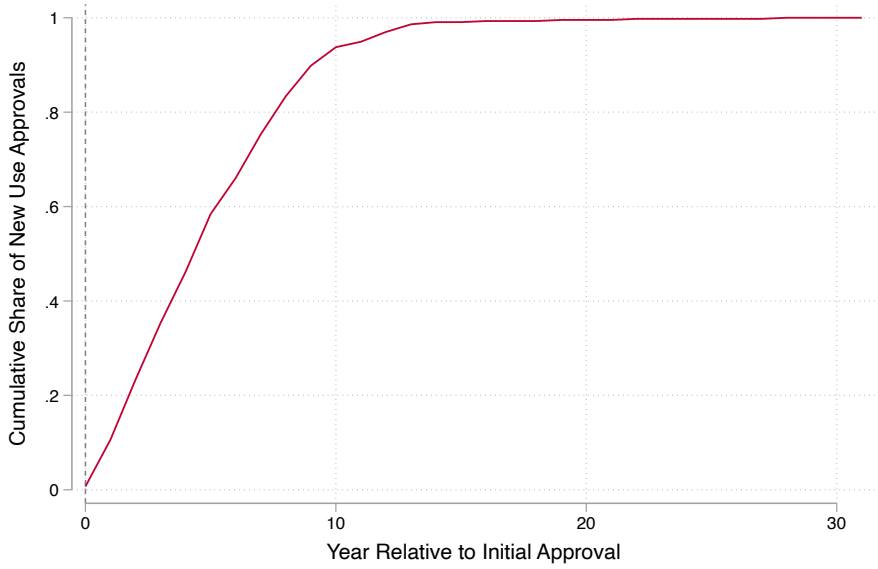
Notes: Panel A plots the probability of re-approval for a new use for the sample of innovator drugs defined in Section 4.5 ($N = 990$). Panel B plots the cumulative share of new use approvals by each year for the same sample. Both panels plot trends relative to market exclusivity expiration. Market exclusivity expiration is defined as the earlier of two events: expiration of all patents and regulatory exclusivities linked to the innovator drug in the FDA's Orange Book, or FDA approval of a generic drug with an identical set of active moieties. We plot the probability of re-approval for a new indication for the twenty years preceding and following exclusivity expiration for all drugs in our sample.

Figure 3: New Use Approvals Relative to Initial Approval

A. Likelihood of New Use Approvals

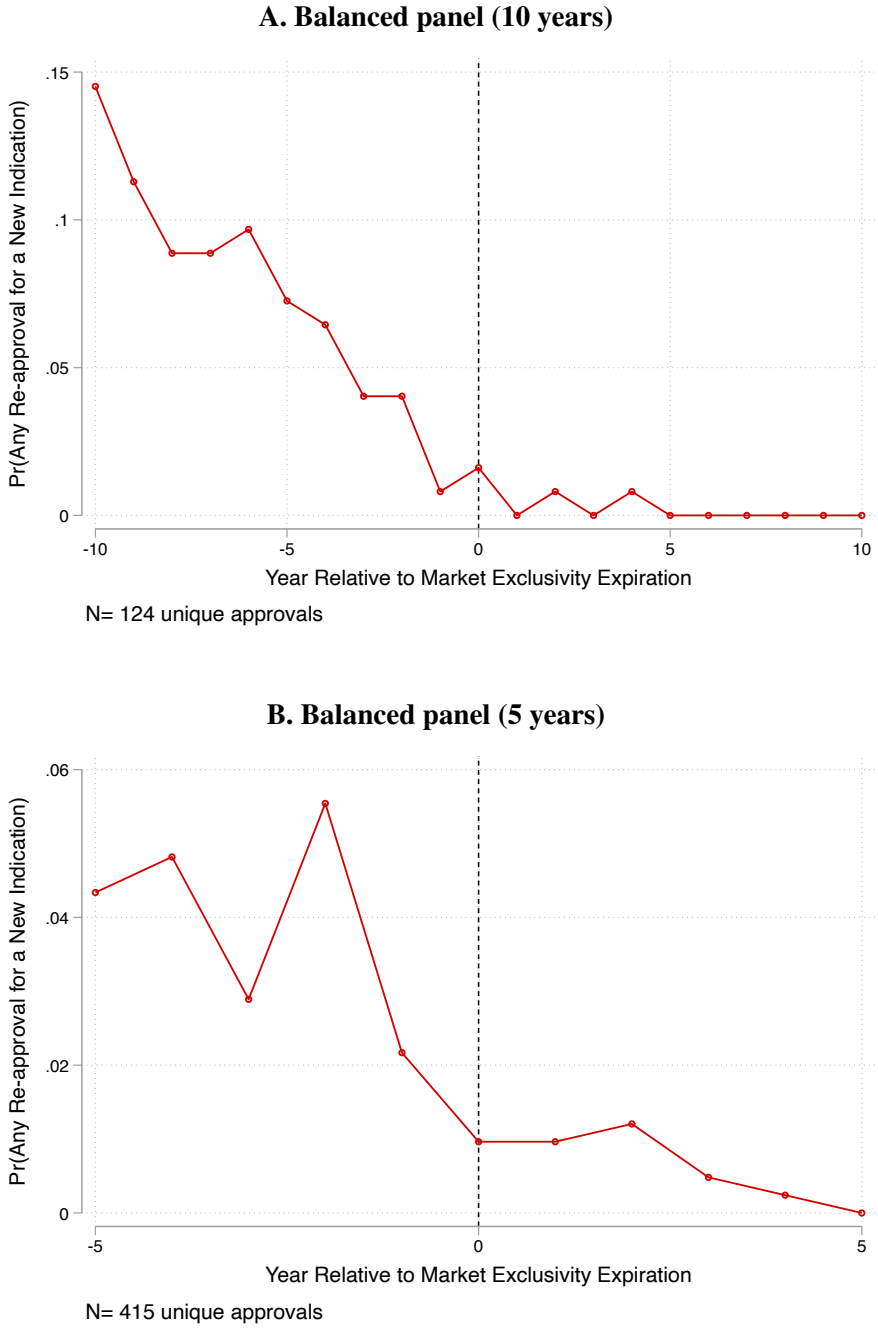


B. Cumulative Share of New Use Approvals



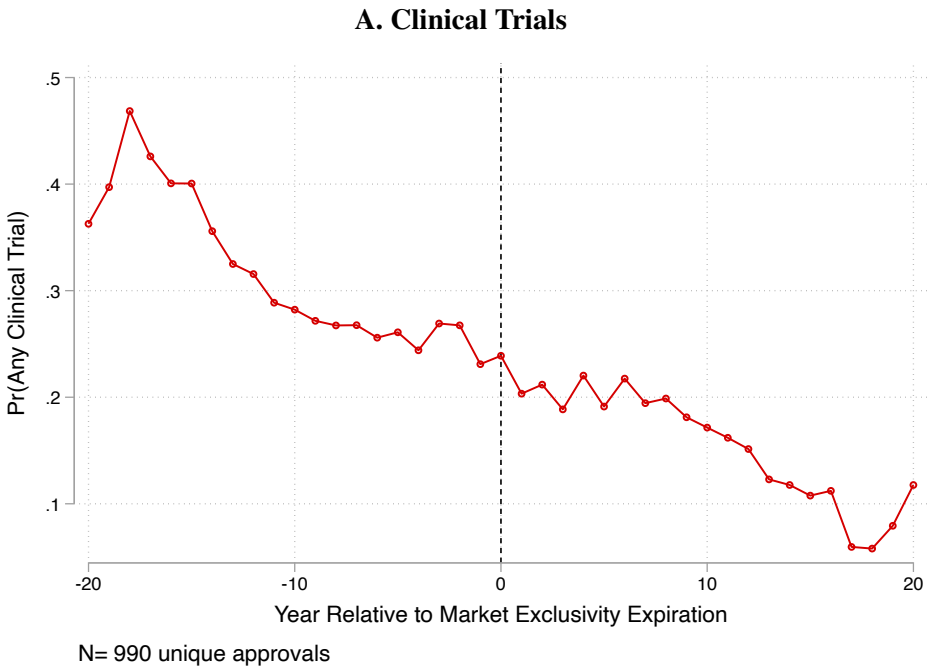
Notes: Panel A plots the probability of re-approval for a new use for the sample of innovator drugs defined in Section 4.5 ($N = 990$). Panel B plots the cumulative share of new use approvals by each year for the same sample. Both panels plot trends relative to the drug's initial approval by the FDA.

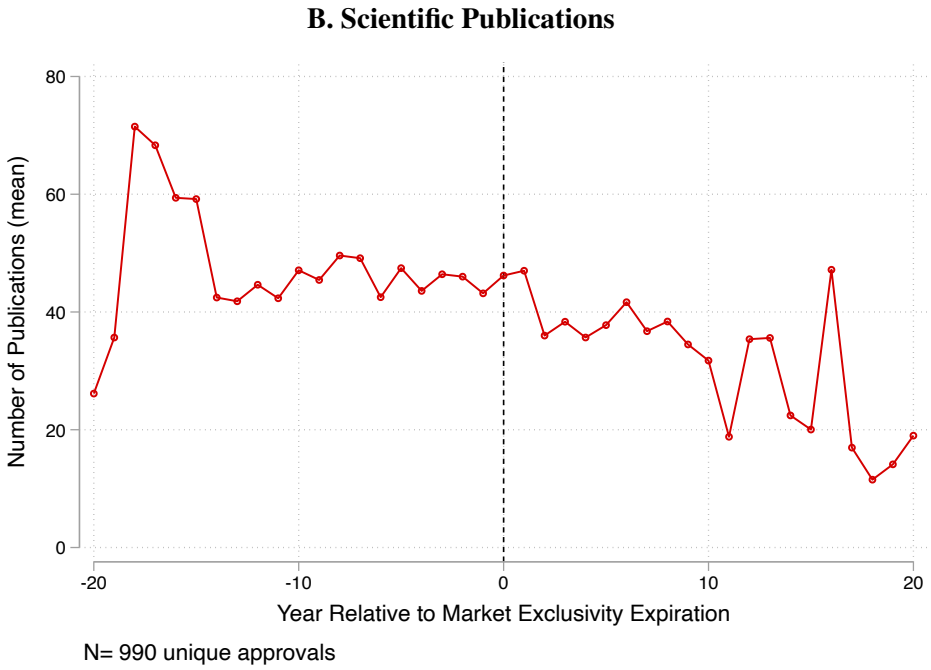
Figure 4: New Use Approvals, Balanced Panel



Notes: Figure plots the probability of re-approval for a new use for the sample of innovator drugs defined in Section 4.5. Panel A is a ten-year balanced panel ($N = 124$). Panel B is a five-year balanced panel ($N = 415$). Market exclusivity expiration is defined as the earlier of two events: expiration of all patents and regulatory exclusivities linked to the innovator drug in the FDA's Orange Book, or FDA approval of a generic drug with an identical set of active moieties.

Figure 5: Other Research Investments Relative to Exclusivity Expiration

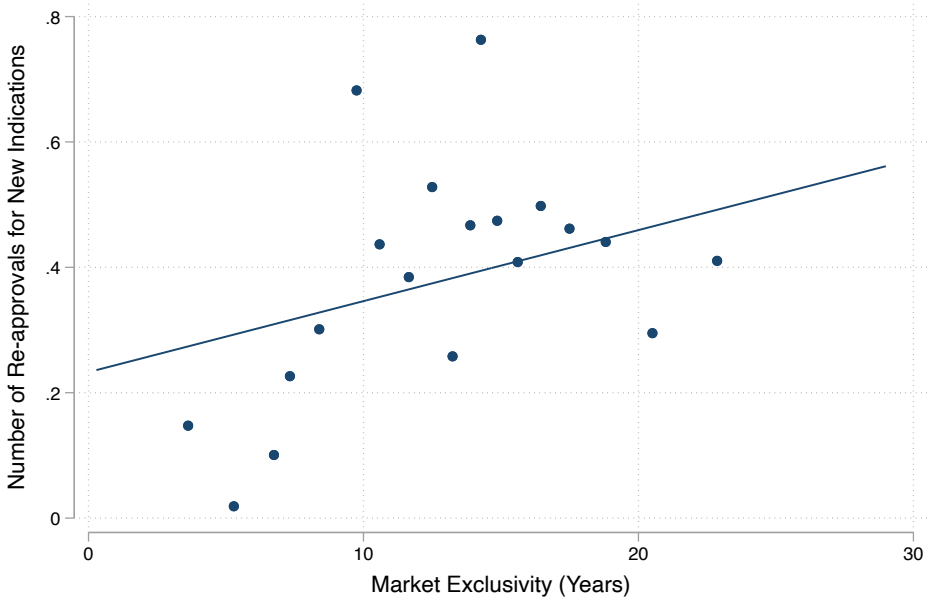




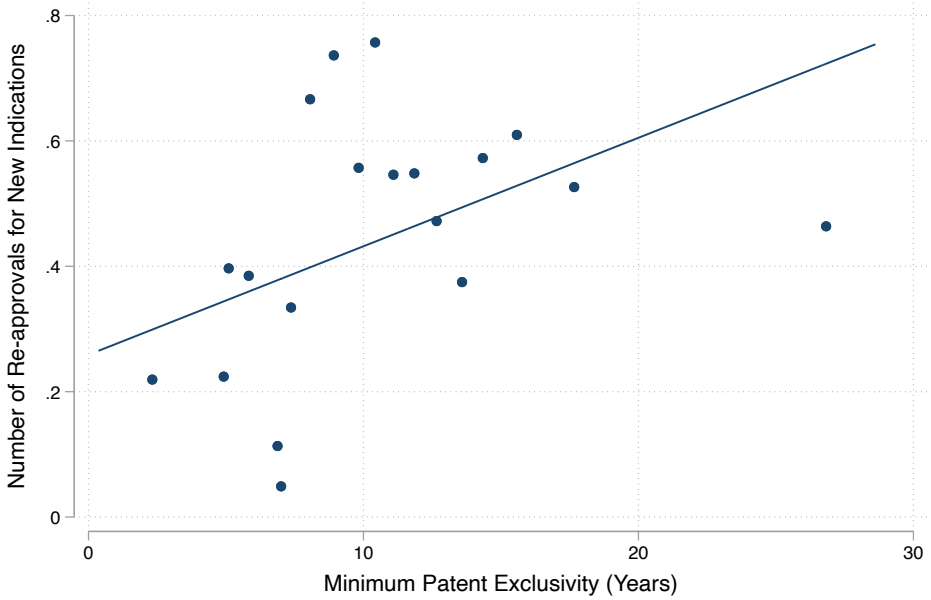
Notes: Panel A plots the probability of any active clinical trial in each relative year for the sample of innovator drugs defined in Section 4.5 ($N = 990$). We plot the probability of any active clinical trial in each of the twenty years preceding and following exclusivity expiration for all drugs in our sample. Panel B plots the average number of scientific publications published in each relative year for the same sample. We plot the average number of scientific publications indexed in PubMed associated with the innovator drug in the twenty years preceding and following exclusivity expiration for all drugs in our sample. Market exclusivity expiration is defined as the earlier of two events: expiration of all patents and regulatory exclusivities linked to the innovator drug in the FDA's Orange Book, or FDA approval of a generic drug with an identical set of active moieties. Appendix Figures A4 and A4 plot trends in clinical trials and publications, respectively, across the drug's lifecycle (year relative to initial approval).

Figure 6: Quantity of New Uses by Exclusivity Duration

Panel A: Market Exclusivity

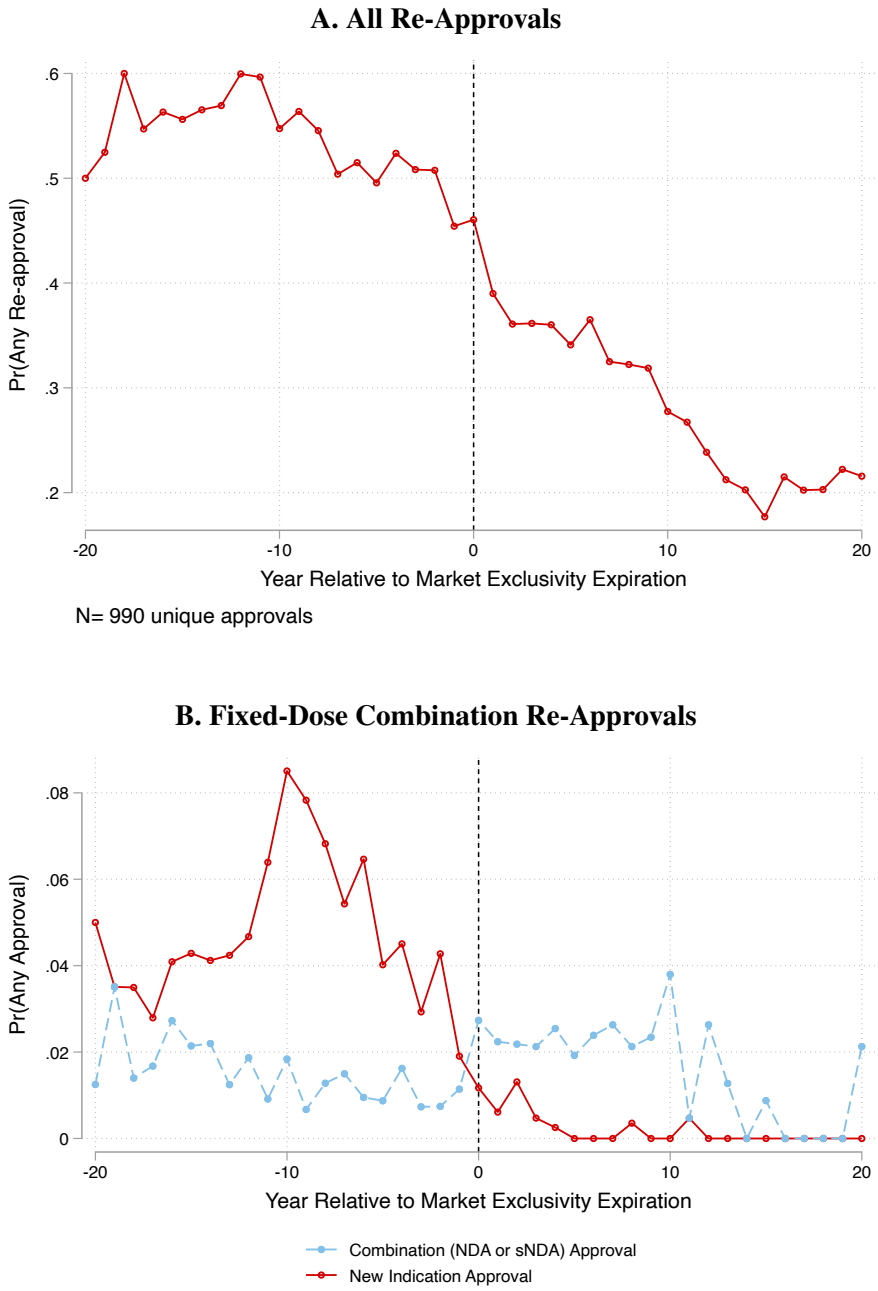


Panel B: Minimum Patent Exclusivity



Notes: Figure plots the relationship between a measure of market exclusivity (x-axis) and the number of re-approvals for new uses (y-axis) for the sample of innovator drugs defined in Section 4.5 ($N = 990$). Both the x- and y-axes are residualized by the year of innovator drug approval, as year of approval is mechanically related to the total number of realized new uses. The scatterplot is binned into ventiles of market exclusivity (20 bins), with the mean value of the exclusivity measure added back to facilitate interpretation. A line of best fit is included. Panel A uses a measure of exclusivity that captures the number of years between the drug's initial approval and the earlier of the expiration of all patents and regulatory exclusivities or generic approval. Panel B uses a measure of exclusivity that captures the number of years between the drug's initial approval and expiration of the earliest expiring patent and all regulatory exclusivities.

Figure 7: Other Re-Approvals Relative to Exclusivity Expiration



Notes: Panel A plots the probability of any re-approval (e.g., new use, new dosage form, etc.) for the sample of innovator drugs defined in Section 4.5 ($N = 990$). Panel B plots two measures of FDA approvals for a subset of innovator drugs with a single active moiety ($N = 792$): re-approval for a new indication and re-approval as a combination drug. In Panel B, we impute zeroes in each relative year if a single-moiety innovator drug never appears in a combination drug. Market exclusivity expiration is defined as the earlier of two events: expiration of all patents and regulatory exclusivities linked to the innovator drug in the FDA's Orange Book, or FDA approval of a generic drug with an identical set of active moieties.

Table 1: Summary Statistics

	Mean	Median	SD	Min.	Max.	# obs.
Years to Generic Entry	11.27	11.47	4.52	0.21	28.19	442
Market Exclusivity (years)	12.90	13.37	5.48	0.18	42.46	990
Minimum Patent Exclusivity (years)	9.35	8.37	4.55	0.18	35.32	956
Number of Re-Approvals	14.87	11	13.43	0	87	990
Number of Re-Approvals (new/modified indication)	0.52	0	1.18	0	10	990

Notes: Table shows summary statistics for the sample of innovator drugs defined in Section 4.5 ($N = 990$). Innovator drugs are defined as the first FDA approvals of unique combinations of active moieties.

Table 2: Market Exclusivity and New Indication Approvals

Panel A: *re-approvals for new/modified indication*

	(1)	(2)	(3)	(4)
Market Exclusivity Exclusivity	0.013	0.020	—	—
(years)	(0.006)	(0.006)		
Minimum Patent Exclusivity	—	—	0.037	0.023
(years)			(0.008)	(0.008)
Approval Year FEs	no	yes	no	yes
Observations	990	990	990	990
Panel B: Two-part model				
Extensive Margin	*any re-approvals for new/modified indication*			
Market Exclusivity Exclusivity	0.026	0.042	—	—
(years)	(0.012)	(0.014)		
Minimum Patent Exclusivity	—	—	0.083	0.059
(years)			(0.015)	(0.018)
Approval Year FEs	no	yes	no	yes
Observations	990	990	990	990
Intensive Margin	*# re-approvals for new/modified indication*			
Market Exclusivity Exclusivity	0.015	0.020	—	—
(years)	(0.017)	(0.018)		
Minimum Patent Exclusivity	—	—	0.017	0.011
(years)			(0.021)	(0.022)
Approval Year FEs	no	yes	no	yes
Observations	256	256	256	256

Notes: Table shows the relationship between measures of market exclusivity and the number of realized new uses, estimating 7 in the text. Observations are the sample of innovator drugs defined in Section 4.5. We report estimates using two measures of exclusivity: "market exclusivity" captures the number of years between the drug's initial approval and the earlier of the expiration of all patents and regulatory exclusivities or generic approval, and "minimum patent exclusivity" captures the number of years between the drug's initial approval and expiration of the earliest expiring patent and all regulatory exclusivities. Panel A reports estimates for a pooled sample, estimated via ordinary least squares. Panel B reports estimates from a two-part model, as described in Section 5.2. All specifications include approval year fixed effects.

References

Abramowicz, Michael, "The danger of underdeveloped patent prospects," *Cornell Law Review*, 2007, *92* (6), 1065–1121.

Abrams, David, "Did TRIPS spur innovation? An analysis of patent duration and incentives to innovate," *University of Pennsylvania Law Review*, 2009, *157* (6), 1613–1647.

Arrow, Kenneth, "Economic Welfare and the Allocation of Resources to Invention," in Universities-National Bureau Committee for Economic Research and the Committee on Economic Growth of the Social Science Research Councils, eds., *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton, NJ: Princeton University Press, 1962, pp. 609–626.

Barder, Owen, Michael Kremer, and Ruth Levine, "Answering concerns about Making Markets for Vaccines," *Geneva: Commission on Intellectual Property Rights, Innovation and Public Health (CIPIH). Available: http:// www. who. int/intellectualproperty/submissions/BarderSubmission. pdf. Accessed*, 2005, *15*.

Berger, Benjamin, Amitabh Chandra, and Craig Garthwaite, "Regulatory approval and expanded market size," Technical Report, National Bureau of Economic Research 2021.

Bronnenberg, Bart J, Jean-Pierre Dubé, Matthew Gentzkow, and Jesse M Shapiro, "Do pharmacists buy Bayer? Informed shoppers and the brand premium," *The Quarterly Journal of Economics*, 2015, *130* (4), 1669–1726.

Budish, Eric, Benjamin N Roin, and Heidi L Williams, "Do firms underinvest in long-term research? Evidence from cancer clinical trials," *American Economic Review*, 2015, *105* (7), 2044–2085.

_ , _ , and _ , "Patents and research investments: Assessing the empirical evidence," *American Economic Review*, 2016, *106* (5), 183–187.

Butcher, Nancy, "Old Drugs and New Tricks: Repurposing Drugs to Treat Psychiatric Disorders," *IMS Magazine*, 2013, *1*, 21–21.

Carr, Austin and Ann Choi, "The Plot to Steal the Secret Coke Can Liner Formula," May 2023. Accessed: March 10, 2025.

Carrier, Michael A and Steve D Shadowen, "Product hopping: A new framework," *Notre Dame L. Rev.*, 2016, *92*, 167.

Cohen, Joshua, L. Cabanilla, and J. Sosnov, "Role of follow-on drugs and indications on the WHO Essential Drug List," *Journal of Clinical Pharmacy and Therapeutics*, 2006, *31* (6), 585–592.

Cohen, Wesley, Richard Nelson, and John Walsh, "Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not)," 2000. Published: National Bureau of Economic Research (NBER) working paper 7552.

Collins, Francis, "Testimony before the US House Subcommittee on Labor - HHS - Education Appropriations," 2012. Published: Department of Health and Human Services, National Institutes of Health.

Collins, Francis S, "Mining for therapeutic gold," *Nature Reviews Drug Discovery*, 2011, *10* (6), 397.

_ , "Reengineering translational science: the time is right," *Science translational medicine*, 2011, *3* (90), 90cm17–90cm17. Publisher: American Association for the Advancement of Science.

Collins, Francis S. and Harold Varmus, "A New Initiative on Precision Medicine," *New England Journal of Medicine*, 2015, *372* (9), 793–795.

Corbett, Anne, James Pickett, Alistair Burns, Jonathan Corcoran, Stephen B Dunnett, Paul Edison, Jim J Hagan, Clive Holmes, Emma Jones, Cornelius Katona et al., "Drug repositioning for Alzheimer's disease," *Nature reviews Drug discovery*, 2012, *11* (11), 833–846.

Critharis, Mary, "Report to Congress on Patent Eligible Subject Matter: Public Views on the Current Jurisprudence in the United States," June 2022. Accessed: 2025-03-19.

Cutler, David, *Your Money or Your Life: Strong Medicine for America's Health Care System*, Oxford University Press, 2004.

Dix, Rebekah and Todd Lensman, "Combining Complements: Theory and Evidence from Cancer Treatment Innovation," 2024.

Dranove, David, Craig Garthwaite, and Manuel Hermosilla, "Pharmaceutical profits and the social value of innovation," Technical Report, National Bureau of Economic Research 2014.

Dudley, Joel T, Tarangini Deshpande, and Atul J Butte, "Exploiting drug–disease relationships for computational drug repositioning," *Briefings in bioinformatics*, 2011, *12* (4), 303–311.

Durvasula, Maya and Lisa Larrimore Ouellette, "Beyond the Pharmaceutical Patent Arms Race," *Yale Journal on Regulation*, 2026.

_ **, C Scott Hemphill, Lisa Larrimore Ouellette, Bhaven Sampat, and Heidi L Williams**, "The NBER Orange Book dataset: A user's guide," *Research Policy*, 2023, *52* (7), 104791.

Eckhause, Melissa, "Fighting Image Piracy or Copyright Trolling? An Empirical Study of Photography Copyright Infringement Lawsuits," *Alb. L. Rev.*, 2022, *86*, 111.

Eguale, Tewodros, David L Buckeridge, Nancy E Winslade, Andrea Benedetti, James A Hanley, and Robyn Tamblyn, "Drug, patient, and physician characteristics associated with off-label prescribing in primary care," *Archives of Internal Medicine*, 2012, *172* (10), 781–788.

Eisenberg, Rebecca, "The problem of new uses," *Yale Journal of Health Policy, Law, & Ethics*, 2005, *717* (5), 717–739.

EveryCure, "EveryCure: Unlocking the Full Potential of Existing Drugs," 2025.

FDA, "List of Patent and Exclusivity for Drugs without Approved Generic," 2025. Accessed: 2025-03-14.

Frakes, Michael D. and Melissa F. Wasserman, "Strategic Patenting: Evidence from the Biopharmaceutical Industry," 2025.

Gale, Jason and Andrea Gerlin, "Five-cent diabetes pill from 1958 may be new cancer drug," *Bloomberg Business*, 2012, *26 September*.

Gawande, Atul, "The checklist manifesto: How to get things right," 2009.

Gelijns, Annetine, Nathan Rosenberg, and Alan Moskowitz, "Capturing the unexpected benefits of medical research," *New England Journal of Medicine*, 1998, *339* (10), 693–697.

Gilbert, Richard and Carl Shapiro, "Optimal patent length and breadth," *RAND Journal of Economics*, 1990, *21* (1), 106–112.

Gilchrist, Duncan, "Patents as a Spur to Subsequent Innovation? Evidence from Pharmaceuticals," *American Economic Journal: Applied Economics*, 2016, *8* (4), 189–221.

Hemel, Daniel J and Lisa Larrimore Ouellette, "The generic drug trilemma," *Entrepreneurship and Innovation Policy and the Economy*, 2023, *2* (1), 41–77.

_ and _, "Valuing medical innovation," *Stan. L. Rev.*, 2023, *75*, 517.

Hemphill, C Scott, "Paying for delay: Pharmaceutical patent settlement as a regulatory design problem," *NYUL Rev.*, 2006, *81*, 1553.

Hemphill, C. Scott and Bhaven Sampat, "Evergreening, patent challenges, and effective market life in pharmaceuticals," *Journal of Health Economics*, 2011, *31* (2), 327–339.

Hemphill, Scott and Bhaven Sampat, "Evergreening, Patent Challenges, and Effective Market Life in Pharmaceuticals," *Journal of Health Economics*, 2012, *31*, 327–339.

Hertel, Larry W, George B Boder, J Stan Kroin, Sharon M Rinzel, Gerald A Poore, Glen C Todd, and Gerald B Grindey, "Evaluation of the antitumor activity of gemcitabine (2', 2'-difluoro-2'-deoxycytidine)," *Cancer research*, 1990, *50* (14), 4417–4422.

Hodgson, Charles, "Information externalities, free riding, and optimal exploration in the uk oil industry," Technical Report, National Bureau of Economic Research 2024.

Hoyt, A Sasha, "The impact of uncertainty regarding patent eligible subject matter for investment in US medical diagnostic technologies," *Wash. & Lee L. Rev.*, 2022, *79*, 397.

IQVIA Institute for Human Data Science, "The Use of Medicines in the U.S.," 2021. Accessed: 2025-03-19.

Johnson, Matthew S, Michael Lipsitz, and Alison Pei, "Innovation and the enforceability of noncompete agreements," Technical Report, National Bureau of Economic Research 2023.

Jones, Charles I, "Growth and ideas," in "Handbook of economic growth," Vol. 1, Elsevier, 2005, pp. 1063–1111.

_ , "Recipes and economic growth: A combinatorial march down an exponential tail," *Journal of Political Economy*, 2023, *131* (8), 1994–2031.

Kapczynski, Amy and Talha Syed, "The continuum of excludability and the limits of patents," *Yale LJ*, 2012, *122*, 1900.

_ , Chan Park, and Bhaven Sampat, "Polymorphs and prodrugs and salts (oh my!): An empirical analysis of 'secondary' pharmaceutical patents," *PLoS One*, 2012, *7* (12), e49470.

Klemperer, Paul, "How broad should the scope of patent protection be?," *RAND Journal of Economics*, 1990, *21* (1), 113–130.

Kremer, Michael, Jonathan Levin, and Christopher M Snyder, "Advance market commitments: insights from theory and experience," in "AEA Papers and Proceedings," Vol. 110 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2020, pp. 269–273.

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

Lemley, Mark A and Carl Shapiro, "Probabilistic patents," *Journal of Economic Perspectives*, 2005, *19* (2), 75–98.

Lerner, Josh, "Patent protection and innovation over 150 years," 2002. Published: National Bureau of Economic Research (NBER) working paper 8977.

Levin, Richard, Alvin Klevorick, Richard Nelson, Sidney Winter, Richard Gilbert, and Zvi Griliches, "Appropriating the returns from industrial research and development," *Brookings Papers on Economic Activity*, 1987, *1987* (3), 783–831.

Lord, Simon R and Adrian L Harris, "Is it still worth pursuing the repurposing of metformin as a cancer therapeutic?," *British journal of cancer*, 2023, *128* (6), 958–966.

Machlup, Fritz, "An economic review of the patent system," 1958. Published: Study of the Subcommittee on Patents, Trademarks, and Copyrights of the Committee on the Judiciary, United States Senate, 85th Congress, Second Session.

Mansfield, Edwin, "Patents and innovation: An empirical study," *Management Science*, 1986, *32* (2), 173–181.

_ , Mark Schwartz, and Samuel Wagner, "Imitation costs and patents: An empirical study," *The Economic Journal*, 1981, *91* (364), 907–918.

Michelman, Valerie and Lucy Msall, "Sex, Drugs, and R&D: Missing Innovation from Regulating Female Enrollment in Clinical Trials," Technical Report, University of Chicago Working paper 2024.

Morton, Fiona M Scott, "Entry decisions in the generic pharmaceutical industry," *The Rand journal of economics*, 1999, pp. 421–440.

Moser, Petra, "How do patent laws influence innovation? Evidence from nineteenth-century world fairs," *American Economic Review*, 2005, *95* (4), 1214–1236.

_ , "Why don't inventors patent?," 2007.

Mossinghoff, Gerald, "Overview of the Hatch-Waxman Act and its impact on the drug development process," *Food & Drug Law Journal*, 1999, *54* (2), 187–194.

Mozaffarian, Dariush, Tao Hao, Eric B Rimm, Walter C Willett, and Frank B Hu, "Changes in diet and lifestyle and long-term weight gain in women and men," *New England journal of medicine*, 2011, *364* (25), 2392–2404.

Mullahy, John, "Much ado about two: reconsidering retransformation and the two-part model in health econometrics," *Journal of health economics*, 1998, *17* (3), 247–281.

Muthyala, Ramaiah, "Orphan/rare drug discovery through drug repositioning," *Drug Discovery Today: Therapeutic Strategies*, 2011, *8* (3-4), 71–76.

Nelson, Richard R, "The simple economics of basic scientific research," *Journal of political economy*, 1959, *67* (3), 297–306.

Nordhaus, William, *Invention, Growth, and Welfare: A Theoretical Treatment of Technological Change*, Cambridge, MA: MIT Press, 1969.

_ , "The optimum life of a patent: Reply," *American Economic Review*, 1972, *62* (3), 428–431.

Oster, Emily, "Why Is Nutrition So Stressful?," ParentData Podcast Dec. 2024. Updated Jan. 3, 2025.

Pearson, Steven D., Bill Dreitlein, and Chris Henshall, "Indication-specific Pricing of Pharmaceuticals in the United States Health Care System: A Report from the 2015 ICER Membership Policy Summit," Technical Report, Institute for Clinical and Economic Review March 2016.

Philipson, Tomas and Anupam Jena, "Who benefits from new medical technologies? Estimates of consumer and producer surpluses for HIV/AIDS drugs," *Forum for Health Economics & Policy*, 2006, *9* (2).

Pollack, Andrew, "Debate Over AIDS Drug's Effect on H.I.V. Infection Rate," July 2006. Accessed: March 10, 2025.

Price, W Nicholson, Arti K Rai, and Timo Minssen, "Knowledge transfer for large-scale vaccine manufacturing," *Science*, 2020, *369* (6506), 912–914.

Price, W Nicholson II, "Making do in making drugs: innovation policy and pharmaceutical manufacturing," *BCL Rev.*, 2014, *55*, 491.

Rai, Arti, "Use patents, carve-outs, and incentives - A new battle in the drug-patent wars," *New England Journal of Medicine*, 2012, *367* (6), 491–493.

Reinmuth, Kate and Emma Rockall, "Innovation through labor mobility: Evidence from non-compete agreements," *Available at SSRN 4459683*, 2023.

Robbins, Christopher, "Fake Luxury Goods: Inside the Secret World of Counterfeit Handbags," August 2018. Accessed: March 10, 2025.

Roin, Benjamin N, "The case for tailoring patent awards based on time-to-market," *UCLA L. Rev.*, 2013, *61*, 672.

_ , "Solving the problem of new uses by creating incentives for private industry to repurpose off-patent drugs," 2014. Published: mimeo, MIT Sloan School of Management.

Romer, Paul M, "The origins of endogenous growth," *Journal of Economic perspectives*, 1994, *8* (1), 3–22.

Sahoo, Alison, *Indication expansion: opportunities for successful lifecycle management*, Business Insights Limited, 2007.

Saiyed, Mohammad M, Pei-Shi Ong, and Lita Chew, "Off-label drug use in oncology: a systematic review of literature," *Journal of clinical pharmacy and therapeutics*, 2017, *42* (3), 251–258.

Sakakibara, Mariko and Lee Branstetter, "Do stronger patents induce more innovation? Evidence from the 1998 Japanese patent law reforms," *RAND Journal of Economics*, 2001, *32* (1), 77–100.

Santo, Brian, "The Consumer Electronics Hall of Fame: Zojirushi Micom Electric Rice Cooker/Warmer," September 2023. Accessed: March 10, 2025.

Scherer, F. M., "Nordhaus' theory of optimal patent life: A geometric reinterpretation," *American Economic Review*, 1972, *62* (3), 422–427.

Scotchmer, Suzanne, "Standing on the shoulders of giants: Cumulative research and the patent law," *Journal of Economic Perspectives*, 1991, *5* (1), 29–41.

_ , *Innovation and incentives*, MIT press, 2004.

Sertkaya, Aylin, John T Eyraud, Anna Birkenbach, Calvin Franz, Nyssa Ackerley, Valerie Overton, and Kevin Outterson, "Analytical framework for examining the value of antibacterial products," *Submitted to the US Department of Health and Human Services*, 2014, pp. 14–25.

Shafrin, Jason, Sabiha Quddus, Moises Marin, and Dennis Scanlon, "A decade of health innovation: the impact of new medicines on patient health and the implications for NICE's size of benefit multiplier," *Value in Health*, 2023, *26* (10), 1435–1439.

Silverman, Ed, "Single Tablet Regimen: How Gilead Scientists Developed a Pill That Changed the Course of HIV Prevention," November 2018. Accessed: March 10, 2025.

Singh, Jasjit and Matt Marx, "The Geographic Scope of Knowledge Spillovers: Spatial Proximity, Political Borders and Non-Compete Enforcement," 2011.

Teece, David J, "Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy," *Research policy*, 1986, *15* (6), 285–305.

Telleria, Carlos M, "Drug repurposing for cancer therapy," *Journal of cancer science & therapy*, 2012, *4* (7), ix.

Trusheim, Mark, Murray Aitken, and Ernst Berndt, "Characterizing markets for biopharmaceutical innovations: Do biologics differ from small molecules?," *Forum for Health Economics and Policy*, 2010, *13* (1), 4.

Tse, Tony, Kevin M Fain, and Deborah A Zarin, "How to avoid common problems when using ClinicalTrials. gov in research: 10 issues to consider," *Bmj*, 2018, *361*.

US Institute of Medicine, *Improving the quality of cancer clinical trials: Workshop summary*, The National Academies Press, 2008.

USPTO, "Patent Term Adjustment," 2025. Accessed: 2025-03-14.

Vokinger, Kerstin N, Camille EG Glaus, Aaron S Kesselheim, Miquel Serra-Burriel, Joseph S Ross, and Thomas J Hwang, "Therapeutic value of first versus supplemental indications of drugs in US and Europe (2011-20): retrospective cohort study," *bmj*, 2023, *382*.

Ward, Erin H, "Defining Active Ingredient: The US Food and Drug Administration's Legal Interpretation of Regulatory Exclusivities.," *Congressional Research Service (CRS) Reports and Issue Briefs*, 2019.

Wermuth, Camille G., "Selective Optimization of Side Activities: the SOSA Approach," *Drug Discovery Today*, 2006, *11*, 160–161.

Yin, Wesley, John R Penrod, Ross Maclean, Darius N Lakdawalla, and Tomas Philipson, "Value of survival gains in chronic myeloid leukemia," *Am J Manag Care*, 2012, *18* (Suppl 11), S257–S264.

Yurukoglu, Ali, Eli Liebman, and David B Ridley, "The role of government reimbursement in drug shortages," *American Economic Journal: Economic Policy*, 2017, *9* (2), 348–382.

Zarin, Deborah A, Tony Tse, Rebecca J Williams, and Thiyagu Rajakannan, "Update on trial registration 11 years after the ICMJE policy was established," *New England Journal of Medicine*, 2017, *376* (4), 383–391.

A Data Construction

A.1 Drug Approvals

Our primary source of data on drug approval and generic entry is the **Drugs@FDA** database, which is provided online and regularly updated by the U.S. Food and Drug Administration (FDA). With the exception of the following categories, this database contains most novel drug products:

- drugs not approved by the FDA
- drugs under review by the FDA for which action has not yet been taken
- over-the-counter (OTC) products approved for marketing through any process other than the submission of a New Drug Application (NDA) or a Biologic License Application (BLA)
- drugs sold outside of the U.S. that are not approved for marketing within the U.S.
- dietary supplements, which do not require FDA approval for domestic sale
- biological products (regulated by the Center for Biologics Evaluation and Research)
- animal drugs (regulated by the Center for Veterinary Medicine)

We use data from three Drugs@FDA files:

1. Application-level data: *application.txt*
2. Product-level data: *Product.txt*
3. Supplemental approval data: *RegActionDate.txt*

The Drugs@FDA data structure changed on 06 November 2016. After this date, many details in the Supplemental Approval data file were no longer available, and the file *RegActionDate.txt* was no longer published.[78] We use the latest available version of the database referred to as "Drugs@FDA 1.0," dated 25 October 2016, obtained via the Internet Archive's Wayback Machine.

A.1.1 Background on the FDA approval process and data files

With few exceptions, drugs that are prescribed in the United States must first be approved by the FDA. FDA approval indicates that, for its intended population, a drug's benefits outweigh both known and potential risks.[79] A brief primer on the vocabulary of drug applications and key features of the approval process provides helpful context for the data in this paper.

Application numbers: The FDA assigns six-digit application numbers to three key types of applications: new drug applications (NDAs); abbreviated new drug applications (ANDAs), and

[78] We have been unable to obtain analogous records from the FDA covering more recent years.

[79] See https://www.fda.gov/drugs/development-approval-process-drugs for details.

biologic license applications (BLAs). Each application may be differentiated by its: (i) unique active ingredients; (ii) dosage form; and (iii) strength.

New drug applications (NDAs): An NDA is submitted when there exists sufficient evidence on the safety and effectiveness of a drug to meet the FDA's requirements for new drug marketing. NDAs are typically submitted for all novel, brand-name drugs.

Abbreviated new drug applications (ANDAs): An ANDA is submitted for the approval of a generic drug that is the "therapeutic equivalent" of any previously approved drug. Drug products are considered to be therapeutic equivalents only if they are pharmaceutical equivalents and if they can be expected to have the same clinical effect and safety profile when administered to patients under the conditions specified in the labeling.

Biologic license applications (BLAs): A BLA is an application submitted for the approval of a biological product.

Application-level data The Application data file contains information on 20,560 drug approvals; 15,341 are NDAs, 5,082 are ANDAs, and 137 are BLAs.

Product-level data The Product data file contains information for each application that includes—among other variables—a record of drug ingredients that are associated with each application. Specifically, the file includes the following variables:

Active ingredients: An active ingredient in a drug is defined by the FDA as any component that provides pharmacological activity or another direct effect in the diagnosis, cure, mitigation, treatment, or prevention of disease. Several NDAs and ANDAs may have the same active ingredients.

Dosage form: A dosage form is the physical form in which a drug is both produced and dispensed. These include tables, capsules, and injectable forms.

Strength: The strength of a particular drug product captures how much of the active ingredient is present in each dosage form.

Supplemental approval data The RegActionDate data file contains records of all approvals associated with each NDA and ANDA. This file contains 128,906 approvals, which correspond to 17,299 unique application numbers. In addition to initial marketing approvals, this file contains records of "supplementary new drug applications" (sNDAs).

An sNDA is an application to the FDA requesting some change to the agency's approval for the drug. Changes may include new or modified indications, modifications to the labeling (i.e., the packaging affixed to a dispensed prescription), a new route of administration, or a new patient population for the drug. All changes of this kind must be approved by the FDA to ensure that the conditions that supported initial approval are still satisfied.

There are more than 50 categories of approval actions that can be granted using an sNDA. These approval types include standard approvals (which appear in our data as "N" type approvals), presubmissions ("M"), and manufacturing changes ("SCM"). For each approval, the FDA includes its action date which is the date at which the application was approved.

A.1.2 Identifying "innovator" drugs

We use data from drugs@FDA to define our base sample of new medicines. We call a drug an **innovator** if it meets the following criteria:[80]

1. It was approved under a New Drug Application (NDA) to the FDA.

2. It is the first approval of a unique combination of components

Determining whether a drug meets the first criterion is straightforward. Records in the Drugs@FDA database (applications file) are assigned one of three application types: "A" (abbreviated new drug application), "B" (biologics license application), or "N" (new drug application). We restrict consideration to applications of type "N."

Determining whether a drug meets the second criterion is more challenging—and foreshadows the medication name matching procedures discussed throughout this appendix. The core challenge is determining what constitutes a *unique component* in a drug. Many drugs are part of "substance sets" that include closely related products, which are pharmacologically and physiologically equivalent. Consider a drug molecule that contains a hydrogen atom. Replacing this hydrogen with "a metal or its equivalent" will form a salt, while replacing it with "an organic radical" will form an ester (Ward (2019), citing Amarin Pharm. Ireland Ltd. v. FDA, 106 F. Supp. 3d 196, 199 n.1 (D.D.C. 2015)). In many cases, these slight modifications have no clinically significant effects. However, these modified forms often appear as distinct active ingredients in the Drugs@FDA database. For example, *acetazolamide* and its salt form *acetazolamide sodium* both appear as active ingredients in our data. If we treated unique active ingredient strings as unique drug components for the purpose of constructing this sample, we would overcount the number of innovator products by incorrectly categorizing salts, esters, and other small chemical modifications to drugs as entirely new inventions.

Instead, we follow the FDA's own practice when evaluating the novelty of products and define innovator drugs in terms of unique *active moieties*: the part of the larger molecule that is responsible for the physiological or pharmacological action of the drug substance.[81] For the example above, *acetazolamide* is the active moiety in both *acetazolamide* and *acetazolamide sodium.* Thus, we construct our sample of innovator drugs by identifying the initial approvals of unique combinations of *active moieties* by the FDA.

[80] This is conceptually similar to, but not the same as, the FDA classification "new molecular entity" (NME). NMEs are defined as new drugs whose active ingredients are chemical substances that are marketed for the first time in the United States. Put differently, NMEs are drugs for which the FDA has not previously approved some component. NMEs are a strict subset of our sample definition. Our modification is to include first approvals of unique combinations of ingredients, even if every component has been previously approved by the FDA.

[81] Debates about whether the FDA should refer to "active ingredients" or "active moieties" when, for example, making determinations about various types of market exclusivity have been ongoing since the 1980s. See Ward (2019) for a more detailed discussion.

Active moiety crosswalk To the best of our knowledge, there is no standard crosswalk from active ingredient strings, as appear in Drugs@FDA, to active moieties.[82] We construct such a crosswalk, using the Drugs@FDA Product file as our input data.[83]

We begin by standardizing strings of active ingredients recorded in the database. We ensure text is lower case, components (for drugs that are combinations of multiple ingredients) are alphabetized and separated by semicolons, and extraneous punctuation is removed. Next, we separate each active ingredient string into its component parts. For example, *abacavir; lamivudine; zidovudine* is split into its three pieces. For each component in a drug, we manually review PubChem, chemistry textbooks, and online resources to identify the active moiety in the drug. We reconstruct strings for combination drugs, again ensuring that the component active moieties are alphabetized and formatted consistently.

Of the 25,976 unique active ingredient strings in our data, we determine that the active moiety and active ingredient are identical in 12,510 cases.

Defining our sample We drop 6 application numbers that are present in the Application file but do not have corresponding data in the Product file. All application numbers in the Product file have matching data in the Application file.

We begin by constructing a sample of initial drug approvals. We focus on records with an "N" (standard approval) doctype designation in the Supplement data file. There are 140 unique applications with multiple associated "N" approvals. For these 140 cases, we select the earliest approval recorded as the initial approval date for that application. We also drop 1,126 applications for which we never observe an "N" approval in our data. Thus, we construct a preliminary dataset that includes 16,171 initial approvals.

We merge these records onto the Application and Product files to yield a dataset with information on initial drug approval and product characteristics for 15,361 applications. Excluded from this sample are 836 applications without records in the Product and Application files, and 4,895 applications that have records in the Product and Application files but no corresponding approval histories in the Supplement file.

Finally, we impose the two restrictions described above to construct a sample of innovator drugs, which are unique at the active moiety level. We begin by dropping 122 drugs approved under

[82] The International Nonproprietary Names (INN) initiative maintained by the World Health Organization attempts to deal with exactly this problem. The description of the INN program reads: "When the INN Programme was initiated, it was decided that in such situations, in order to limit the number of published INNs, an INN should be selected for one member of such a group only. This approach, which concerns especially substance sets formed by salts or esters of the same active moiety was validated in the 20th report of the WHO Expert Committee on Nonproprietary Names for Pharmaceutical Substances (Technical Report Series No. 581)." See https://www.who.int/publications/i/item/inn-05-167-3 for details. Note that INNs may not always map onto drugs names as used in the United States. For example, "paracetamol" is the INN for "acetaminophen."

[83] In the interest of developing resources that can be re-used in research, we use the most current version of the Drugs@FDA database at the time this work was completed, dated 05 January 2023. Recall that this version of the Drugs@FDA database lacks information on product re-approvals central to our analysis. All other fields remain complete, however.

BLAs.[84] We drop 77 instances in which active moiety groups have no products that were approved under NDAs.[85] We drop 737 drugs based on their dates of approval: 250 instances in which drugs were initially approved before 1962, when the Kefauver-Harris Drug Amendment was signed into law; and 487 drugs approved between 1962 and 1985, when the regulatory structure introduced by the 1984 Hatch Waxman Act was implemented.[86]

We identify **1,102** unique combinations of active moieties that meet our criteria for "innovator" drugs. These drugs are our primary sample.

A.1.3 Linking "innovator" drugs to generic competitors

We designate any drug that is approved under an ANDA a "generic."[87] There are 8,322 application numbers in our dataset approved under ANDAs. To determine the first generic entrant for an innovator drug, we consider the set of all ANDAs with identical sets of active moieties and select the drug product with the earliest approval date. If there are multiple generic drugs within an active ingredient group being approved on the earliest date—which we observe for 186 unique products—we pick the drug with the lowest application number as our earliest generic entrant. We observe at least one generic entrant for 451 drugs in our sample—41%.

A.1.4 Counting approvals and re-approvals

We construct two measures of drug approvals and reapprovals. As noted above, a drug's initial approval is identified by searching for records with an "N" (standard approval) doctype, and we ensure that there is only one initial approval per drug. We construct a measure of "all approvals" associated with a drug, by identifying all approvals without an "N" doctype. Supplementary approvals associated with doctype "SE1" are approvals for new indications. We count SE1 approvals to generate a measure of approvals for new uses.

A.2 Scientific publications

We collect a measure of research investments from the National Library of Medicine's PubMed database. PubMed collects information on biomedical scientific publications, with curated links to

[84] Most BLAs are approved through the FDA's Center for Biologics Evaluation and Research (CBER). The sample of BLAs approved through CDER and indexed in Drugs@FDA is not representative of biologic drugs as a class.

[85] Examination of these cases indicates that these are instances where products were either (i) never approved as FDA brand-name drugs (e.g., application number 077394 was processed as an abbreviated new drug application in 2005, corresponding to "sodium bicarbonate" - baking soda); or (ii) correspond to innovator drugs that were approved before our dataset begins in 1962 (e.g. various applications related to penicillin, which was widely used beginning in the 1940s).

[86] The Kefauver-Harris Drug Amendment, often referred to as the "Drug Efficacy Amendment," introduced the requirement that drug manufacturers demonstrate the effectiveness and safety of drugs before approval. Manufacturers were also required, after this point, to disclose accurate information about side effects. Cheap generic drugs could no longer be re-branded as brand name (expensive) drugs and re-marketed as "breakthroughs."

[87] This designation is standard and consistent with FDA language. In FDA documents, ANDA and 'generic' are often used interchangeably. See https://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/default.htm for drugs@FDA descriptions of ANDAs for generic products

chemicals mentioned in the text of each paper. Chemical names are standardized to match terms in the "Medical Subject Headings" (MeSH) vocabulary, a controlled thesaurus maintained by the NLM. As of 2022, there were more than 30,000 entries in MeSH. Each MeSH term has an associated unique ID (UI). It is important to note that although each term corresponds to at most one UI, one UI can be associated with multiple synonymous terms. If chemical names are tagged with UIs, then, the numerical identifiers can allow us to identify and properly group synonyms.

For each active moiety in our innovator drug sample, we identified the corresponding MeSH term and UI. We used a crosswalk between MeSH terms and UIs drawn from the PubMed database.[88] When active moiety strings matched the text description of MeSH terms precisely, we assigned both the MeSH term and UI to the active moiety. When there was no precise match, we searched the web-based MeSH browser, to determine if there may be an alternative formatting, spelling, etc. that resulted in a match. In our sample of 1,102 innovator products, we successfully identified a MeSH UI for 810 (74%).

Comparatively few combination drugs have associated MeSH UIs. When MeSH terms exist for combination drugs, formatting is standardized (e.g. "hydrocortisone, urea drug combination" (MeSH UI: C010017)). However, of the 351 combination drugs in our sample of innovator products, only 78 have a MeSH UI.

For the set of non-combination drugs, we impute zero publications if there is no associated MeSH UI. For the sample of combination drugs, we proceed differently. We collect counts of all publications that reference every component of the combination drug (e.g., for "hydrocortisone, urea drug combination," we collect all publications that reference both "hydrocortisone" and "urea"). In our primary analyses, we use these counts as our measures of scientific publications.

We also collect flags from PubMed that indicate whether a publication received any funding from a public agency. We draw on information from the "grants" table of PubMed and designate a publication as have any public funding if there is any entry associated with one of three variables in PubMed: Grant ID, Acronym (Grant Agency), Grant Agency.

A.3 Clinical trials

A.3.1 Datasets

We collect information on clinical trials as one measure of drug development. To maximize coverage over our sample period, we draw records from four datasets. Table A1 summarizes the coverage of each dataset.

[88] Specifically, we used the "chemicals" table from the bulk version of PubMed. We constructed a list of all unique MeSH terms reported in the "substname" field and their associated UIs reported in the "substUI" field. As these fields are added by indexers at the National Library of Medicine (and not constructed by authors themselves), we did not verify the accuracy of these substname-substUI links. However, we did note that, as expected, each substname is associated with one substUI, as expected, and that the same substUI is associated with multiple substnames.

Table A1: Clinical Trials Datasets

dataset	version	file names	years of coverage	proprietary
NDA Pipeline	received in April 2007	nda_trials.dta	1982–2001	no
Pharmaprojects	current through May 2011	trend_accnum_year.dta	1995–2010	yes
Cortellis	current through 2016	trialid_singledata.dta trialid_drug_primary_intervention.dta	1995–2016	yes
ClinicalTrials.gov	downloaded from AACT; version dated 20220728	studies.txt; browse_interventions.txt; interventions.txt; interventions_other_names.txt	2010–2016	no

Notes:

NDA Pipeline

We collect records from the NDA Pipeline dataset, maintained by F-D-C Reports. The NDA Pipeline is a reference document covering drug developments during the previous calendar year. Each observation includes a text description of the stage of development for a given drug compound. We use data from 1982 to 2001.

Pharmaprojects

We collect additional information on drug development between 1995 and 2010 from Pharmaprojects, a proprietary dataset that records events associated with drugs at various stages of the development pipeline.[89] Pharmaprojects is compiled and sold by Citeline. Employees gather information from company websites, reports, and press releases. Every company in the database verifies information related to their listed drugs. Pharmaprojects reports trials that are newly initiated and ongoing in each annual report.

Citeline's documentation associated with the dataset notes: "There is continual two-way communication between Pharmaprojects staff and their contacts in the pharmaceutical and biotechnology industries; both to gather new data and importantly, to verify information obtained from other sources."

Cortellis

We also collect records on drug development between 1995 and 2015 from Cortellis, a proprietary dataset that, like Pharmaprojects, records details on drug development activity.[90]

[89] Our version of Pharmaprojects is current through May 2011.

[90] Details on the contents of Cortellis are available here: https://cortellislabs.com/api/clinical/. There are a small number of records for trials initiated before 1995, which appear to be incomplete (e.g., there is one trial in the Cortellis data that is active in the year 1981). We do not include these records in our sample.

ClinicalTrials.gov

We supplement these records with data drawn from ClinicalTrials.gov, the largest repository of clinical trial records in the world, for clinical trials initiated between 2010 and 2016. Since 2007, sponsors of trials regulated by the U.S. Food and Drug Administration are required, by law, to register clinical trials with ClinicalTrials.gov. Since 2005, registration in a global registry has been required as a condition of publication in any journal associated with the International Committee of Medical Journal Editors.

A.3.2 Data construction

NDA Pipeline

NDA Pipeline records include the following information: document publication year, firm name, trade name of drug, generic name of drug, abbreviation of drug name, description of study, status, and miscellaneous notes. We identify developments about clinical trials by searching for keywords within the "status" variable. Specifically, we keep records where the status variable includes either "phase" (as in, "Phase I trial") or "clinical." We drop records where the description includes either "preclinical" or "pre-clinical."

As these records are intended to provide annual updates on drug development pipelines, clinical trial records are best interpreted as reports of trials that were active in the year prior to publication. That is, they reflect firm decisions about project *continuation*, rather than project *initiation.* Since these documents were published one year after the relevant clinical developments, we subtract one from the year of publication to generate a variable capturing the year in which a trial was conducted.

We link records to our baseline sample of FDA-approved active moieties using the generic name of the drug. We standardize the formatting of the string (remove punctuation, alphabetize components, remove extraneous characters, etc.) and attempt two merges. First, we attempt to merge the generic name of the drug in this dataset to the active ingredients associated with our baseline sample. For observations that do not successfully merge, we next attempt to link generic names to the active moieties associated with our baseline sample. We impute annual trial counts of zero for all active moieties approved during this time period.

Our final dataset includes records associated with 543 FDA-approved active moieties (or combinations) tested between 1981 and 2000. The most studied drugs in our sample are paclitaxel and thalidomide.[91]

[91] Paclitaxel is a type of chemotherapy that can treat various forms of cancer, originally approved in 1992. See https://www.nytimes.com/2005/01/08/business/new-form-of-breast-cancer-drug-approved.html. Thalidomide was approved by the FDA in 1998 to treat leprosy and has since been reapproved to treat various forms of cancer. Given its (notorious) history of causing birth defects, its distribution is tightly controlled. See https://www.nytimes.com/1998/07/17/us/thalidomide-approved-to-treat-leprosy-with-other-uses-seen.html.

Pharmaprojects

We use data from Pharmaprojects on the status of drug development projects. We collect information on clinical trial developments by, similarly, searching within a variable about project status ("originator status"). As with NDA Pipeline data, we keep records where the status variable includes either "phase" (as in, "Phase I trial") or "clinical." We drop records where the description includes either "preclinical" or "pre-clinical." Pharmaprojects records provide annual updates on drug development projects and, thus, also capture project *continuation*, rather than project *initiation*.

Our linking procedure with Pharmaprojects data is identical to the linking project for NDA Pipeline. We link records to our baseline sample of FDA-approved active moieties using the generic name of the drug. We standardize the formatting of the string (remove punctuation, alphabetize components, remove extraneous characters, etc.) and attempt two merges. First, we attempt to merge the generic name of the drug in this dataset to the active ingredients associated with our baseline sample. For observations that do not successfully merge, we next attempt to link generic names to the active moieties associated with our baseline sample. We impute annual trial counts of zero for all active moieties approved during this time period.

Our final dataset includes records associated with 613 FDA-approved active moieties (or combinations) tested between 1995 and 2011. The most studied drug in our sample is paclitaxel.

Cortellis

We also collect data from Cortellis that—like the Pharmaprojects data—collects data on drug development typically sold for the purposes of "competitive intelligence." Cortellis includes one record per clinical trial and includes variables that capture start year and end year. We transform these data into records that indicate whether a given trial on a drug was active in each year.

Our linking procedure with Cortellis data is identical to that used for NDA Pipeline and Pharmaprojects. We, again, impute annual trial counts of zero for all active moieties approved during this time period.

Our final dataset includes records associated with 990 FDA-approved active moieties (or combinations) tested between 1995 and 2016. The most studied drug in our sample is docetaxel.

ClinicalTrials.gov

We use data from ClinicalTrials.gov to supplement records from NDA Pipeline, Pharmaprojects, and Cortellis. Although ClinicalTrials.gov was made available to the public in 2000, we collect records beginning in 2010, by which time trial registration had become more common.[92] We collect

[92] In 2005, the International Committee of Medical Journal Editors (ICMJE) began requiring trial registration as a condition of publication. In 2007 the requirements for submission to ClinicalTrials.gov were expanded after Congress passed the Food and Drug Administration Amendments Act of 2007 (FDAAA) (PDF). Section 801 of FDAAA (FDAAA 801) required more types of trials to be registered and additional trial registration information. For more details, see https://clinicaltrials.gov/ct2/about-site/history.

records of all trials, regardless of trial phase, that were active between 2010 and 2016. ClinicalTrials.gov allows us to construct two measures of trial activity: trial initiation decisions (using trial start years) and trial continuation decisions (using trial active years). To facilitate integration with NDA Pipeline and Pharmaprojects data, our main analyses use trial continuation decisions measured using trial active years. There are a number of concerns about the completeness and accuracy of ClinicalTrials.gov records (see, for example, Zarin et al. (2017); Tse et al. (2018)). Given these limitations, we report analyses that consider the likelihood of any clinical trial associated with a given drug in each year, rather than considering the total number of trials.

As ClinicalTrials.gov is administered by the National Library of Medicine, we follow the same procedure to link records used for PubMed data. Trial sponsors report the interventions (therapies, behavioral treatments, drugs, etc.) associated with each study using standardized MeSH terms, in the file *browse_interventions*. We link MeSH terms associated with clinical trials to the MeSH terms associated with each drug in our baseline sample. We impute annual trial counts of zero for all active moieties approved during this time period with no matching trial records.

Our final dataset includes records associated with 1,013 FDA-approved active moieties (or combinations) tested between 2010 and 2016. The most studied drug in our sample is docetaxel.

A.3.3 Clinical trials sample

We construct a sample of clinical trials by appending records for NDA Pipeline, Pharmaprojects, and ClinicalTrials.gov. Our primary measure of trial activity is an indicator variable that takes a value of one if there was any active clinical trial associated with an innovator drug in a given year. Although our data do enable us to consider the total volume of clinical trials, we prefer to focus on a binary measure to avoid introducing concerns about differences in measurement and coverage across the four datasets.

A.4 Drug-disease crosswalk

We construct data on the specific therapeutic uses (indications) associated with each drug approval and re-approval in our sample. For each approval/re-approval event, we collect text descriptions of the associated indications from a variety of sources. Where possible, we use FDA application number and date of (re-)approval to retrieve the relevant drug label from the FDA's full-text drug product labeling service, FDALabel.[93] We supplement data on labels—which are primarily available from 2005 forward—with data from drug approval packets[94], dated company press releases, scientific articles, and other internet searches.

For each text description, we identify a set of associated alphanumeric codes from the International Classification of Diseases, Tenth Revision (ICD-10), a classification system of diagnosis codes representing conditions and diseases, related health problems, abnormal findings, signs and

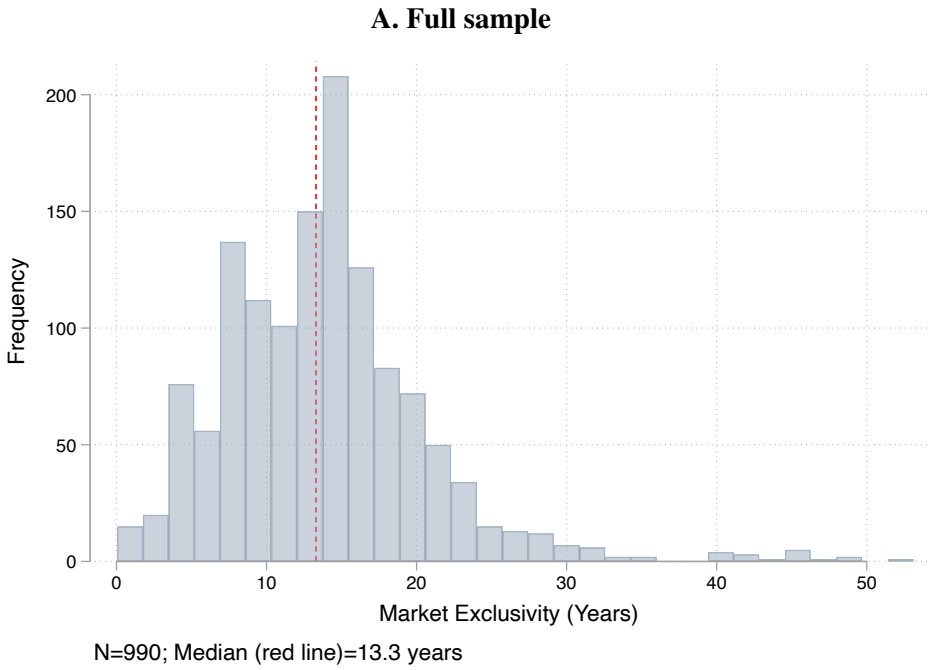
[93] https://www.fda.gov/science-research/bioinformatics-tools/fdalabel-full-text-search-drug-product-labeling

[94] Drug approval records are available for many drugs beginning in 1995 through the Drugs@FDA interactive search tool: https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm

symptoms, injuries, and external causes of injuries and diseases. We use online searches to identify an initial set of "likely" ICD-10 codes. We verify the accuracy of these codes using online code look-up tools, intended for hospital staff who are completing billing tasks.[95] A second researcher verified the accuracy of a random 20 percent sample of codes. We report the most granular ICD-10 code available, but though it is common for analyses using ICD-10 codes to use truncated versions. Concretely, instead of using **G43.1** (Migraine with Aura), we use **G** (Diseases of the Nervous System) and **G43** (Migraine).

A.5 Distribution of Market Exclusivity

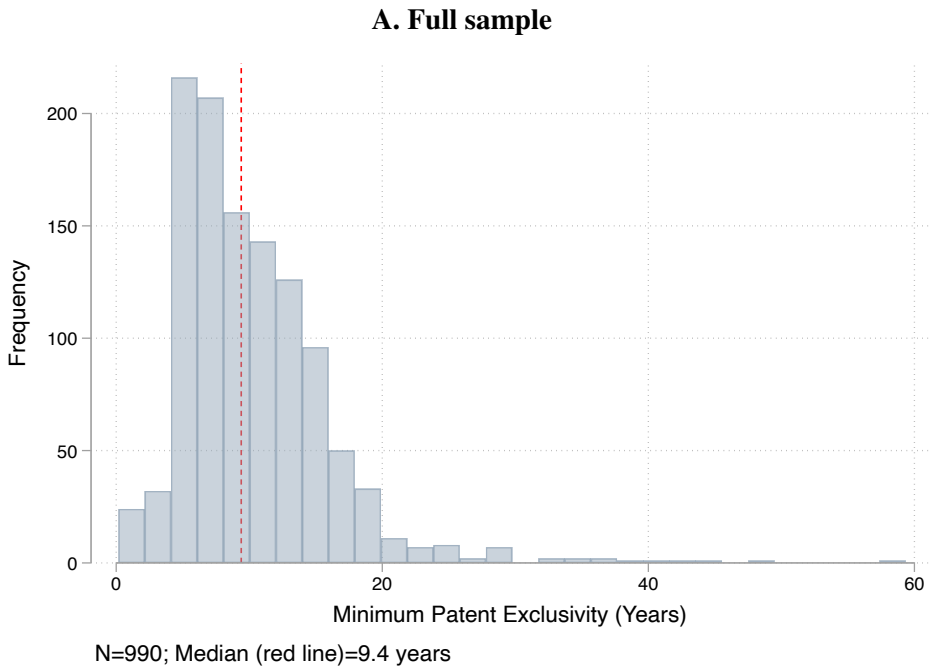
Figure A1: Market Exclusivity



Notes: Figure is a histogram that plots the distribution of the market exclusivity measure defined in the text as the number of years between an innovator drug's initial approval and the earlier of the expiration of all patents and regulatory exclusivities and generic entry. The distribution of this measure is plotted for the sample of innovator drugs defined in Section 4.5.

[95] For example, https://icdcodelookup.com/icd-10/codes

Figure A2: Minimum Patent Exclusivity

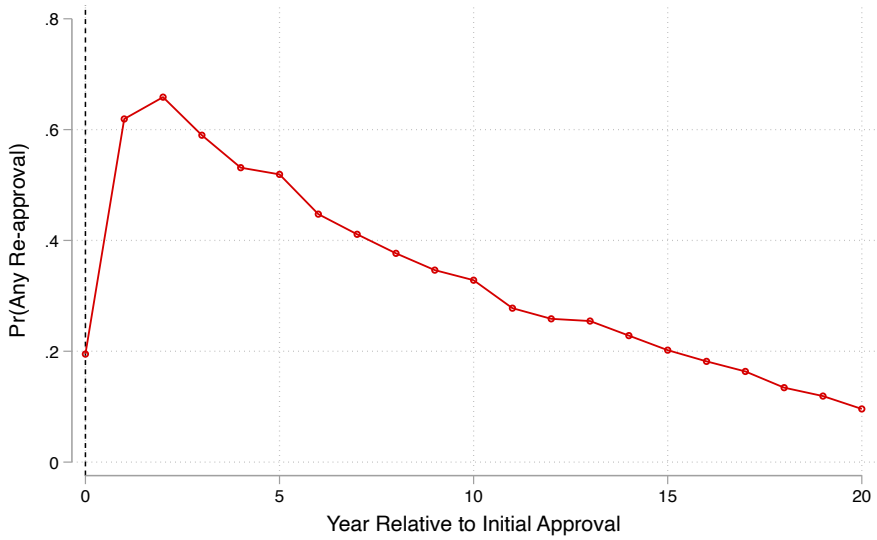


Notes: Figure is a histogram that plots the distribution of the minimum patent exclusivity measure defined in the text as the number of years between an innovator drug's initial approval and the expiration of all regulatory exclusivities and the earliest expiring patent. The distribution of this measure is plotted for the sample of innovator drugs defined in Section 4.5.

B Additional Results

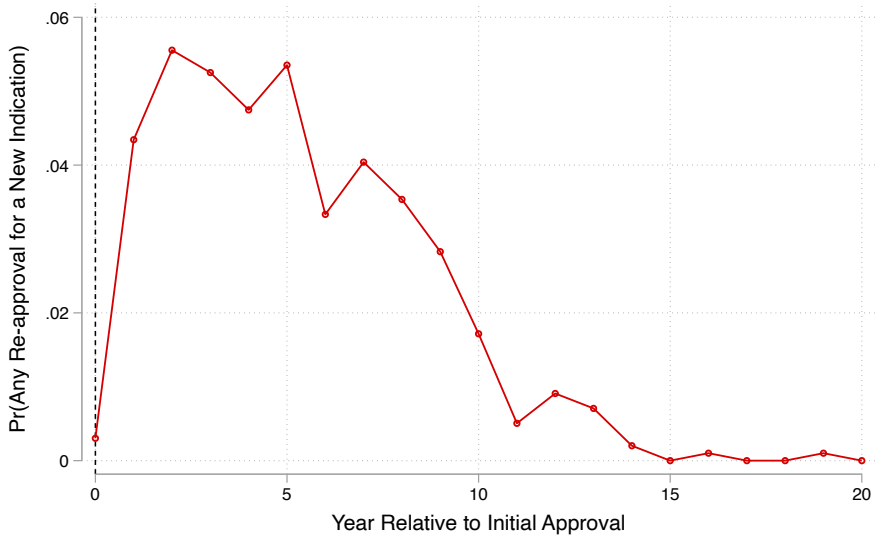
Figure A3: FDA Approvals Relative to Initial Approval

A. All Re-approvals



N= 990 unique approvals

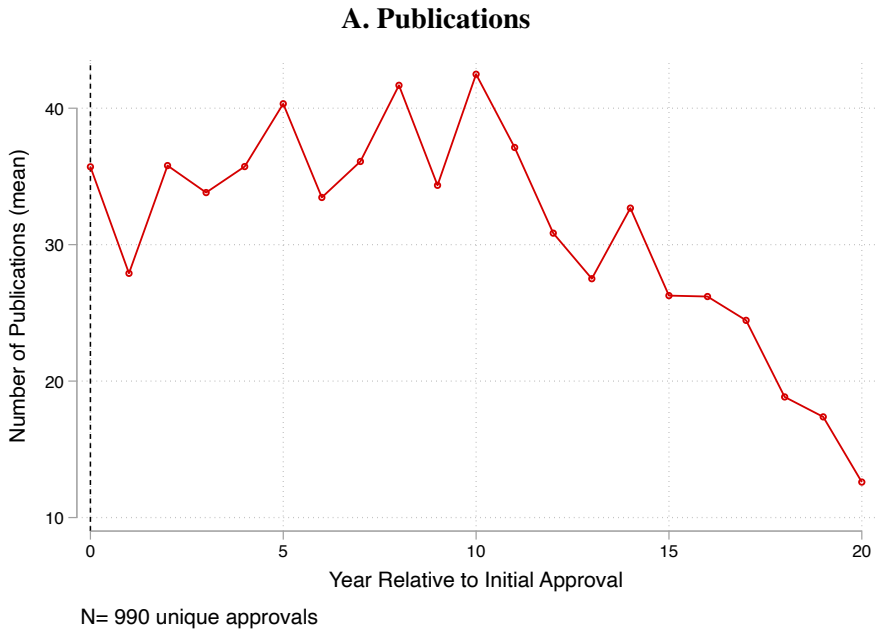
B. New Use Approvals

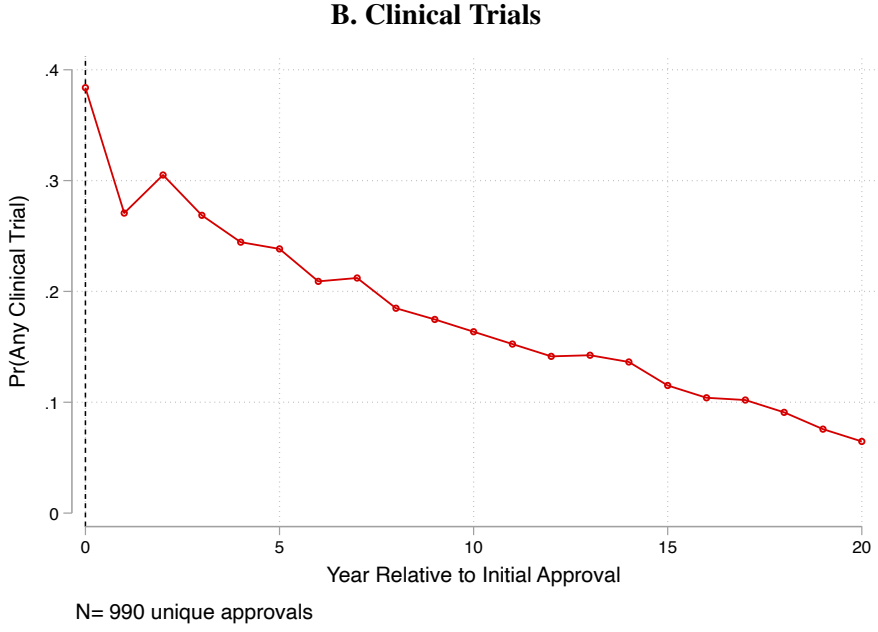


N= 990 unique approvals

Notes: Panel A plots the probability of any re-approval (e.g., new use, new dosage form, etc.) for the sample of innovator drugs defined in Section 4.5 in calendar time. Panel B plots the probability of re-approval for a new use for the sample of innovator drugs defined in Section 4.5 ($N = 990$).

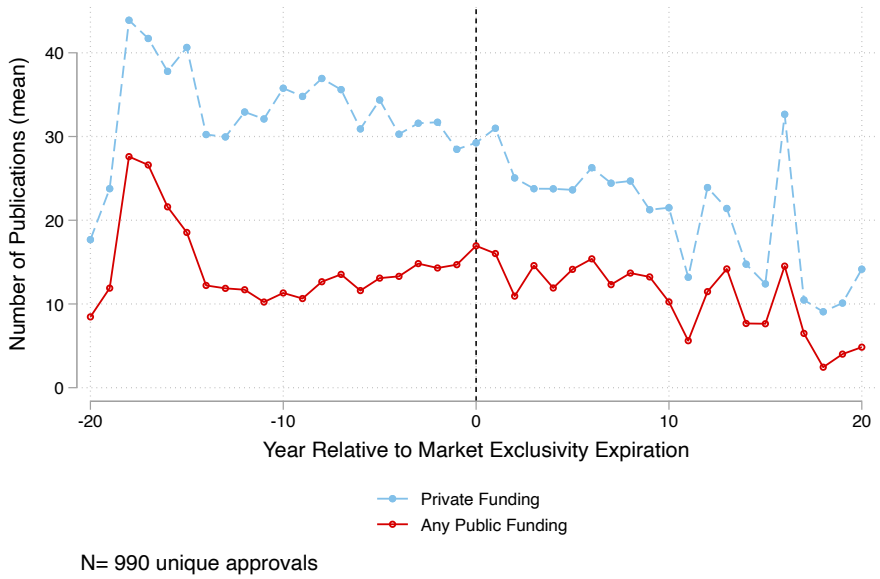
Figure A4: Research Investments Relative to Initial Approval





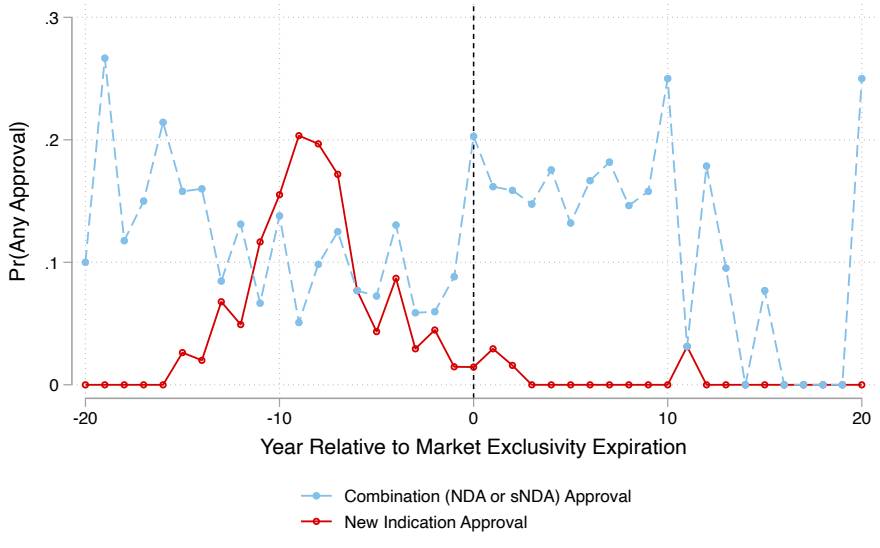
Notes: Panel A plots the average number of scientific publications published in each year for the sample of innovator drugs defined in Section 4.5 relative to initial approval. Panel B plots the probability of any active clinical trial in each year for the same sample of innovator drugs, also in each year relative to initial approval.

Figure A5: Scientific Publications Relative to Exclusivity Expiration



Notes: Figure plots the average number of scientific publications published in each relative year for the sample of innovator drugs defined in Section 4.5 ($N = 990$). Publication counts are plotted separately for publications that acknowledge receiving any public research support (red solid line) and publications that do not (blue dashed line). Market exclusivity expiration dates are defined as the earlier of two events: expiration of all patents and regulatory exclusivities linked to the innovator drug in the FDA's Orange Book, or FDA approval of a generic drug with an identical set of active moieties.

Figure A6: Combination Approvals



Notes: Figure plots two measures of FDA approvals for a subset of innovator drugs with a single active moiety. We restrict consideration to the set of single-moiety innovator drugs that ever appear in a combination drug ($N = 140$). Market exclusivity expiration dates are defined as the earlier of two events: expiration of all patents and regulatory exclusivities linked to the innovator drug in the FDA's Orange Book, or FDA approval of a generic drug with an identical set of active moieties. Panel A shows the probability of re-approval for a new indication for the twenty years preceding and following exclusivity expiration for all drugs in our sample. Panel B takes a balanced panel approach, including only those drugs that can be observed for a full five years before and after market exclusivity expiration.