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

Eric Budish Maya Durvasula Benjamin Roin Heidi Williams

February 2026

Abstract

We investigate missing markets for innovation: contexts where intellectual property rights necessary to incentivize research investment are either unavailable or unenforceable. A simple theoretical model formalizes the conditions that yield missing investment. We identify an empirical setting—research into new uses for existing drugs—with sharp variation in the enforceability of intellectual property rights on comparable inventions. We show that when intellectual property rights become unenforceable, research investment and commercialization nearly cease. Estimates of the value of this missing investment are in the trillions of dollars. We discuss the implications of these findings for the design of innovation policy.

*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, Oren Bar-Gill, Daniel Fetter, Amy Finkelstein, Peter Ganong, Matthew Gentzkow, Daniel Hemel, Scott Hemphill, David Kamin, Louis Kaplow, Mark Lemley, Paul Milgrom, Abhishek Nagaraj, Matt Notowidigdo, Lisa Larrimore Ouellette, Al Roth, Elizabeth Sepper, Carolyn Stein, Chad Syverson, Melissa Wasserman, Brian Wright, and seminar participants at AALS (Law & Economics), Chicago Booth, Harvard Law & Economics, MIT Sloan, NBER Market Design, NBER Productivity Lunch, NBER Summer Institute (Law & Economics), Northwestern Kellogg, NYU Law, NYU Law & Economics, Penn Carey Law, Stanford, Texas Law, the Radcliffe Institute, TNIT, UC-Berkeley Haas, and UCLA, 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, the B.F. Haley and E.S. Shaw Fellowship for Economics via the Stanford Institute for Economic Policy Research, the Sloan Foundation Graduate Fellowship on Innovation and Productivity via the NBER, and the National Science Foundation through grant DGE-1656518 (Durvasula) to Stanford University. While this paper was written, Williams was a visiting scholar and is currently working on contract as a research advisor to 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 incentivizes firms to engage in these processes of research and development, competitive markets will under-provide socially valuable innovation (Nelson, 1959; Arrow, 1962; Nordhaus, 1969). Policies based on this idea go back at least half a millennium to 15th century Italy and are enshrined in the U.S. Constitution.[1]

In this paper, we investigate the idea that, for large classes of potentially valuable inventions, intellectual property rights that exist "on paper" are impossible to enforce in practice. When this is the case, the key mechanism used to induce private investment in research and development provides no meaningful incentive for innovation. We show that there are many scientifically feasible, socially valuable inventions that are missing as a result.

Consider three hypothetical inventions: a rigorously tested regimen of nutritional supplements that reduces the risk of developing Alzheimer's disease for adults or that improves cognitive development for babies; a manufacturing process that would substantially 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.[2] 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.[3] Yet any resulting grants of intellectual property would offer little protection from imitation: in each case, inventors would find it nearly impossible to enforce those rights in practice.

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 widely-available supplements in a way that infringes on a patent or that a bottle of pills was manufactured using a protected

[1] Article I, Section 8, Clause 8 ("[t]o promote the Progress of Science and useful Arts, by securing for a limited Time to Authors and Inventors the exclusive Right to their respective Writings and Discoveries").

[2] We choose, for the sake of illustration, three types of socially valuable inventions for which the enforceability problem of interest to this paper can arise: medical inventions, business process inventions, and product inventions. For a discussion of the value of rigorous evaluation of health interventions, 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); for an analysis of the 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.

Note that 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).

[3] It may surprise some readers that all three kinds of inventions described in the text are eligible for patent protection. This reflects, in part, the distinction this paper draws between the availability of patent protection in principle and the enforceability of patent protection in practice. Section 2 provides a primer on the standards for patentability in the United States.

method, inventors will be unable to enforce their patents simply because they cannot observe infringement. And even when an 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 there is no viable party against whom she can 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 earn an economic return on their investments in research and development without formal grants of intellectual property—so long as they can maintain 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.

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 completely 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.[4] Generic versions of the drug entered the market in 2002. In the years that followed, oncologists began to collect tentative evidence that metformin may also be effective at treating or preventing cancer. Though these researchers proceeded with various "preclinical" 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.[5]

To be clear, the USPTO would likely have granted a patent on the use of metformin to treat cancer, and the FDA would then have barred other firms from marketing generics for this use.[6] But once the information about metformin's value as a treatment for cancer became publicly known, any patient or their physician could make use of the information freely—by using or prescribing the generic version—without compensating the firm that undertook the expensive clinical trial. And even if that firm could detect unauthorized use, enforcement would be practically impossible: the only parties on whom liability would definitively fall would be the patients using the drug to treat their own cancer.[7] Decades after preliminary evidence first emerged, whether metformin has potential as a treatment or even potentially as a preventative for cancer remains an open question (O'Connor et al., 2024).[8]

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 consequence 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

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

[5] 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*" (emphasis added).

[6] 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.

[7] 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 liability, though suits against them for patent infringement are rare for the simple reason that it is rarely advantageous for firms to sue their own (repeat) customers. A generic manufacturer that encourages infringement of a patent may also be liable, though there are several statutory protections that limit the viability of many such cases (see Section 3.2).

[8] Since at least Gale and Gerlin (2012) reported on the difficulties of garnering commercial support for clinical trials, an academic literature has explored whether metformin may have anti-cancer effects. Much of this work consists of observational studies, secondary analyses, and relatively small, exploratory randomized trials—often conducted in diabetic populations (where metformin use is correlated with disease severity, comorbidities, and baseline cancer risk) or using surrogate endpoints. Though well suited to generating hypotheses and characterizing potential biological mechanisms, these studies have not been designed to provide definitive causal evidence on metformin's efficacy for a specific cancer indication. Accordingly, meta-analyses have catalogued problems with scale, design, and publication bias (Lord and Harris, 2023); a recent, comprehensive meta-analysis in the *Journal of the National Cancer Institute* concludes as follows: "Metformin may be associated with a decreased risk of many cancer types, but high heterogeneity and risk of publication bias limit confidence in these results. Additional studies in populations without diabetes are needed to better understand the utility of metformin in cancer prevention" (O'Connor et al., 2024). Metformin is far from the only widely-used drug for which there is suggestive evidence of high-value new uses: Roin (2014, p. 43) collects examples of "possible treatments for cancer (including cancer prevention), Alzheimer's disease, depression, diabetes, stroke, tuberculosis, malaria, multi-drug resistant bacteria," among others.

set a price for its drug that reflects all commercialized uses. So long as the firm retains market exclusivity, it is compensated for both the original use and any 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 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 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.[9]

Using these data, we show that investments in new uses for existing drugs closely track our predictions. At peak, ten years prior to market exclusivity expiration, seven percent of drugs in our sample are re-approved for at least one new use. In each subsequent year the likelihood of re-approval declines monotonically;[10] this likelihood falls to roughly zero following market exclusivity expiration. The vast majority of new uses ever approved by the FDA for our sample of drugs were approved prior to exclusivity expiration.

There are two potential alternative explanations for why investment in new uses drops off at the time of exclusivity expiration. First, the set of "scientific opportunities" for new uses may simply be exhausted by then. Alternatively, an innovator firm that knows all potentially viable new uses may strategically sequence research and development to ensure that all new uses come to market during their period of enforceable market exclusivity. 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

[9] A separate paper, Durvasula et al. (2023), serves as a "user's guide" to the U.S. Food and Drug Administration's "Orange Book" dataset, which offers a unique linkage between products (drugs) and patents, which is key in enabling analyses of the kind discussed in this paper.

[10] We document similar trends in clinical trial activity and privately-funded scientific publications: both measures are high early in innovator drugs' lifecycles and monotonically decreasing over time, consistent with declining private sector investment at each stage of the commercialization process.

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 generated by clinical trial requirements 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., 2015a). We use this 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 document evidence consistent with the idea that the enforceability of intellectual property rights is the mechanism driving our results. At the time of market exclusivity expiration, intellectual property rights on new uses become unenforceable because generic drugs are perfectly substitutable for brand-name drugs. But not all re-approvals of existing drugs face the same enforceability problem when market exclusivity on the brand-name drug itself is lost. Fixed-dose combinations, for example, bundle multiple active ingredients into a single dosage form and are regulated as distinct new drugs, each receiving its own enforceable protections that block generic entry. Because generic versions of the individual components are not perfect substitutes for the bundled product, intellectual property rights on these combinations remain enforceable after exclusivity on the underlying drugs expires. Consistent with the idea that enforceability, rather than scientific opportunity, drives the patterns we document, we find that firms substantially increase their investment in fixed-dose combinations at the point of exclusivity expiration.[11]

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.

These empirical exercises allow us not only to establish the existence of a missing market but also to quantify the social value of missing research investments. Our estimated elasticity implies that if intellectual property rights on new uses were enforceable, there would be between 0.1 and 0.4 additional new uses per drug. 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. To offer a sense of magnitudes, we draw on existing estimates of the social value of new drugs and adopt an intentionally conservative figure, valuing each missing use at $0.5 billion per year. The value of missing research and development in this market, then, is on the order of $100–400 billion per year. At a conservative social discount rate, the present value of this missing

[11] 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 sometimes 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 anticompetitive conduct and other solutions that legal scholarship often characterizes as "self help" (Lichtman, 2005).

innovation is on the order of $2.5 to $10 trillion.[12]

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 have ever been 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.

The idea that there are insufficient private incentives 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; Sachs et al., 2017; Conti et al., 2020; EveryCure, 2025). 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, 2011a,b, 2012; Collins and Varmus, 2015).

Evidence of a quantitatively significant missing market for new uses provides, as well, clear empirical support for the paper's simple model of intellectual property and innovation incentives. More specifically, 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., 2015a).[13] 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., 2015a).

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 greater investment—there is remarkably little direct empirical support, owing to the difficulty

[12] Prior work has estimated comparably large magnitudes for the value of missing medical innovation. See, for example, Murphy and Topel (2006), Hall and Jones (2007), and Budish et al. (2015a).

[13] In particular, we formalize and test an idea conjectured in Kapczynski and Syed (2012)—that there are "domain[s] of innovation that patents, whatever their scope, cannot adequately address."

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. (2015a) document evidence consistent with the idea that longer patent terms induce additional research investments, but are unable to rule out corporate short-termism as an alternative explanation. 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. 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. Thus, we contribute more broadly to the empirical literature on the determinants of innovation (for surveys, see Williams, 2017 and Bryan and Williams (2021)) and, more specifically, to the strands of this literature that study how property rights affect incentives for investment (e.g., Moser, 2005; Budish et al., 2015a; Hodgson, 2024; Dix and Lensman, 2024) and how under-provision of incentives distorts research and development (e.g., Budish et al., 2015a; Krieger et al., 2022; Michelman and Msall, 2024; Dix and Lensman, 2024).

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 sets of evidence 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. Section 7 uses our framework to summarize opportunities to "fix" the problem of missing markets when standard intellectual property levers are unavailable, and Section 8 concludes.

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 practice.[14] 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

[14] 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 primary goal the provision of incentives for innovation.

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

Our starting point is a Nordhaus (1969)–style framework, as adapted by Budish et al. (2015a). A representative firm must decide whether to attempt to commercialize a potential invention, the characteristics of which it regards as exogenous.[15] Commercialization refers to the costly and risky process of bringing a product or idea to market—running clinical trials to assess efficacy and satisfy regulatory requirements, running focus groups to iterate on a product's design and features, building manufacturing capacity to allow for production at scale, etc.

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 in year t_{invent} whether or not to attempt to commercialize the potential invention. Commercialization involves a deterministic delay t_{comm}. We normalize $t_{\mathsf{invent}} = 0$.
- **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.[16] 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}.
- **Discount factor**: The firm discounts future profits by a factor δ.
- **Monopoly profits and social value**: A successfully 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.

[15] 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).

[16] In our empirical context, for example, clinical trials may indicate that a drug is ineffective as a treatment for a particular disease.

- **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 $p_{\text{secret}} \in [0,1]$ captures the likelihood that the invention remains secret—and, thus, profitable—upon commercialization.[17] 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 successful 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 $p_{\text{eligible}} \in [0,1]$. The associated protection lasts for t_{patent} 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 $p_{\text{enforce}} \in [0,1]$.

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. The firm's expected time under monopoly 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

$$T_{\text{secret}} = p_{\text{comm}} \cdot p_{\text{secret}} \cdot \sum_{t_{\text{comm}}}^{\infty} (\delta\gamma\eta)^t. \tag{1}$$

In words, time under secrecy T_{secret} 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).

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

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

In words, time under intellectual property protection T_{patent} is determined by (i) the likelihood of successful commercialization p_{comm}, (ii) the likelihood that the invention is eligible for intellectual property protection p_{eligible}, times the likelihood that the intellectual property rights can be enforced

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

p_{enforce}, and (iii) the expected duration of effective intellectual property protection $t_{\text{patent}} - t_{\text{comm}}$, adjusting for discounting and obsolescence risk.

2.1.3 Incentives to Invest

A profit-maximizing firm will attempt to commercialize a potential invention if and only if the expected profits exceed the costs of commercialization. Define T^* as the firm's expected time under monopoly under whichever of trade secrecy and patent protection yields more time in expectation:

$$T^* = \max(T_{\text{secret}}, T_{\text{patent}}). \tag{3}$$

The firm will pursue the invention if and only if

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

If $T_{\text{secret}} > T_{\text{patent}}$ then the firm chooses trade secrecy and T^* is given by (1). Otherwise, the firm chooses intellectual property protection and T^* is given by (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. That is, both $T_{\text{secret}} = 0$ and $T_{\text{patent}} = 0$.

With (1)–(3), we can write the expected time under monopoly T^* for all potential inventions as:

$$T^* = \max\left(p_{\text{comm}} \cdot p_{\text{secret}} \cdot \sum_{t_{\text{comm}}}^{\infty} (\delta\gamma\eta)^t, p_{\text{comm}} \cdot p_{\text{eligible}} \cdot p_{\text{enforce}} \cdot \sum_{t_{\text{comm}}}^{t_{\text{patent}}} (\delta\gamma)^t\right). \tag{5}$$

Assume that the invention is plausibly successful $p_{\text{comm}} > 0$, not immediately obsolete $\gamma > 0$, and that the firm is not completely impatient $\delta > 0$. Under what more substantive conditions is $T^* = 0$? There are two cases.

The first case arises if trade secrecy is ineffective and the invention's patent term is too short relative to the time it takes to bring the invention to market. Formally, $p_{\text{secret}} = 0$ and $t_{\text{patent}} < t_{\text{comm}}$. This case is discussed in Budish et al. (2015a), who provide evidence that firms underinvest in projects that are completed over long time horizons.

A second case, our focus here, arises if trade secrecy is ineffective and the invention is either not eligible for intellectual property protection or the available intellectual property rights are unenforceable. Formally, $p_{\text{secret}} = 0$ and either $p_{\text{eligible}} = 0$ or $p_{\text{enforce}} = 0$.

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_{patent} (e.g., by extending patent terms), or the relationship between the invention's timing

parameters t_{invent}, t_{comm}, and t_{patent} (e.g., by starting patent terms at commercialization instead of invention). These shifts, discussed in Budish et al. (2015a), 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 (No Nordhaus Parameters)**.** *If* $p_{\text{secret}} = 0$ *and either* $p_{\text{eligible}} = 0$ *or* $p_{\text{enforce}} = 0$*, then neither longer* $(t'_{\text{patent}} > t_{\text{patent}})$ *nor broader* $(\pi' > \pi)$ *patent terms make commercialization profitable.*

Theoretical studies of innovation and intellectual property rights often characterize optimal policy in terms of two "Nordhaus parameters"—t_{patent} (capturing patent duration) and π (capturing patent breadth).[18] 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_{patent} and π have no impact on the firm's investment decision in (4).

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 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 invention 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.[19]

[18] See Gilbert and Shapiro (1990) and Klemperer (1990) for models of patent breadth. Gilbert and Shapiro (1990) define patent breadth as the profit parameter π—that is, as the flow rate of profit available to the firm so long as the patent is in force. Klemperer (1990) defines patent breadth in a spatial horizontal differentiation model, where greater breadth gives the firm greater market power and hence greater profits in equilibrium.

[19] When secrecy is a commercially viable option, state and federal laws in the United States provide formal trade secret protection. For example, Uniform Trade Secrets Act (with 1985 Amendments), 14 U.L.A. 437 (2005); Cal. Civ. Code

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).[20] 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., Mansfield, 1986; Levin et al., 1987; Cohen et al., 2000; Moser, 2005, 2007).[21] Manufacturing processes used entirely within a firm—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 disclosure.[22] 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).[23] Outside markets for health technologies, engineered secrecy is quite common. Every aspect of the "fuzzy logic" used to modulate cook 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 disassembly will destroy the relevant component (Santo, 2023).

§§ 3426–3426.11; Defend Trade Secrets Act of 2016, 18 U.S.C. §§ 1836–1839. 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.

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

[21] 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.

[22] 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 inspect only from the outside: https://www.worldofcoca-cola.com/explore-inside/explore-vault-secret-formula. A very small number of employees have access to the formula itself, 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.

[23] 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. Sequence data alone was insufficient to enable imitation or reproduction without additional reverse-engineering of the methods used to produce the commercially-available version of the product. For a longer discussion, see Gault (2021).

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.[24] 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,[25] methods for processing loan information through a clearinghouse,[26] and computer programs that automatically categorize digital images[27]—are ineligible for patent protection.

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 on inventions that are novel, non-obvious, and useful (35 U.S.C. §§ 101, 102, 103). Roin (2009) observes that statutory requirements in the patent context 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, in particular, 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).

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

[24] 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).

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

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

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

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 protection, not all intellectual property rights are enforceable. Though many factors affect this enforceability, we highlight two: whether the firm can observe instances in which its rights are infringed and whether enforcement is practical.

If a firm cannot observe instances in which its intellectual property rights are infringed, such rights are unenforceable, and p_{enforce} takes a value of zero. Of course, for many inventions, unauthorized uses—like "knockoffs" of luxury handbags (see Hemphill and Suk, 2008) and counterfeit consumer electronics (see 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).

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 that license methods of petroleum refining to other parties within the 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.[28] To understand what this means in practice, consider the case of a firm that holds a patent on a method for using orthodontic retainers to straighten teeth. A competitor that copies the product and sells its own retainers to be used in the same way may not, technically, infringe the patent. Instead, infringement will occur each time a patient uses the copycat retainer to straighten her teeth—thus *performing* the method. To enforce its patents, the innovator would have to file suit against each retainer-wearing defendant. In general, consumers are rarely attractive defendants in

[28] Direct infringement is governed by 35 U.S.C. § 271(a), which states: "[W]hoever without authority makes, uses, offers to sell, or sells any patented invention, within the United States or imports into the United States any patented invention during the term of the patent therefor, infringes the patent." Patent holders can also sue parties that have facilitated infringement, even if they themselves have not directly infringed. Federal law recognizes two forms of indirect infringement: inducement (35 U.S.C. § 271(b)) and contributory infringement (35 U.S.C. § 271(c)). For a primer, see Part III of Masur and Ouellette (2025).

such cases, both because these suits are likely to engender backlash and because these defendants have comparatively shallow pockets that limit the extent of any recovery.

2.2.4 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.

3 The New Uses Problem

3.1 Background

Modern medicine relies on a relatively small set of chemical compounds—roughly 2,000 in total—that have been shown to be safe for use in humans and effective for at least one therapeutic indication. Scientists often describe compounds that meet both of these demanding criteria as "promiscuous," by which they mean that if a drug can successfully treat one disease without causing serious harm, it is likely to be useful in treating others as well (Wermuth, 2006).

In some cases, multiple therapeutic uses are anticipated at the outset of drug development—for example, when related conditions (e.g., depression and anxiety disorders) are expected to respond similarly to treatment. Often, however, additional uses are identified only with substantial delay. New opportunities may emerge as advances in basic science improve understanding of how different diseases are related, or as clinical experience with an approved drug reveals previously unrecognized benefits in new patient populations.[29] Technological change can accelerate both processes by enabling new scientific insights and by uncovering patterns in accumulated clinical data (EveryCure, 2025). Even so, the recognition and validation of new uses often occurs many years—or decades—after a drug's initial approval. As Gelijns et al. (1998) observe, "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."

Aspirin, initially developed as a treatment for pain and inflammation and now widely used to reduce the risk of heart attack and stroke, is widely cited by scientists to support the claim that many existing drugs have additional—yet unrealized—therapeutic uses Gelijns et al. (1998). Several high-profile examples make clear how wide-ranging new uses may be. Thalidomide, initially developed as a treatment for morning sickness (and famously kept off the U.S. market before its global recall in the early 1960s), was approved in 1998 as a treatment for leprosy and again in 2006 as a treatment for

[29] For example, a large recent literature has rethought how cancers are classified and treated. Traditionally, cancers have been classified by the organ in which they first appear (e.g., "breast cancer" or "lung cancer"). Recent work shows that this approach can be misleading: tumors within the same organ can differ substantially, while tumors arising in different organs can sometimes share important similarities. As a result of this shift in how cancers are understood, treatment decisions and drug development have increasingly moved away from organ-based categories and toward shared characteristics across tumors. This has created new, concrete opportunities to develop and test existing drugs for specific patient groups—opportunities that were not meaningfully defined at the time of initial approval (Garraway and Lander, 2013; Hoadley et al., 2014, 2018).

multiple myeloma.[30] Minoxidil, initially approved in 1979 as an antihypertensive, was reapproved in 1988 for male pattern baldness and in 1991 for female pattern hair loss.[31] Recently, the biologic drug omalizumab (marketed under the brand name Xolair), initially approved 2003 to treat allergic asthma and related conditions, was reapproved as the first medicine that reduces the frequency and severity of allergic reactions, including potentially life-threatening anaphylaxis (Wood et al., 2024).[32]

Prominent researchers have argued 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 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).

Despite this scientific promise, investment in the systematic development of new uses for existing drugs remains limited (Eisenberg, 2005; Rai, 2012; Butcher, 2013; Pahud et al., 2014; Roin, 2014; Sachs et al., 2017; Conti et al., 2020; EveryCure, 2025).[33] While a small number of new indications ultimately receive regulatory approval,[34] historical accounts and descriptive evidence suggest that many potential uses identified through advances in basic science and clinical observation are never explored in rigorous clinical trials.[35]

3.2 A Missing Market for New Uses

In this section, we show that the conditions that give rise to a missing market are satisfied in the case of new uses for existing drugs. To illustrate, we walk through each step of the stylized 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

[30] For regulatory documents, see U.S. FDA NDA # 020785, NDA # 021430.

[31] For regulatory documents, see U.S. FDA NDA # 019501, NDA # 020834, NDA #021812.

[32] For regulatory documents, see U.S. FDA BLA #103976. Section 4 describes our construction of datasets that catalog approvals of new uses systematically.

[33] Roin (2014, pp. 18-20, 43) collects detailed histories of potential new uses identified through advances in basic science and clinical observation, as well as dozens of examples of promising new uses that stalled in the commercialization process.

[34] See Section 5.1 for an extended discussion.

[35] Drugs may be prescribed "off label" even where rigorous clinical trial evidence does not support their use for a particular indication. There have been many historical examples where such off label use was ex post validated by strong scientific evidence—for example, the use of an existing heart medication to treat a common childhood vascular condition, later confirmed in large randomized trials (Léauté-Labrèze et al., 2015). There have also been many historical examples where widespread off label use was shown to cause severe harm once more rigorous evidence became available—for example, the widespread use of hormone therapy to prevent heart disease in postmenopausal women, which was ultimately shown to increase the risk of stroke, heart attack, and blood clots (Manson et al., 2003). Thus, while off label use may generate hypotheses or early signals about potential therapeutic applications, it rarely substitutes for investment in large-scale clinical trials—the margin of investment that is the focus of this paper.

carry labels specifying each FDA-approved therapeutic indication.[36] 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 must know when to dispense it, patients must know how to take it, and insurers must know how to process reimbursements. Thus, $p_{\mathsf{secret}} = 0$.

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.[37]

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, and formulations of previously approved drugs. During this three-year exclusivity period, the FDA will not approve applications from generic competitors for the specific use protected by NCI exclusivity.[38] 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.

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

p_{enforce} : **Can the firm profitably enforce any intellectual property grant?** ***Yes, then No.***

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 Hatch Waxman Act of 1984 created the regulatory system for generic drugs in the United States.[39] Its key provisions included a change to clinical trial requirements for generic firms: since

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

[37] In re Marshall, 578 F.2d 301, 304 (C.C.P.A. 1978). Recall that our definition of p_{eligible} captures only subject matter eligibility. New uses may be ineligible for patent protection for other reasons, including failure to satisfy statutory requirements of novelty and non-obviousness. Later in a drug's lifecycle, as observational evidence cataloging potential leads in scientific journals accumulates, both may be important barriers to patent protection. As a general matter, though, federal courts have consistently granted 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"); 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").

[38] 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 C.F.R. §314.108.

[39] 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.).

1984, generic firms have not been required to conduct independent, costly clinical trials as a condition of approval. Instead, sponsors must satisfy two requirements. First, the generic firm must prove that its product is chemically equivalent to the original brand name drug. Second, for whatever scope of approval it seeks—intended use(s), product form, patient population, etc.—it must demonstrate that any relevant patents associated with the brand name drug are invalid, not infringed, or expired. The list of relevant patents is provided by the sponsor of the brand name drug and reported in a regulatory document maintained by the FDA, colloquially referred to as the "Orange Book."[40] While challenges to listed patents are in progress—as required by statute, for at least a thirty-month period—the FDA will not approve a competitor. Until the competitor is approved, patents associated with the brand name drug, including patents on its uses, are thus perfectly enforceable (and enforced, in effect, by the FDA as an intermediary).[41]

Once any generic entry occurs, enforceability of intellectual property rights on new uses collapses. The problem for the brand name firm is two-fold. First, they must identify instances in which a generic drug approved by the FDA for some use a is instead used for the new use b. Second, they must be able to successfully file and win a lawsuit or otherwise put an end to infringement.

Neither is practical. First, the brand name firm can neither observe when the generic is prescribed or purchased, nor what disease target it is used for. Second, even if the brand name firm could observe infringement of a patent on a new use, the nature of liability in this setting makes enforcement unattractive. Generic firms typically rely on the Hatch Waxman Act's "skinny labeling" provisions, which allow them to market drugs without formally infringing method-of-use patents by omitting patented uses from their labels.[42] Courts have consistently shielded generic manufacturers from patent infringement liability unless they explicitly induce infringing use. This means that a generic manufacturer likely 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.[43] 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

[40] We provide details on this patent list in Section 4.4.

[41] 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, and recommend criminal prosecution.

[42] Generic applicants submit "section viii" statements, which certify that patents do not cover the uses of the drug for which the generic seeks approval. The generic firm must submit, alongside a section viii statement, a proposed label for their product that omits any part of the brand name drug's label associated with uses still protected by patent.

[43] Warner-Lambert v. Apotex Corp., 316 F.3d 1348 (Fed. Cir. 2003). A series of Federal Circuit decisions that postdate our analysis—and therefore do not complicate the interpretation of our empirical results—have called into question this approach to so-called "skinny labeling." Most notably, in GlaxoSmithKline LLC v. Teva Pharmaceuticals USA, Inc., 976 F.3d 1347 (Fed. Cir. 2020), the court held that a generic manufacturer could be found liable for induced infringement based on its labeling and marketing conduct notwithstanding a formal carve-out of the patented indication—in effect, narrowing both Warner-Lambert and the practical scope of the Hatch Waxman Act's skinny labeling provisions. For discussion of additional cases in this evolving area of patent law, see Sherkow and Gugliuzza (2026). In January 2026, the Supreme Court granted certiorari in Hikma Pharmaceuticals USA Inc. v. Amarin Pharma, Inc., which squarely presents questions concerning the permissibility and limits of skinny labeling.

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 to sue their own customers en masse. For extended discussions of these institutional details, see Eisenberg (2005) and Roin (2014).

Thus, before generic entry, p_{enforce} takes a value of one: the brand name firm can fully enforce its portfolio of intellectual property rights. After generic entry, p_{enforce} 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 period of time, the innovator has effective market exclusivity. And during this time, patents on new uses are fully enforceable. Since each new use that is commercialized yields monopoly profits until exclusivity expires, incentives for investing in new uses are strongest at $t = 0$. As 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 immediately following 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 decline monotonically towards zero, as the expected time under monopoly (T_{patent}) for a new use decreases with 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.[44] The molecule in Gemzar responsible for its pharmacological action—its *active moiety*—is called gemcitabine.[45]

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

[45] 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 quite technical, it is significant for regulatory determina-

In the 1980s, scientists at Lilly began to develop gemcitabine as an antiviral. In early preclinical studies in the 1990s, however, an internal research group documented that the drug was effective at killing leukemia cells *in vitro*, and effort shifted to investigate the drug's potential as a treatment for cancer (for a history, see Hertel et al., 1990). Lilly sponsored a set of clinical trials that supported its initial approval for the treatment of pancreatic cancer, but also investigated other potential uses, including as a treatment for smallpox, hepatitis, and various types of cancer. The FDA re-approved 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.[46] By 2010, at the time of first generic entry, sales of Gemzar generated approximately $800 million per year.[47]

To test for a missing market for new uses, we would like to collect data on each of these investments into Gemzar, over time and in relation to the end of its period of market exclusivity. In light of Section 3's discussion about the difficulty of enforcing certain pharmaceutical patents, the expiration of all Gemzar patents may be an imperfect proxy for the firm's actual exclusivity period. The timing of generic gemcitabine approval validates this concern. Per the 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.[48] 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.[49] 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, thus did not fully block generic entry.[50]

tions that hinge on a drug's novelty and, as relevant here, for our data construction. 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.

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

[47] 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").

[48] 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 C.F.R. 314.54 lays out the listing requirements for patents and regulatory exclusivities.

[49] 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.

[50] 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.

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 (U.S. Food and Drug Administration, 1985-2014).[51] We classify a drug as 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, complex molecules derived from living organisms—which are subject to distinct legal and regulatory frameworks in the United States.[52]

The second restriction ensures that we focus on research incentives associated with each drug's first approval.[53] 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.

[51] 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 (from the middle of 2016 on), 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).

[52] Biologic drugs include gene therapies, monoclonal antibodies, and vaccines. A different set of statutes govern federal research incentives for biologics, and the 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.

[53] 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.

(2021).[54] 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 (U.S. Food and Drug Administration, 2005-2016b). 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) (World Health Organization, 2019). These alphanumeric codes systematically classify diseases, symptoms, and related health conditions.

4.3 Measuring Research Investments

For each drug in our sample, we construct measures of research and development into new uses. Our primary outcome of interest, FDA re-approvals of existing drugs for new uses, captures trends in successful commercialization. Two secondary outcomes, clinical trials and scientific publications, capture trends in commercialization activity and underlying scientific research.

New Use Re-Approvals We collect records of new use re-approvals—our primary outcome of interest—from FDA administrative records that classify applications into three types: new applications, supplemental applications, and generic drug applications (U.S. Food and Drug Administration, 1985-2014). 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.

Clinical Trials We use clinical trial activity as a proxy for investment in commercialization. To maximize coverage over our sample period, we compile trial records from four sources: NDA Pipeline (F-D-C Reports, 1982-2001), Pharmaprojects (Citeline, 1995-2010), Cortellis (Clarivate, 1995-2016), and ClinicalTrials.gov (National Library of Medicine, 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, which relies largely on standardization by hand, 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 drawn from 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.

[54] 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.

Scientific Publications We collect records of scientific publications from the National Library of Medicine's PubMed database (National Library of Medicine, 1966-2022). 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. This process successfully maps approximately 75 percent of drugs in our final sample to a PubMed term.

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 have no 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.[55]

The most natural empirical proxy for a firm's expectation of the time when it will lose market exclusivity is the date on which generic entry actually occurs. This should reflect market participants' rational expectations. However, the difficulty in using the 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 percent) do not experience generic entry during our sample period. Similar trends in generic entry have been well-documented and, thus, are not peculiar to our data. Although a series of legal reforms in the 1980s—including the 1984 Hatch Waxman Act—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.[56]

An alternative proxy might be the date on which the patent or patents on the drug expire.[57] There are three challenges with constructing such a measure. 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,

[55] Durvasula et al. (2023) offers an extended discussion of both the legal and empirical challenges associated with measuring market exclusivity.

[56] 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). There are fewer small-molecule drugs without generic entry in FDA (2025) than in our data because FDA (2025) covers an additional eleven years of generic drug approvals.

[57] As Section 4.5 discusses, given our substantive interest in the relationship between research investments and periods of market exclusivity provided by intellectual property rights, we restrict consideration to those drugs for which we observe at least one form of intellectual property. Section 4.1 of Durvasula et al. (2023) investigates the very small number of drugs for which there is no linked patent or regulatory exclusivity in the Orange Book.

and these exclusivities may run sequentially or concurrently with patents.[58] Second, new drugs are, increasingly, protected by many patents, not all of which will block generic competition.[59] The median innovator drug approved in 1990 was protected by a single patent grant. In 2014, the median drug had five 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 highlights, not every patent 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.[60] 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. We use the actual date of generic approval if generic entry occurs before the expiration of all of the drug's patents and exclusivities. Otherwise, we assume that firms expect all patents and regulatory exclusivities to bind and use the latest expiration date.[61] Formally, for each drug, we define the year that market exclusivity expires as

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

In effect, our definition of *market exclusivity* selects an upper bound on the firm's expected period of exclusivity. We define an additional measure of exclusivity, *minimum patent exclusivity*, that takes the earlier of generic approval and the expiration date of the first-filed patent. This second measure functions, in effect, as a lower bound. We return to its interpretation in Section 5.2, where it is used in one of our key empirical tests.

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

[58] Regulatory exclusivities are often called quasi-patents, as they also allow a firm to exclude competitors from the market for some fixed period of time. They differ from patents in two key ways. First, regulatory exclusivities in the pharmaceutical sector are awarded by the FDA on the basis of its own assessments of a product (e.g., the FDA itself determines whether a drug is a "new chemical entity" and thus eligible for a five-year exclusivity grant). Second, exclusivities are enforced by agency: so long as exclusivities are in force, the FDA will limit the actions that may be taken by generic competitors.

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

[60] 21 U.S.C. § 355(j)(5)(B)(iv). For instance, the first (successful) generic challenger is potentially eligible for a bounty in the form of exclusive rights to the generic market for a 180-day period.

[61] 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 (NCE), orphan drug exclusivity (ODE), pediatric exclusivity (PED—only instances of PED that extend a form of exclusivity that is generic blocking), and Generating Antibiotic Incentives Now exclusivity (GAIN). Technically, there are sixteen types of FDA-granted regulatory exclusivities. Table 5 in Durvasula et al. (2023) provides a list.

pathway.[62] 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 (U.S. Food and Drug Administration, 1984-2016a). Appendix Figure A1 plots the distribution of this market exclusivity measure for our sample of innovator drugs. The median drug in our sample receives 13.37 years of market exclusivity (mean, 12.90 years) and receives 8.37 years of protection from its first-expiring patent (mean, 9.35 years).

4.5 Sample Construction & Summary Statistics

There are 1,102 drugs that meet our criteria for innovator drugs. We further 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.[63] This yields a final sample of 990 drugs. Table 1 presents summary statistics.

5 Empirical Evidence

This Section provides three pieces of empirical evidence that show that there is a missing market for new uses of existing drugs. We begin in Section 5.1 by establishing that each of the three empirical signatures of a missing market for new uses, as introduced in Section 3.3, is borne out in our data: firms have the strongest incentives to invest in new uses soon after initial approval; incentives decline as market exclusivity expiration approaches; and incentives fall to zero after exclusivity expiration. Section 5.2 presents causal evidence that extending the duration of market exclusivity increases the number of new uses that are commercialized. Section 5.3 shows that, at the time of exclusivity expiration, investment activity shifts toward projects for which some enforceable intellectual property protection remains. Section 5.4 argues that these empirical patterns are inconsistent

[62] 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 (2012) 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.

[63] There are nine drugs that are excluded because their periods of measured market exclusivity are 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). For each, we review FDA records and drug approval histories to understand why generic entry appears to predate NDA approval. 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 capture an earlier NDA approved by the pre-1962 FDA. Generic versions of these drugs were approved before the NDAs that we flagged as innovators. There are a small number of other conceptual difficulties that arise when drugs in our sample are grandfathered into new regulatory schemes introduced in the 1960s and 1980s. We offer an extended discussion in Durvasula et al. (2023).

with alternative explanations.

5.1 Empirical Signatures of a Missing Market

Figure 2, Panel A and Figure 3, Panel A present trends in the probability that a drug is re-approved for any new use in each year (y-axis). Panel A of Figure 2 plots this probability in each year relative to the drug's initial approval (x-axis). Consistent with our predictions, the probability of any re-approval peaks shortly after initial approval, then declines monotonically. At this peak ($t = 2$ through $t = 5$ years after initial approval), approximately five percent of drugs are re-approved for a new use each year. Recall that different drugs in our sample receive different periods of market exclusivity, with a median and mean of about 13 years. Within 15 years of initial approval, the likelihood of re-approval for a new use is approximately zero.

Panel A of Figure 3 plots the same data capturing probability of re-approval for a new use, but instead uses the time relative to the drug's market exclusivity expiration as the x-axis. With this reorientation, note again that the probability that a new drug is re-approved for a new use peaks early in its lifecycle, declines monotonically as exclusivity approaches, and then is approximately zero after exclusivity expiration. The peak is ten years prior to exclusivity expiration ($t = -10$), when seven percent of drugs are re-approved for at least one new use.[64] 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 for five- and ten-year windows around expiration. Appendix B shows that our results are robust to several other cuts of the data. We also examine trends in two secondary measures of research investments, clinical trials and scientific publications. Figure 5 plots trends in both measures relative to market exclusivity expiration.[65] Panel A shows that the probability an innovator drug is studied in any active clinical trial peaks 18 years prior to exclusivity expiration when nearly half of all innovator drugs are under study in at least one trial, then declines steadily across the drug's lifecycle. As our clinical trial variable is noisy and captures both trials intended to establish the efficacy of existing drugs for new uses and also many other categories of trials, we do not expect this series to fall to zero after

[64] In Figure 3, we include each of the 990 innovator drugs in our primary sample for every year in which data are available. This means that the sample of drugs shifts across points in Figure 3. This changing sample composition generates the non-monotonic trend between $t = -20$ and $t = -17$. Consider, for instance, a drug approved in the year 2000 that receives ten years of market exclusivity, and recall that our data include observations through 2014. In Figure 3, this drug appears in the sample plotted at each point between $t = -10$ and $t = 4$. Other versions of this exercise aim to address these changes in sample composition in various ways. Figure 4 takes a balanced panel approach, in which we require data on each drug in each year and confirms that this non-monotonicity is corrected when we address changing sample composition. Appendix B.1 collects additional exercises that cut our data differently. Appendix Figure A3 splits each drug's market exclusivity into five periods and plots trends across periods, rather than over calendar time. With this modification, we find, as predicted, that the probability of approval is highest just after initial approval, then declines monotonically before falling nearly to zero following exclusivity expiration. Appendix Figure A4 addresses the issue of changing sample composition by inspecting trends in the probability of re-approval for drugs that receive exactly the same period of market exclusivity (e.g., for the cohort of drugs that receive 12 years of exclusivity). This restriction eases interpretation at the cost of substantially increasing the noisiness of these plots; though each cohort restriction drops upwards of ninety percent of our data, the three empirical signatures are qualitatively unchanged.

[65] Appendix Figure A6 plots trends in both measures relative to each drug's year of initial approval.

exclusivity expiration; indeed, the likelihood of an active clinical trial remains positive across our sample period.[66] Panel B of Figure 5 plots the number of scientific publications studying each innovator drug in each year relative to exclusivity expiration. Publication activity declines by roughly a factor of two over the twenty-year window around exclusivity expiration. Appendix Figure A7 clarifies the source of this decline by separately plotting trends for publications with and without public funding. The volume of publicly funded publications remains essentially constant—around ten per year across the entire period. The decline is driven entirely by a drop off in privately funded publications, suggesting that scientific attention persists even as private interest wanes.

While Figure 2, Panel A and Figure 3, Panel A focus on the likelihood that any new use is re-approved in each year, Panel B in each presents the cumulative share of new uses approved over time. Of the 511 new uses of innovator drugs approved by the FDA in our data, roughly 95 percent were approved before market exclusivity expiration. Put differently, less than 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.

Figures 2 through 4 document patterns that are consistent with—though not dispositive of—the existence of a missing market for new uses. There remain two sets of questions. First, do trends in FDA approvals fully capture relevant research and development? Second, could alternative explanations—such as a natural decline in scientific leads over a drug's lifecycle, or strategic sequencing of investments to commercialize uses before exclusivity expiration—account for the observed patterns? We take up the first question here, then turn to the second in Section 5.4.

On the question of whether our data capture all relevant research and development, one possibility is that firms shift investment away from the original innovator product at exclusivity expiration and continue developing new uses under separate NDAs for the same active moiety. Our main analysis tracks re-approvals associated with the 990 innovator drugs in our primary sample for which measures of market exclusivity are cleanly defined. Any reallocation of research effort would lead our analyses to underestimate investment. In general, the scope for this concern in our data appears limited: the median innovator drug is associated with a single NDA.[67] In Appendix B.4, we show that the three empirical signatures are unchanged when we track re-approvals at the active-moiety level: 92 percent of new use approvals occur before exclusivity expiration with this change, compared to 95 percent when we track re-approvals at the innovator drug level. This pattern is consistent with our model's prediction that incentives to invest in new uses tied to any form of an active moiety collapse once generics enter, since generic competitors are substitutes for all formulations of the drug.

Taken together, trends in other measures of research investment are consistent with the concern that private firms stop investing in new uses after exclusivity expiration. The remainder of this Section investigates whether alternative explanations, other than a missing market created by

[66] Our measure of clinical trial activity captures, for example, early-phase trials, large late-phase efficacy trials, post-approval surveillance studies, academic studies intended to inform clinical practice, and trials required by foreign regulators. For details on our clinical trial measure, see Section 4.3 and Appendix A.3.

[67] Appendix A.1.4 offers additional details.

unenforceable intellectual property rights, could account for the observed patterns.

5.2 Estimating the Causal Effect of Exclusivity on Research Investment

If gaps in private incentives cause the trends in Figures 2 through 5, then drugs with longer periods of market exclusivity should have a larger number of realized new uses, as longer exclusivity periods (T^*) increase expected profitability ($T^*\pi$). 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

We leverage a source of variation in the *effective* market exclusivity provided to innovator drugs that arises from the interaction of patent law and regulatory law. In the pharmaceutical sector, patent protection is typically secured early in the research and development process, while regulatory approval is obtained only after the completion of extensive preclinical and clinical testing. Because patent terms elapse during this regulatory development period, the length of required clinical testing mechanically determines the amount of effective protection remaining when a drug is ultimately approved. Put differently, drugs that require longer clinical trials to establish safety and efficacy enter the market with shorter remaining periods of effective patent protection.

Pharmaceutical firms face strong incentives to file for patent protection very early in the research and development process—in the notation of Section 2, at t_{invent} rather than at t_{comm}. Delays in patent filing risk not only that a competitor will patent first, but also that the firm's own subsequent activities may enter the prior art record and, thus, undermine the drug's patentability. In particular, disclosures associated with drug development—including scientific presentations, publications, and the legally mandated public disclosure of clinical trial protocols and results—can render later-filed claims non-novel or obvious.[68] As a result, core patents—most often, compound

[68] Patent validity is assessed relative to the body of publicly available information at the time of filing, and later-filed claims may be invalidated if intervening disclosures render them non-novel or obvious. See 35 U.S.C. §§ 102, 103. Courts have interpreted "prior art" broadly to include a wide range of publicly accessible research disclosures, including scientific presentations and other materials disseminated in the course of research and development. See, e.g., In re Hall, 781 F.2d 897 (Fed. Cir. 1986); In re Klopfenstein, 380 F.3d 1345 (Fed. Cir. 2004).

In the pharmaceutical sector, this constraint is particularly salient because regulatory requirements mandate increasing public disclosure as a drug advances through clinical testing. Courts have treated disclosures associated with clinical trials as capable of anticipating or rendering obvious later-filed claims, even where those disclosures were not directed at commercialization. See Pfaff v. Wells Elecs., Inc., 525 U.S. 55 (1998); Helsinn Healthcare S.A. v. Teva Pharms. USA, Inc., 586 U.S. 123 (2019). In addition, under the doctrine of inherent anticipation, prior disclosures may defeat later claims even if the therapeutic significance of the invention was not recognized at the time. See Schering Corp. v. Geneva Pharms., Inc., 339 F.3d 1373 (Fed. Cir. 2003).

Recent decisions suggest that these principles may extend further still. In a 2025 inter partes review—an administrative proceeding before the Patent Trial and Appeal Board that permits third parties to challenge the validity of issued patents—case involving patents held by Johns Hopkins University, the Patent Trial and Appeal Board treated the public disclosure of a clinical trial protocol itself, as distinct from any trial results, as anticipatory prior art and rejected arguments that such disclosures were insulated by experimental-use considerations. See Merck Sharp & Dohme LLC v. The Johns Hopkins Univ., IPR2024-00240, Paper 90, at 28–31 (P.T.A.B. June 9, 2025); Merck Sharp & Dohme LLC v. The Johns Hopkins Univ., IPR2024-00647, Paper 53, at 30–32 (P.T.A.B. Sept. 23, 2025).

or molecule patents that provide the most certain protection from generic competition (Lemley and Shapiro, 2005; Hemphill and Sampat, 2011; Durvasula et al., 2023)—are typically filed prior to the initiation of clinical trials (Mossinghoff, 1999; Patrick, 2005; Thomas, 2007).

The effective period of protection provided by these core patents, then, is determined by the length of the regulatory development process required to bring the drug to market. Approval requires completion of a sequence of preclinical and clinical testing stages whose duration varies substantially across drug–disease pairs and reflects how quickly therapeutic efficacy can be credibly established. For some conditions, relevant clinical endpoints are observable over relatively short horizons; for others—particularly chronic or slowly progressing diseases—demonstrating efficacy requires extended follow-up, larger samples, or both (Roin, 2013; Budish et al., 2015a). These features of the regulatory process are tied to disease biology, drug characteristics, and prevailing evidentiary standards; from the firm's perspective, conditional on pursuing a particular drug-disease pair, the required duration of clinical testing is essentially fixed.[69]

These institutional features motivate our empirical test. To isolate the component of market exclusivity determined by the interaction of early patent filing and heterogeneous regulatory development timelines, we focus on the measure of minimum patent exclusivity defined in Section 4.4, which captures the period of protection provided by a drug's earliest-filed patent and thus excludes potential exclusivity extensions provided by subsequent intellectual property grants.

Consistent with the heterogeneity in testing requirements across drug-disease pairs, this minimum patent exclusivity measure varies substantially across our sample: the average innovator drug in our sample has 9.4 years of minimum patent exclusivity, while drugs at the 10th and 90th percentiles receive 5 and 15 years, respectively.[70] The key identifying assumption is that cross-drug variation in regulatory development timelines is not systematically related to underlying scientific potential for new uses.

Panel A of Figure 6 plots the relationship for drugs in our sample between this measure of minimum patent exclusivity and the number of realized new uses as a binned scatterplot with an overlaid line of best fit. Panel A of Table 2 formalizes this relationship in a regression framework.

[69] FDA regulations require "adequate and well-controlled" studies designed to distinguish treatment effects from other influences, including spontaneous change in the course of the disease. 21 C.F.R. § 314.126. FDA guidance on "substantial evidence" similarly emphasizes that the evidentiary showing depends on disease context and the clinical meaningfulness and measurement properties of endpoints. FDA, Demonstrating Substantial Evidence of Effectiveness for Human Drug and Biological Products (Sept. 11, 2023). See also FDA, Early Alzheimer's Disease: Developing Drugs for Treatment (Mar. 5, 2024) (recognizing that longer trials may be needed in early AD given minimal/absent deficits and the challenge of establishing clinically meaningful effects).

[70] Appendix Figure A2 plots the distribution of this measure. In 18 cases (two percent), 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 categorize Orange Book patents as either method-of-use or 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.

In particular, for innovator drug i, we estimate the regression

$$y_i = \beta \cdot \text{exclusivity}_i + \mu + \varepsilon_i \tag{6}$$

where the dependent variable y_i captures the number of re-approvals of innovator drug i for a new or modified indication and μ is an approval year fixed effect. The estimate of interest is β, the coefficient on our measure of each drug's market exclusivity duration in years, exclusivity_i. Columns (1) and (2) report results using minimum patent exclusivity; Columns (3) and (4) instead use the measure of market exclusivity defined in Section 4.4 as 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.

In what follows, we focus on the interpretation of the estimate in Column (2) in Table 2, as the specification with an approval year fixed effect accounts for the concern that earlier-approved drugs have had more years in which to realize additional new uses.

There is one additional econometric challenge for this exercise, which arises because nearly three-quarters of innovator drugs in our sample ($N = 725$, 73 percent) have zero re-approvals for new uses: the existence of a mass point at zero may bias our estimates (e.g., Mullahy, 1998). To address this concern, in Panel B of Table 2, 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—0.026 in Column (1) and 0.042 in Column (2)—suggest that an additional year of market exclusivity increases the odds of observing any re-approval for a new use by a factor on the order of two to four percent.[71] Though the intensive margin restriction leaves us underpowered, 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 investments in research and development? That more new uses are commercialized when drugs receive longer periods of market exclusivity provides, to our knowledge, the first direct evidence that the answer to this question is yes.

[71] 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.

5.2.2 Interpreting these estimates as an elasticity

This empirical exercise also yields an estimate of the elasticity of research investments with respect to market exclusivity. This elasticity is a key parameter for the design of optimal patent policy that has been challenging to estimate (Nordhaus, 1969; Budish et al., 2016).[72] Though it is necessarily specific to our empirical context, 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, to offer a rough quantification of the social costs of inadequate incentives for investment.

Returning to Table 2, Panel A, Column (2), 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. Denote by $\bar{y} = 0.52$ the average number of new use approvals per innovator drug and denote by $\bar{x}$ the average years of exclusivity (9.35 for minimum patent exclusivity and 12.90 for our market exclusivity measure that includes second patents and exclusivities). The point elasticity at the mean is given by $\beta \cdot \frac{\bar{x}}{\bar{y}}$. Given our coefficient estimate of 0.02, this implies elasticities of approximately 0.35 and 0.49, using respectively our minimum patent exclusivity measure and standard market exclusivity measure.[73] These estimates suggest that doubling the average market exclusivity would increase the expected number of re-approvals by approximately 35 to 49 percent.

Another interpretation of the coefficient estimate of 0.02 is that, relative to the sample mean of 0.52 new use approvals per drug, adding one year of market exclusivity increases the average number of approvals for new uses by roughly 3.8 percent. It is this semi-elasticity—expressed as a percent change in realized new uses per additional year of exclusivity—that we compare to prior estimates in the literature.

Budish et al. (2015a) and Abrams (2009) offer two points of comparison. Budish et al. (2015a) 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 percent. 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_{patent} in Section 2. In that context, a one year

[72] A survey article by Budish et al. (2016) writes: "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."

[73] Minimum patent measure: $= 0.02 \times \frac{9.35}{0.52} \approx 0.35$.
Market exclusivity measure: $= 0.02 \times \frac{12.90}{0.52} \approx 0.49$.

increase in patent term is estimated to generate a 66 percent increase in patent filings.

That our estimate is smaller than that in either Budish et al. (2015a) 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. (2015a) 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 as well, given the context of the policy change used to generate this elasticity: 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. It is unsurprising that our estimate, drawn from a more precise research design, is smaller in magnitude.That being said, a second explanation for the difference in magnitude is that, in many ways, our context is narrow. Inventions in our setting are clinically validated links between drugs and new diseases, which in effect function as improvements to existing products.

Eisenberg (2005) provides a third 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 the 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.

Not all re-approvals of existing drugs face the same enforceability problem at the time of generic entry. Crucially, enforceability depends on the object of protection. Generic competitors must individually challenge patents and demonstrate bioequivalence for each distinct formulation. If other patents and exclusivities that protect certain dosages, forms, and patient populations of a brand name drug remain in force, those dimensions of the product remain legally and practically excludable, and generic products will be 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. Similarly, for many innovator drugs in our data, generic competitors are initially approved only for a subset of patient populations (e.g., adult but not pediatric) or delivery forms (e.g., tablet but not liquid).

Thus, at the time of initial generic entry, intellectual property rights over certain drug features—such as formulation, dosage, or delivery—may remain partially enforceable ($p_{\text{enforce}} \in (0,1)$), even as intellectual property rights over information about new therapeutic uses become effectively un-

enforceable ($p_{\text{enforce}} = 0$).

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 Panel A of Figure 7. Visually, we find a much more gradual decline in re-approvals around the time of market exclusivity expiration than that documented in Figure 3, 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 linked to innovator drugs 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.[74] The widely-used HIV treatment and preventive Truvada, for instance, is a single tablet that contains two distinct drugs, emtricitabine and tenofovir disoproxil fumarate, neither of which on its own offers the same therapeutic benefits.[75] Even if the components of a fixed dose combination are available in generic form, the combination form is often preferred by patients and providers, as it can simplify complex treatment regimens and increase prescription adherence.[76] Existing generics, then, are not perfect substitutes.

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 Panel B documents that, in the year of exclusivity expiration, the likelihood that a new indication for a combination drug is approved more than doubles relative to the previous year. 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 additional 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

[74] Note a point of potential confusion. A recent paper by Dix and Lensman (2024) studies incentives to invest in combinatorial innovation using data on combinations of existing cancer drugs. These kinds of combinations are distinct in form, function, and regulation from the fixed dose combinations that we consider in this test of our mechanism. Combination cancer therapies include components that are dispensed separately in in-patient settings; each component is regulated as a distinct drug. Fixed dose combinations are dispensed as single products, which are regulated themselves as new drugs.

[75] Truvada (emtricitabine; 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. Emtricitabine 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.

[76] 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).

market exclusivity expiration suggests that we are accurately capturing firm expectations.[77]

5.4 Discussion

On its own, the evidence in Section 5.1 provides strong evidence of a missing market for new uses. New use re-approval patterns in our data closely track the empirical signatures discussed in Section 3.3: re-approvals are concentrated soon after initial approval, decline as exclusivity expiration approaches, and fall to approximately zero thereafter. If opportunities to commercialize new uses arise with substantial delays that exceed the duration of market exclusivity—as basic science progresses and clinical experience accumulates (see Section 3.1)—then the sharp drop-off in re-approvals at the time of exclusivity expiration is prima facie evidence that private investment does not fully track scientific opportunity.

There are two potential alternative explanations for some of these patterns. If the set of scientific opportunities for new uses is gradually exhausted, research investments may naturally decline over a drug's lifecycle. This story easily fits the first two empirical signatures (i.e., a level of investment that is initially high then declining), but to rationalize the third requires that scientific opportunities coincidentally drop off precisely when enforceable intellectual property protection drops off. This story also fails to rationalize the evidence from Sections 5.2 and 5.3. Were it the case that the three empirical signatures in Section 5.1 captured only the natural depletion of potential new uses, we would not expect to find a relationship between the duration of market exclusivity and the number of commercialized uses (Section 5.2). Additionally, we would not expect a coincidental spike in the scientific opportunities associated only with those products on which intellectual property rights remain enforceable (Section 5.3).

Alternatively, a second story is that strategic firms that expect generic entry may sequence research and development to bring all viable new uses to market before exclusivity expires. This story is arguably consistent with all three empirical signatures documented in Section 5.1. Indeed, our framework assumes that firms are strategic. However, the key implication of this alternative is that there are no scientifically feasible new uses remaining at the time of market exclusivity expiration. That is, there is no missing research and development. This is inconsistent with the evidence in Sections 5.2 and 5.3. It is also inconsistent with the widely accepted idea that opportunities for new uses often arise only with substantial delay (see Section 3.1).

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

[77] 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.

investment with respect to market exclusivity calculated in Section 5.2 to develop a rough quantification of the social value of this missing innovation. We proceed in three steps. First, we use our elasticity estimate to estimate how many additional new uses would have been realized, had they received longer periods of market exclusivity. Second, we use estimates from the existing literature to give a sense of magnitudes for the value of a missing use. Last, we multiply the estimates from the first two steps and take a present value. Given the considerable uncertainty in each stage of this exercise, our objective in this section is not to arrive at a point estimate that captures the extent of missing research and development but, rather, to offer a sense of magnitudes.

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. We use this elasticity to develop an estimate of how many new uses would have been developed had intellectual property rights on these inventions been fully enforceable.

To estimate the number of missing new uses, ideally, we would ask how many additional new uses would have been commercialized had intellectual property rights on new uses been enforceable for some number of years after approval. As a lower estimate, how many additional new uses would have been developed if new uses received five additional years of enforceable protection following the expiration of any existing exclusivity? Five-year extensions are 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). And as an upper estimate, how many additional new uses would have been developed if enforceable protection for uses lasted an additional twenty years (i.e., the duration of a statutory patent term)?

To operationalize this thought experiment using observable variation in our data, we use the measured relationship between the number of new uses and the period of minimum patent exclusivity for the innovator drug reported in Figure 6 and Table 2—in effect, shifting each observation in Figure 6 to the right by the relevant number of years. We emphasize that this exercise is intended to offer a sense of the magnitude of missing research and development, not to analyze a specific policy. Section 7 takes up the question of potential policy mechanisms.

Estimates from Columns (2) and (4) of Table 2, Panel A imply that a five year extension would yield 0.10 new uses per innovator drug (i.e., $0.02 \times 5 = 0.10$) and that a twenty year extension would yield 0.40 new uses per drug (i.e., $0.02 \times 20 = 0.40$). Since 1962, when the modern FDA was established, the agency has approved roughly 2,000 drugs that satisfy our sample criteria. Thus, our estimates indicate that between 200 and 800 additional new uses would have been developed had enforceable intellectual property rights for new uses been extended along these lines.

6.2 How Valuable are New Uses?

Though previous studies have concluded that, in general, new drugs significantly increase life expectancy, there is enormous heterogeneity in estimates of social value. 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; 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 immediately following initial approval. Other studies suggest that many "me-too" drugs—products that provide little clinical differentiation from existing therapies—contribute little in the way of social value (Dranove et al., 2014). This heterogeneity in therapeutic value is compounded by an additional complication: new drugs that offer limited therapeutic advancement may nonetheless represent real scientific advancements, which offer little incremental benefit for current cohorts of patients but enable important follow-on innovation.[78]

Our goal with this back-of-the-envelope calculation, then, is modest: to implement a transparent calculation that offers a sense of the magnitude of underinvestment. Rather than attempting to aggregate estimates from available case studies, we follow a standard approach in studies of cost-effectiveness, which values new therapeutics in terms of quality-adjusted life years (QALYs) that capture improvements in both well-being and life expectancy in standardized units.

Based on a survey of existing literature, we select values for the following: (i) the incremental QALYs generated by a new drug relative to the standard of care; (ii) the social value of one QALY, per patient, per year; and (iii) the patient population treated. Throughout, where the literature admits a range of plausible values, we err on the side of conservative assumptions. These quantities, scaled by a factor that reflects the lower average value of new uses compared to initial approvals, give our estimate of social value as $V = Q \times P \times N \times \theta$, where Q denotes incremental QALYs per patient, P is the value per QALY, N the number of patients, and θ a scaling factor for new uses.

Shafrin et al. (2023) estimate that the average drug approved by the FDA between 2011 and 2021 added 1.04 QALYs per patient relative to the standard of care, based on published cost-effectiveness analyses available for roughly half of all approvals.[79] For our back-of-the-envelope calculation, we round this to $Q = 1$. In the United States, cost-effectiveness analyses typically value one QALY at $P = \$100,000$.[80] For patient population, we assume that $N = 10,000$, a conservative estimate given industry reports suggest that specialty drugs are used to treat 10,000 to 50,000 patients per year, with an average market size of 40,000. (IQVIA Institute for Human Data Science, 2021). Under these assumptions, the annual social value of an initial approval is roughly $1 billion ($1\times$ 100,000

[78] 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."

[79] The authors note that many drugs provided less than 0.5 QALYs and a small number provided multiple QALYs; drugs at the 25th percentile provided 0.33 QALYs and drugs at the 75th percentile provided 2 QALYs. Note that the Shafrin et al. (2023) sample does not include the restriction that new drugs are "innovators" (the first approvals of a unique combination of active moieties) and hence is more expansive.

[80] 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."

$\times$ 10,000).

We next scale these values to reflect existing evidence that new uses are, on average, less valuable than initial approvals. Vokinger et al. (2023) find that roughly one-third of new uses have high therapeutic value compared to half of initial indications, based on a sample of drug approvals from the United States and European Union between 2011 and 2020. In oncology, Michaeli et al. (2022) find that first indications deliver nearly twice the QALY gain per patient—around 0.99 QALYs—relative to second or third indications, which average approximately 0.5–0.6 QALYs.[81] On the basis of these estimates, and recognizing the considerable heterogeneity across indications, we set $\theta = 0.5$, thus assuming that new uses are half as valuable as initial approvals on average. Applying this factor implies that the average social value of a new use is \$0.5 billion.

6.3 Valuing the Missing Market for New Uses

To quantify the social value of missing investment in new uses, we begin by calculating the annual social value of missing new uses. Using estimates from Section 6.1 and Section 6.2, then, we multiply \$0.5 billion per new use per year (Section 6.2) by the 200–800 new uses that we estimate could have been developed with alternative exclusivity periods (Section 6.2). This yields an estimate of \$100–400 billion in missing social value per year.

To translate these annual flow estimates into a measure of total social value, we compute the present value of the missing stream of benefits using a standard constant-growth formulation. With a social discount rate r and long-run population growth rate g, the present value of the missing market for new uses is given by the annual social value divided by $(r-g)$.

With standard, conservative values for r and g—$r = 0.05$ and $g = 0.01$—the total social value of the missing market for new uses is on the order of \$2.5 to \$10 trillion. This estimate is comparable to estimates of the value of missing medical innovation in existing work (see, for example, Murphy and Topel (2006), Hall and Jones (2007) and Budish et al. (2015a)).

It is useful to benchmark this magnitude relative to the existing stock of approved drug–disease combinations. There are approximately 2,000 existing innovator drugs, each with an average of roughly 1.5 approved uses, implying approximately 3,000 approved drug–disease combinations in the current market. Against this benchmark, the 200–800 missing new uses that we estimate correspond to an expansion of the set of approved uses by roughly 7–25 percent. Assuming once more that new uses generate roughly half the social value of initial uses, this implies an increase in aggregate social value on the order of 3–12 percent relative to the value generated by currently approved uses.

[81] 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. In roughly one quarter (26 percent) 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 percent that of each drug's initial indication. This estimate is consistent with the findings of a similar exercise in Berger et al. (2021), which examines the effects of re-approvals for new uses on market size.

While our estimates are necessarily imprecise, intended only to give a sense of magnitudes, the qualitative conclusion is clear: in just one market, the social cost of the missing markets problem appears to be very large.

7 Fixing Missing Markets

The potential scale of the missing markets 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, typically deployed 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.

Our model does, however, offer a structured way to identify alternatives. Consider a stylized version of (4) that captures the conditions under which a firm receiving a grant of intellectual property will commercialize an invention—if and only if

$$p_{\text{eligible}} \cdot p_{\text{enforce}} \cdot (\text{timing parameters})\pi - c \geq 0.$$

To create incentives for innovation, policymakers could intervene on the intellectual property parameters that directly generate the missing market, p_{eligible} and p_{enforce}, the profitability of commercialization π, or the cost of commercialization c. In certain cases, it may also be possible for policymakers to intervene on p_{secret}, creating trade secrecy protection where it otherwise did not exist. We discuss each of these possibilities, briefly, in the remainder of this section.

7.1 Intellectual Property: p_{eligible}, p_{enforce}, and p_{secret}

Eligibility (p_{eligible}) Changes to p_{eligible} are a topic of active debate. A series of Supreme Court decisions imposed novel restrictions on patentability in the early 2000s,[82] and policymakers have proposed new legislation that would expand the scope of patentable subject matter.[83]

Expanding eligibility entails a Nordhaus-style tradeoff: increasing the scope of patentable inventions raises expected returns and can stimulate investment at the margin, but it risks over-rewarding inframarginal inventions that would have been developed without stronger protection (Nordhaus, 1969; Budish et al., 2016). Broader eligibility may generate both static deadweight loss and dynamic costs if higher prices or broader exclusion slow cumulative innovation (Scotchmer (1991); Williams (2013); Sampat and Williams (2019)). Importantly, because patent law, in general, applies uniformly across technologies, reforms that expand eligibility to address missing markets in one domain—such as diagnostics or software—inevitably affect others.

[82] See Section 2.2 for an overview.

[83] See, for example, S. 1546 – Patent Eligibility Restoration Act of 2025.

Enforceability (p_{enforce}) Alternatively, policymakers could target enforceability by expanding opportunities to observe and police instances of infringement—both through broad changes to legal doctrine and court practice that currently make enforcement difficult, and through context-specific interventions tailored to particular markets.

To illustrate, consider the case of new uses for existing drugs. Here, enforceability fails not because patents are unavailable, but because infringement is challenging to observe and address: physicians can prescribe generic drugs "off-label" for unapproved indications, and firms cannot observe or prevent such uses. Policy responses that increase observability could, in principle, restore some fraction of p_{enforce}. Integrated electronic health records, for example, might allow firms or regulators to monitor prescribing patterns and facilitate indication-specific reimbursement, effectively enabling pricing that reflects use. Roin (2014) sketches how such data infrastructures could be designed, and Pearson et al. (2016) describe one attempt to implement indication-based pricing for cancer therapies in Italy.

Changes in legal doctrine could have similar effects. Adjustments to statutory and judicial rules that govern infringement liability—such as limiting the scope of "skinny labeling" exemptions under the Hatch-Waxman Act or expanding inducement liability for generics (see Sherkow and Gugliuzza (2026))—would likewise increase the effective enforceability of intellectual property rights in practice, even without altering their formal scope.

As with changes to eligibility, these interventions entail a Nordhaus-style tradeoff: increasing enforceability raises expected returns and thus can induce investment at the margin, but it also risks over-rewarding inframarginal inventions that would have been developed without stronger protection. Many context-specific costs may also be nontrivial. Enhanced observability requires extensive data collection and surveillance, with corresponding privacy and governance risks. Legal reforms that expand the opportunities for brand-name firms to litigate could slow generic entry, increase compliance costs, and raise drug prices for patients.

Secrecy (p_{secret}) A third possibility is to alter the environments in which secrecy can substitute for formal intellectual property rights. Although the probability that an invention could be profitably commercialized under secrecy p_{secret} does not appear in this stylized equation, it is implicit. Recall that, in our framework, a firm will choose secrecy if it offers a longer expected period of time under monopoly than does a grant of intellectual property rights. When intellectual property rights are unavailable or unenforceable, it may be possible to create innovation incentives by enabling secrecy where it was not previously available.

Recent advances in cryptography—specifically, "zero knowledge proofs," which allow one party to prove to another that they possess certain information without revealing the information itself (see e.g., Goldwasser et al., 2019)—suggest one potential class of mechanisms. Of course, in many domains—including our empirical context—complete secrecy is infeasible, as commercialization requires disclosure to regulators, intermediaries, or end users.

As with changes to eligibility or enforceability, expanding opportunities for secrecy requires

consideration of a Nordhaus-style tradeoff. With secrecy, however, there is a second key cost. One of the theoretical benefits of the patent system is its requirement of disclosure: in exchange for temporary exclusivity, inventors must make their discoveries public, allowing knowledge to diffuse and enabling follow-on innovation (Williams, 2017). Secrecy eliminates disclosure. Policies that make secrecy more viable can therefore incentivize innovation where rights are unenforceable, but do so with both traditional Nordhaus costs and reduced knowledge spillovers.

7.2 Pull Incentives

When changes to p_{eligible}, p_{enforce}, and p_{secret} are infeasible or unattractive, "pull mechanisms"—which reward successful outcomes ex post—can substitute for enforceable intellectual property rights (see, for example, Barder et al., 2005 and Snyder et al., 2020). Any pull mechanism must replicate the functions of the parameters p_{eligible}, p_{enforce}, and π in our framework. A funder must determine the conditions under which an inventor is eligible for, say, a prize or advanced market commitment payout; develop a credible mechanism that guarantees compensation to the first inventor who satisfies the requisite conditions (thereby enforcing against copycats); and offer a payment that substitutes for the quasi-rents promised by intellectual property rights.[84]

Setting the reward size is the core design challenge. If the reward is too small, few (if any) innovators will participate; if it is too large, the funder overpays. Optimal design therefore involves balancing the innovation gains from higher rewards against the fiscal cost of overpayment. We formalize this tradeoff with two stylized models below.

7.2.1 Full Information Case

Let V denote the expected social value of successful investment, i.e., the prize to society if the pull mechanism indeed successfully elicits the hoped-for invention. Let there be a single representative risk-neutral firm considering whether to proceed with research and development. As above, let c denote the expected cost of investment, and let p_{comm} denote the probability of success. We assume throughout that $V > c$, i.e., the social value of successful investment exceeds its cost; else, the social planner clearly should not bother.

Consider first the case where society has full information about the probability of success p_{comm} and the cost of investment c. Society wants the representative firm to invest if and only if

$$p_{\text{comm}} V - c \geq 0. \tag{7}$$

Let Π denote the prize society pays for a successful invention. The representative firm will invest if and only if

$$p_{\text{comm}} \Pi - c \geq 0. \tag{8}$$

[84] For details on how these three elements were operationalized to incentivize investment in the pneumococcal vaccine, see Kremer et al. (2020).

Under the maintained assumption of full information about c and p_{comm}, the optimal pull mechanism takes a very simple form: if condition (7) obtains and society wants to elicit the investment, then society sets a prize such that condition (8) obtains with equality. The optimal prize is

$$\Pi^* = \frac{c}{p_{\mathsf{comm}}}. \tag{9}$$

Intuitively, society elicits the invention by offering a prize equal to the firm's risk-adjusted cost. Next, we consider a stylized version of incomplete information.

7.2.2 Incomplete Information Case

Now assume that society knows V and c, as above, but not the probability of success p_{comm}. This may capture cases in which the social planner can estimate research and development costs ex ante or audit costs ex post, but nonetheless intrinsically knows far less about the probability of successful investment than the firm itself. We can solve for the choice of Π that minimizes the worst-case loss relative to the first best under full information, considered over all possible realizations of p_{comm}.

If the planner sets Π too low, it loses the social value of the invention, $p_{\mathsf{comm}}V - c$. In the worst case, its choice is too low by just an epsilon (i.e., $p_{\mathsf{comm}}\Pi - c = -\varepsilon < 0$), and its loss is therefore $p_{\mathsf{comm}}V - c = c(\frac{V}{\Pi} - 1)$ as we take ε to zero.

If the planner sets Π too high, it overpays relative to what was needed. That is, it pays Π but could have paid $\frac{c}{p_{\mathsf{comm}}}$. In the worst case of overpayment, $p_{\mathsf{comm}} = 1$ and society has overpaid by $\Pi - c$. Society elicits the invention for a social gain of $V - \Pi$ but could have gotten a larger social gain of $V - c$.

In Appendix B.5 we prove that the worst-case optimal choice of Π takes the following simple form:

Proposition 3. *The prize that minimizes the worst-case loss relative to first best is given by*

$$\Pi^* = \sqrt{cV}. \tag{10}$$

Intuitively, the optimal prize is larger than the firm's costs c to reflect uncertainty about eventual success; the amount of this "overpayment" is increasing in the ratio of social value V to firm cost c. Precisely, society overpays relative to costs by a factor of $\sqrt{\frac{V}{c}}$ to balance the harm from potentially overpaying against the harm from underinvestment.

For example, if the social value of the realized invention is 100 times its cost (i.e., $V = 100c$), then the worst-case optimal prize is ten times cost (i.e., $\Pi = 10c$). This prize overpays by $\Pi - c = 9c$ if successful commercialization was guaranteed ($p_{\mathsf{comm}} = 1$), and it fails to elicit the invention if $p_{\mathsf{comm}} < 0.10$, for a social loss as large as $p_{\mathsf{comm}}V - c = 9c$.

7.2.3 Discussion

The simple analysis above considers the case of a single representative firm. One additional consideration in the case where multiple firms could potentially invest is that a too-large prize Π might induce a socially wasteful race. This risk is especially salient in the case where V is much higher than c and where firms understand that p_{comm} is high.

An auction or "batching" mechanism might help mitigate overpayment in this instance by inducing potential innovators to compete on price (i.e., a lower prize Π) rather than speed (see Budish et al. (2015b) for a discussion). An excellent topic for future research is the design of pull mechanisms with multiple potential innovator firms and a richer informational environment than is considered in Sections 7.2.1-7.2.2 above.[85]

7.3 Push Incentives

A complementary approach focuses on the other side of the inequality—reducing the cost of commercialization (c) rather than increasing expected returns. Most directly, either a public or private sector funder could directly fund the costs c of research, with the condition that any resulting inventions immediately be placed in the public domain. This is often called "push" financing, and includes grants, contracts, and advance procurement agreements that pay for research and development upfront.[86] Kremer et al. (2020) argue that push financing is especially attractive in settings where missing inventions are "technologically close," in the sense that the potential invention is sufficiently close to the point of commercialization that the funder can specify, with precision, the investment of interest. For example, if the social planner can specify that it is interested in a clinical trial testing metformin as a treatment for breast cancer, on the grounds that the likely return is high (i.e., $p_{\mathsf{comm}}V \gg c$), then it can directly fund the costs c of conducting the appropriate clinical trial.

The core challenge is informational: can the funder properly determine c, identify an appropriate party to receive c and invest in research, and monitor the party's effort? There is a clear analogy between push financing and other forms of public procurement, the challenges associated with which have been well studied (see, e.g., Laffont and Tirole (1993); Athey and Levin (2001); Lewis and Bajari (2011)). In the context of investments in new uses, for example, the government must decide whether to fund a researcher (e.g., a university scientist or firm), contract with an intermediary (e.g., a clinical research organization), or run a procurement auction eliciting bids to conduct such a study from qualified organizations.

Thus, where a social planner knows that the social return on investment is high (in the sense of $p_{\mathsf{comm}}V \gg c$) and has the capacity to ensure that the public resources c are spent effectively, push mechanisms can directly induce investment in socially valuable research. Pull mechanisms, by contrast, require less information on the funder's part but introduces the risk that the funder will

[85] We are grateful to Chris Snyder for conversations about pull mechanisms for new uses for existing drugs. For details on ongoing efforts to design such a mechanism in practice, see https://marketshaping.uchicago.edu/winner/repurposing-generic-drugs/.

[86] See, for example, Ahuja et al. (2021) and Castillo et al. (2021), on the design of push financing mechanisms in the context of COVID-19 vaccines.

either overpay for socially valuable investment by setting a prize that is too large or fail to induce investment by setting a prize that is too small. Developing a more thorough understanding of the optimal use of both pull and push mechanisms is another excellent topic for future research.

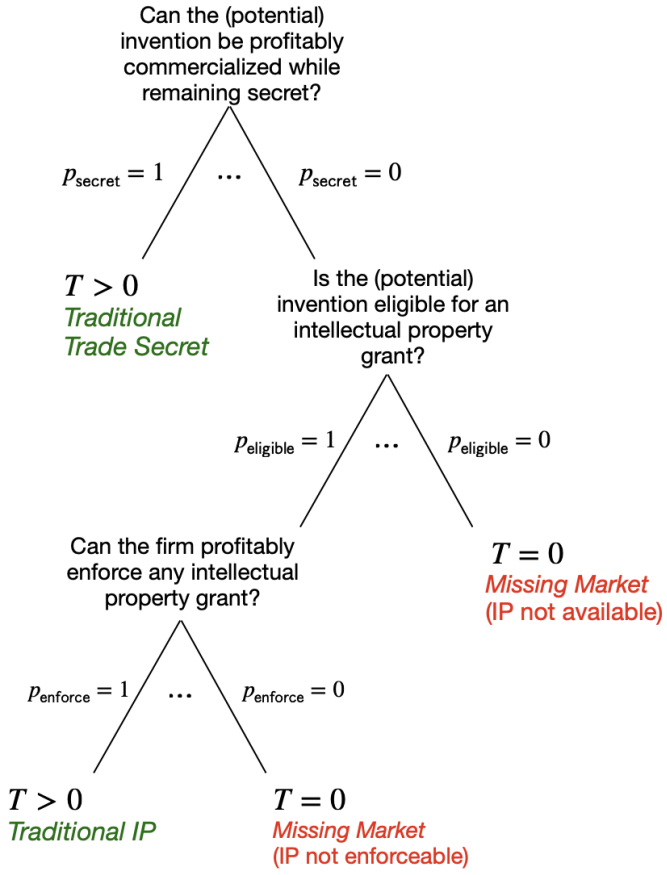
8 Conclusion

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 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 sets of ideas. 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 the number of additional new uses that would have been commercialized had intellectual property rights been fully enforceable. That is, we quantify, roughly, 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 $2.5 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 very large.

Figures and Tables

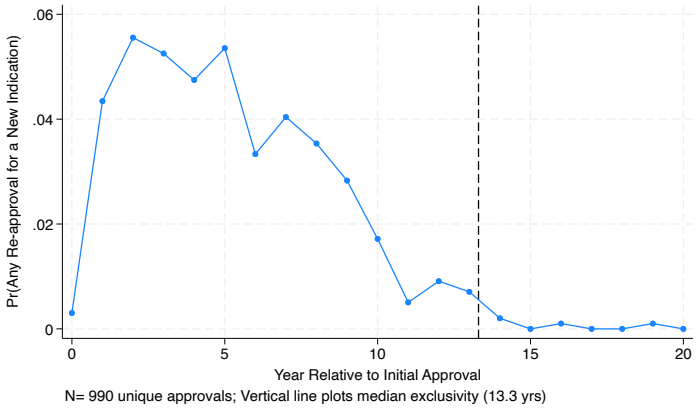
Figure 1: Missing Markets for Innovation



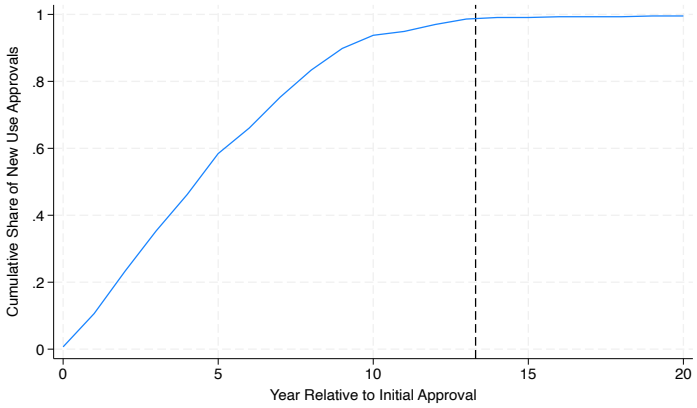
Notes: Figure summarizes the framework introduced in Section 2.1.

Figure 2: New Use Approvals Relative to Initial Approval

A. Likelihood of New Use Approvals



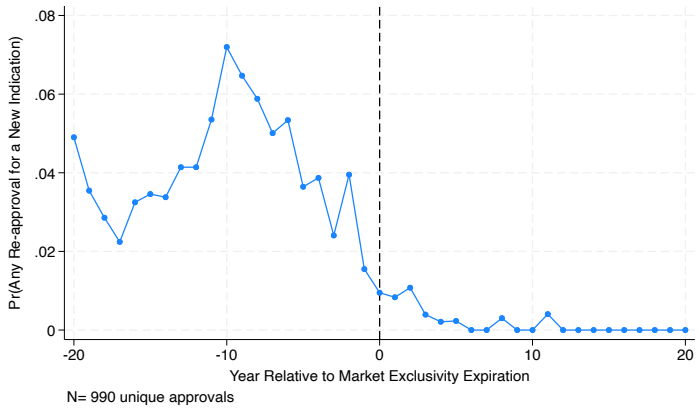
B. Cumulative Share of New Use Approvals



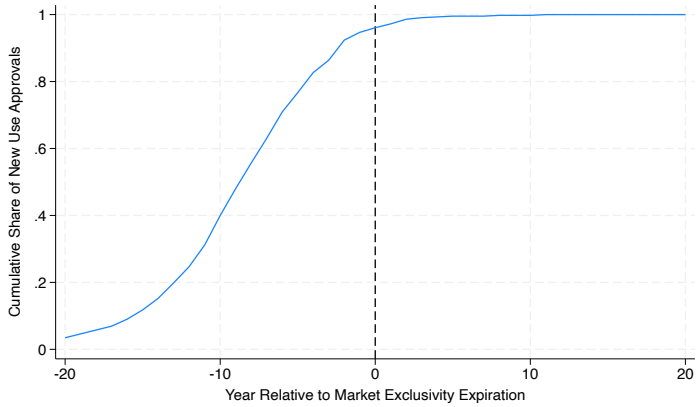
Notes: Panel A plots the probability of re-approval for a new use for the sample of innovator drugs defined in Section 4.5 ($N = 990$). Panel B plots the cumulative share of new use approvals by each year for the same sample. Both panels plot trends relative to the drug's initial approval by the FDA. Vertical lines plot median market exclusivity duration ($t = 13.3$ years).

Figure 3: New Use Approvals Relative to Exclusivity Expiration

A. Likelihood of New Use Approvals



B. Cumulative Share of New Use Approvals

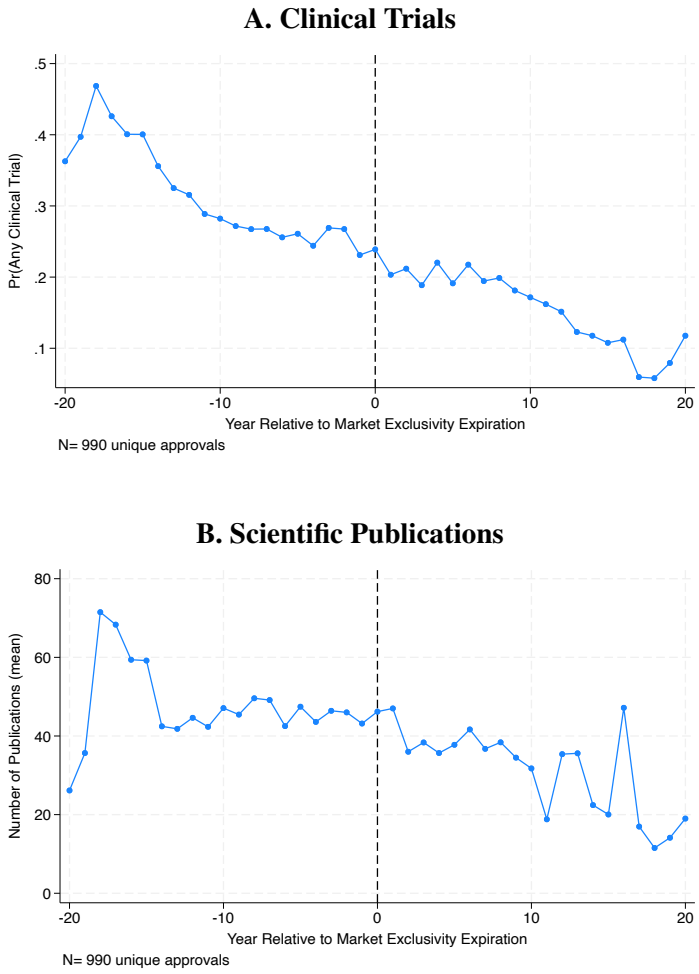


Notes: Panel A plots the probability of re-approval for a new use for the sample of innovator drugs defined in Section 4.5 ($N = 990$). Panel B plots the cumulative share of new use approvals by each year for the same sample. Both panels plot trends relative to market exclusivity expiration. Market exclusivity expiration, marked with vertical lines in both panels, 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 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 ($N = 124$), meaning that we observe data for each drug in each of the twenty years plotted. Panel B is a five-year balanced panel ($N = 415$), meaning that we observe data for each drug in each of the ten years plotted. 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.

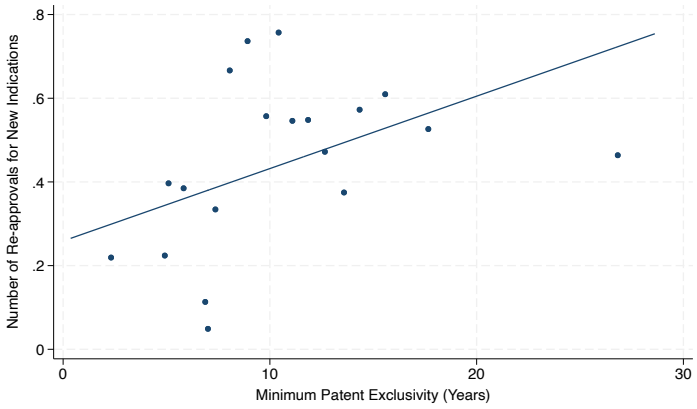
Figure 5: Other Research Investments Relative to Exclusivity Expiration



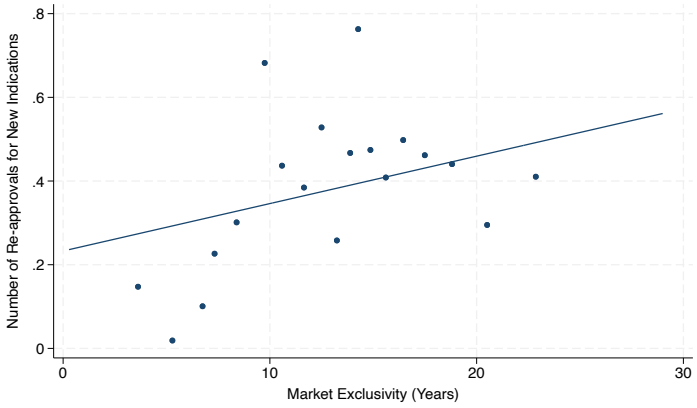
Notes: Panel A plots the probability of any active clinical trial in each relative year for the sample of innovator drugs defined in Section 4.5 ($N = 990$). We plot the probability of any active clinical trial in each of the twenty years preceding and following exclusivity expiration for all drugs in 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 Figure A6 plots trends in clinical trials and publications relative to initial approval.

Figure 6: Quantity of New Uses by Exclusivity Duration

A. Minimum Patent Exclusivity

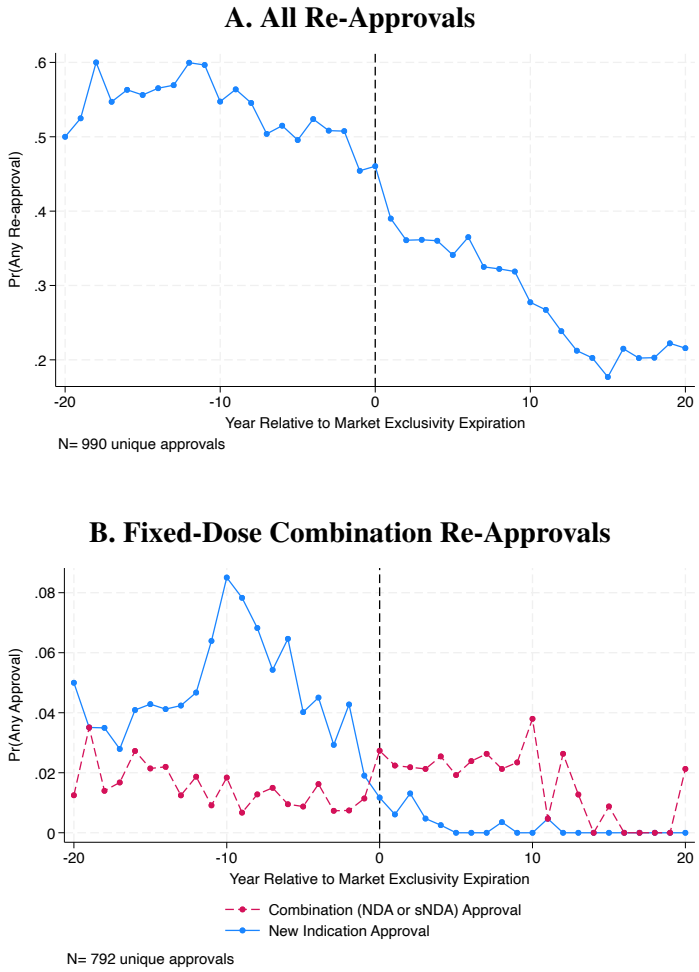


B. Market Exclusivity



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 expiration of the earliest expiring patent and all regulatory exclusivities. Panel B 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.

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. 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 (total)	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)
Minimum Patent Exclusivity	0.037	0.023		
(years)	(0.008)	(0.008)		
Market Exclusivity	—	—	0.013	0.020
(years)			(0.006)	(0.006)
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*			
Minimum Patent Exclusivity	0.083	0.059	—	—
(years)	(0.015)	(0.018)		
Market Exclusivity	—	—	0.026	0.042
(years)			(0.012)	(0.014)
Approval Year FEs	no	yes	no	yes
Observations	990	990	990	990
Intensive Margin	*# re-approvals for new/modified indication*			
Minimum Patent Exclusivity	0.017	0.011	—	—
(years)	(0.021)	(0.022)		
Market Exclusivity	—	—	0.015	0.020
(years)			(0.017)	(0.018)
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 (6) in the text. Observations are the sample of innovator drugs defined in Section 4.5. Robust standard errors reported in parentheses. 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.

Ahuja, Amrita, Susan Athey, Arthur Baker, Eric Budish, Juan Camilo Castillo, Rachel Glennerster, Scott Duke Kominers, Michael Kremer, Jean Lee, Canice Prendergast et al., "Preparing for a pandemic: Accelerating vaccine availability," in "AEA Papers and Proceedings," Vol. 111 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2021, pp. 331–335.

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.

Athey, Susan and Jonathan Levin, "Information and competition in US forest service timber auctions," *Journal of Political economy*, 2001, *109* (2), 375–417.

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.

Bryan, Kevin A and Heidi L Williams, "Innovation: market failures and public policies," in "Handbook of industrial organization," Vol. 5, Elsevier, 2021, pp. 281–388.

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 Association Papers and Proceedings*, 2016, *106* (5), 183–187.

_ , **Peter Cramton, and John Shim**, "The high-frequency trading arms race: Frequent batch auctions as a market design response," *The Quarterly Journal of Economics*, 2015, *130* (4), 1547–1621.

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.

Castillo, Juan Camilo, Amrita Ahuja, Susan Athey, Arthur Baker, Eric Budish, Tasneem Chipty, Rachel Glennerster, Scott Duke Kominers, Michael Kremer, Greg Larson et al., "Market design to accelerate COVID-19 vaccine supply," *Science*, 2021, *371* (6534), 1107–1109.

Citeline, "Pharmaprojects," Database 1995-2010. Proprietary data; version current through May 2011.

Clarivate, "Cortellis Clinical Trials Intelligence," Database 1995-2016. Proprietary data; accessed 2016.

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.

Conti, Rena M, Susan Athey, Richard G Frank, and Jonathan Gruber, "Generic drug repurposing for public health and national security: COVID-19 and beyond," *Health Affairs Forefront*, 2020.

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.

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.

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.

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.

F-D-C Reports, "NDA Pipeline," Database 1982-2001. Proprietary data; accessed April 2007.

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.*

Garraway, Levi A. and Eric S. Lander, "Lessons from the Cancer Genome," *Cell*, 2013, *153* (1), 17–37.

Gault, Matthew, "Stanford Scientists Reverse Engineer Moderna Vaccine, Post Code on Github," *VICE*, mar 2021. 9:35 a.m. (PDT).

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.

Goldwasser, Shafi, Silvio Micali, and Chales Rackoff, "The knowledge complexity of interactive proof-systems," in "Providing sound foundations for cryptography: On the work of shafi goldwasser and silvio micali" 2019, pp. 203–225.

Hall, Robert E and Charles I Jones, "The value of life and the rise in health spending," *The Quarterly Journal of Economics*, 2007, *122* (1), 39–72.

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

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

_ and Jeannie Suk, "The law, culture, and economics of fashion," *Stan. L. Rev.*, 2008, *61*, 1147.

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.

Hoadley, Katherine A. et al., "Multiplatform Analysis of 12 Cancer Types Reveals Molecular Classification within and across Tissues of Origin," *Cell*, 2014, *158* (4), 929–944.

_ and _, "Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer," *Cell*, 2018, *173* (2), 291–304.

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.

Laffont, Jean-Jacques and Jean Tirole, *A theory of incentives in procurement and regulation*, MIT press, 1993.

Léauté-Labrèze, Christine, Peter Hoeger, Josiane Mazereeuw-Hautier, Laurent Guibaud, Eva Baselga, Giedre Posiunas, Robert J. Phillips, Horacio Caceres, Juan Carlos Lopez Gutierrez, Rafael Ballona et al., "A Randomized, Controlled Trial of Oral Propranolol in Infantile Hemangioma," *New England Journal of Medicine*, 2015, *372* (8), 735–746.

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.

Lewis, Gregory and Patrick Bajari, "Procurement contracting with time incentives: Theory and evidence," *The Quarterly Journal of Economics*, 2011, *126* (3), 1173–1211.

Lichtman, Douglas, "How the Law Responds to Self-Help," *JL Econ. & Pol'y*, 2005, *1*, 215.

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.

Manson, JoAnn E., Judith Hsia, Karen C. Johnson, Jacques E. Rossouw, A. Richey Assaf, Norman L. Lasser, Maurizio Trevisan, Howard R. Black, Susan R. Heckbert, Robert Detrano et al., "Estrogen plus Progestin and the Risk of Coronary Heart Disease," *New England Journal of Medicine*, 2003, *349* (6), 523–534.

Masur, Jonathan S and Lisa Larrimore Ouellette, *Patent Law: Cases, Problems, and Materials (2025)* 2025.

Michaeli, Daniel Tobias, Mackenzie Mills, and Panos Kanavos, "Value and price of multi-indication cancer drugs in the USA, Germany, France, England, Canada, Australia, and Scotland," *Applied health economics and health policy*, 2022, *20* (5), 757–768.

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.

Murphy, Kevin and Robert Topel, "The value of health and longevity," *Journal of Political Economy*, 2006, *114* (5), 871–904.

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

National Library of Medicine, "PubMed," Database 1966-2022. Accessed February 2026.

_ , "ClinicalTrials.gov," Database 2010-2016. Accessed July 28, 2022.

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.

O'Connor, Lauren, Maeve Bailey-Whyte, Manami Bhattacharya, Gisela Butera, Kaitlyn N Lewis Hardell, Andrew B Seidenberg, Philip E Castle, and Holli A Loomans-Kropp, "Association of metformin use and cancer incidence: a systematic review and meta-analysis," *JNCI: Journal of the National Cancer Institute*, 2024, *116* (4), 518–529.

Pahud, Dominique, Lesa Mitchell, John Wilbanks, Melissa Stevens, Scott Weir, John McBride, Lili Portilla, Ed Pezalla, Michael A Nameth, Bernadette O'Donoghue, and others, "A New Market Access Path for Repurposed Drugs," *Available at SSRN 2437502*, 2014.

Patrick, Graham, *An Introduction to Medicinal Chemistry*, 2 ed., Oxford University Press Oxford, 2005.

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, "Unpatentable drugs and the standards of patentability," *Texas Law Review*, 2009, *87* (3), 503–570.

_ , "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.

Sachs, Rachel E, Paul B Ginsburg, and Dana P Goldman, "Encouraging new uses for old drugs," *Jama*, 2017, *318* (24), 2421–2422.

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.

Sampat, Bhaven and Heidi L Williams, "How do patents affect follow-on innovation? Evidence from the human genome," *American Economic Review*, 2019, *109* (1), 203–236.

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.

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.

Sherkow, Jacob S and Paul R Gugliuzza, "Infringement by Drug Label," *Stanford Law Review*, 2026, *78*.

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.

Snyder, Christopher M, Kendall Hoyt, Dimitrios Gouglas, Thomas Johnston, and James Robinson, "Designing Pull Funding For A COVID-19 Vaccine: Article proposes a "pull" program that incentivizes late-stage development and manufacturing of COVID-19 vaccines by awarding advance purchase commitments to bidding firms.," *Health Affairs*, 2020, *39* (9), 1633–1642.

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.

Thomas, Gareth, *Medicinal Chemistry: An Introduction*, Wiley-Interscience, 2007.

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*.

U.S. Food and Drug Administration, "Approved Drug Products with Therapeutic Equivalence Evaluations (Orange Book)," Database 1984-2016. Accessed February 2026.

_ , "Drugs@FDA: FDA-Approved Drugs," Database 1985-2014. Accessed February 2026.

_ , "FDALabel: Full-Text Drug Product Labeling," Database 2005-2016. Accessed February 2026.

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

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.

Williams, Heidi L, "Intellectual property rights and innovation: Evidence from the human genome," *Journal of Political Economy*, 2013, *121* (1), 1–27.

_ , "How do patents affect research investments?," *Annual Review of Economics*, 2017, *9*, 441–469.

Wood, R. A. et al., "Omalizumab for the Treatment of Multiple Food Allergies," *New England Journal of Medicine*, 2024. Phase III OUtMATCH trial showing superiority over placebo in reducing allergic reactions.

World Health Organization, "International Classification of Diseases, Tenth Revision (ICD-10)," Classification system 2019. Accessed February 2026.

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.[87] 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.[88] 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

[87] We have been unable to obtain analogous records from the FDA covering more recent years.
[88] 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 tablets, 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, the date on 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:[89]

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, our data includes both *acetazolamide* and its salt form *acetazolamide sodium* as active ingredients. 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.[90] 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.

[89] 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.

[90] 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.[91] We construct such a crosswalk, using the Drugs@FDA Product file as our input data.[92]

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

[91] 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. INNs, however, may not always map onto drugs names as used in the United States. For example, "paracetamol" (commonly used in countries where British English is used) is the INN for "acetaminophen," the drug often marketed under the brand name Tylenol.

[92] 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, for this particular exercise. All other analyses use an older version of Drugs@FDA (Drugs@FDA 1.0) that contains information on product re-approvals central to our analysis.

BLAs.[93] We drop 77 instances in which active moiety groups have no products that were approved under NDAs.[94] 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.[95]

We identify **1,102** unique combinations of active moieties that meet our criteria for "innovator" drugs. Our analyses restrict consideration to the **990** drugs with at least one patent or regulatory exclusivity.

A.1.3 Linking "innovator" drugs to generic competitors

We designate any drug that is approved under an ANDA a "generic."[96] 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 occurs for 186 unique products—we designate the drug with the lowest application number as the earliest generic entrant. This selection procedure has no impact on our analysis, which does not depend on the identity of the first generic entrant.

A.1.4 Counting approvals and re-approvals

We construct two measures of drug approvals and re-approvals. 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. Our primary analyses record re-approvals associated with an innovator drug (i.e., regulatory actions under the same new drug application associated with the innovator approval). In robustness checks, we instead collect re-approvals associated with

[93] 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.

[94] 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).

[95] 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."

[96] 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

all NDAs that include the same active moiety. Concretely, consider the case of ranitidine, most commonly marketed under the brand name Zantac. Ranitidine was first approved by the FDA in 1983 as a treatment for peptic ulcers. It was subsequently re-approved under 11 additional NDAs, with distinct brand names. Our primary analysis focuses on re-approvals for new indications associated with only the first NDA. In robustness analyses, we consider re-approvals associated with any ranitidine NDA. Note that ranitidine is an outlier in terms of total number of associated NDAs: the median innovator drug in our data is approved under just one NDA and only 43 drugs are approved under more than five NDAs.

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 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.[97] 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 percent).

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

[97] 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 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	received in 2016	trialid_singledata.dta trialid_drug_primary_intervention.dta	1995–2015	yes
ClinicalTrials.gov	downloaded from AACT; version dated 20220728	studies.txt; browse_interventions.txt; interventions.txt; interventions_other_names.txt	2010–2016	no

from a public agency. We draw on information from the "grants" table of PubMed and designate a publication as having 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.

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.[98] 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.

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

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.[99]

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 "preclinical."

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

[99] 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.

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.[100]

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 "preclinical." 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.

[100] 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.

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.[101] We collect 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

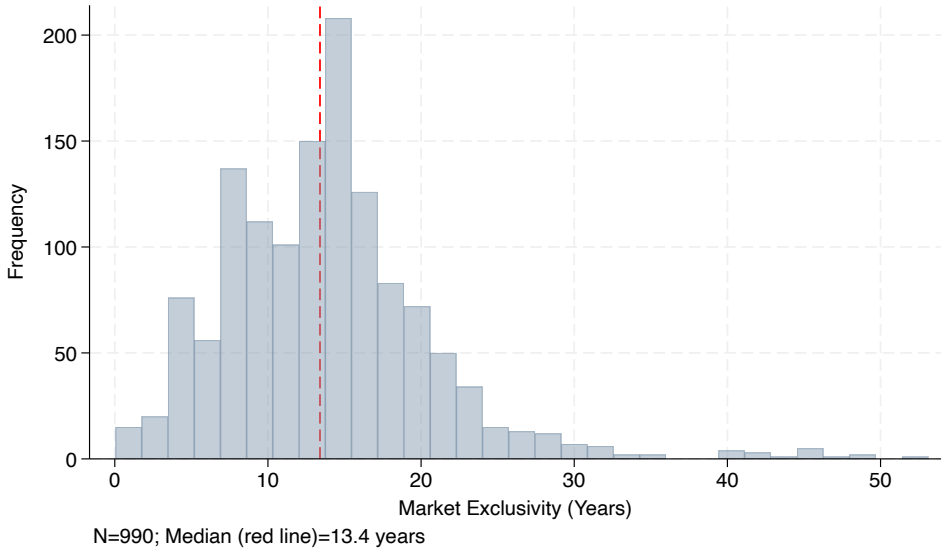
[101] In 2005, the International Committee of Medical Journal Editors (ICMJE) began requiring trial registration as a condition of publication. In 2007 the requirements for submission to ClinicalTrials.gov were expanded after Congress passed the Food and Drug Administration Amendments Act of 2007 (FDAAA) (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.

labeling service, FDALabel.[102] We supplement data on labels—which are primarily available from 2005 forward—with data from drug approval packets[103], 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 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.[104] A second researcher verified the accuracy of a random 20 percent sample of codes. We report the most granular ICD-10 code available, 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



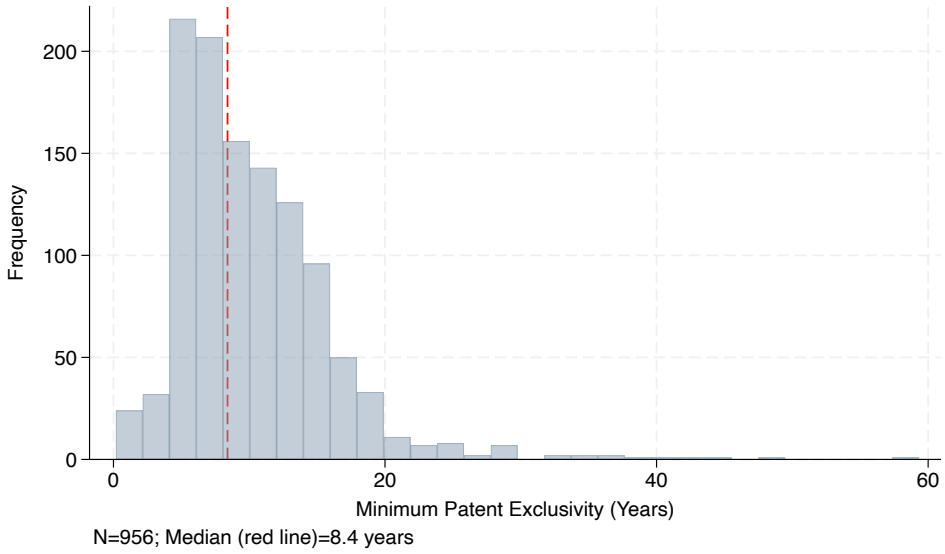
Notes: Figure is a histogram that plots the distribution of the market exclusivity measure defined in the text as the number of years between an innovator drug's initial approval and the earlier of the expiration of all patents and regulatory exclusivities and generic entry. The distribution of this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

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

[103] 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

[104] 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

B.1 Empirical signatures

Figures 3 and 4 plot trends in the probability that an innovator drug is re-approved for a new use around market exclusivity expiration using two different samples. In Figure 3, we include each of the 990 innovator drugs in our primary sample for every year in which data are available. This means that the sample of drugs shifts across points in Figure 3. This changing sample composition generates the non-monotonic trend between $t = -20$ and $t = -15$. Consider, for instance, a drug approved in the year 2000 that receives ten years of market exclusivity, and recall that our data include observations through 2014. In Figure 3, this drug appears in the sample plotted at each point between $t = -10$ and $t = 4$. This same drug will not appear in Figure 4, as the balanced panel specifications require, respectively, ten and twenty years of data both before and after market exclusivity expiration. In this Section, we introduce two additional ways of presenting trends in the probability of re-approval, which confirm that the three empirical signatures introduced in Section 3.3 are robust to alternative cuts of our data.

First, we modify the exercise in Figure 3 to address the fact that drugs in our sample receive varying durations of market exclusivity. For each drug, we normalize market exclusivity to one

and consider trends in the likelihood of re-approval for a new use in each of five periods prior to exclusivity expiration. Under this scaling, the date of initial approval is assigned $t = -1$ and the date of market exclusivity expiration occurs at time $t = 0$.[105] We divide the period of market exclusivity into five equal intervals: $t = [-1.0, -0.8)$, $t = [-0.8, -0.6)$, $t = [-0.6, -0.4)$, $t = [-0.4, -0.2)$, and $t = [-0.2, 0)$. To facilitate comparison with Figure 3 and the three empirical signatures introduced in Section 3.3, we include an additional time period after exclusivity expiration that, for each drug, has the same duration as one of these bins (i.e., $[0, 0.2)$). For example, for a drug with twenty years of market exclusivity, each period includes four years of data; for a drug with five years of market exclusivity, each period includes one year. We thus compare trends in the likelihood of re-approval across relative phases of a drug's lifecycle rather than in calendar time.

Figure A3 plots the probability that a drug is re-approved for a new use in each of these six periods. The probability that drug is re-approved for a new use is highest in the second period (the first period that does not include initial approval), at which time nearly seven percent of drugs are re-approved at least once. The probability of re-approval declines monotonically in each subsequent period and falls below one percent in the first period after exclusivity expiration. These trends are consistent with the three empirical signatures of a missing market for new uses introduced in Section 3.3.

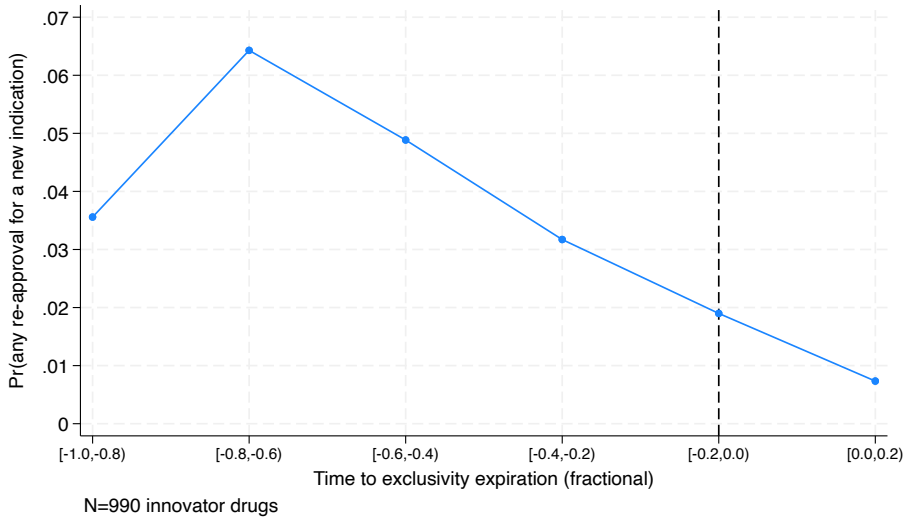
Second, we inspect trends in the probability of re-approval for a new use within cohorts of drugs that receive the same number of years of market exclusivity. Specifically, for all drugs that receive x years of market exclusivity, we plot all data prior to exclusivity expiration and include five years of data following expiration. That is, for each cohort of drugs with x years of exclusivity, we plot $t = -x$ through $t = 5$. As the median drug in our sample receives 13 years of market exclusivity and the modal drug receives 14 years, we separately consider $x = \{12, 13, 14, 15\}$. There are 59 drugs with $x = 12$, 70 drugs with $x = 13$, 120 drugs with $x = 14$, and 85 drugs with $x = 15$.

Although each of these cohorts includes at least 50 innovator drugs, it is worth emphasizing a tradeoff inherent in this style of sample restriction. We gain, with this approach, the ability to compare investments in new uses over time for drugs with identical lifecycle durations. However, we lose more than ninety percent of our sample with each of these cuts. Given that many drugs are never approved for a new use—the rationale behind the two-part model in Section 5.2—these data are likely to be quite noisy.

Nonetheless, the trends in Figure A4 for drugs with exactly $x = \{12, 13, 14, 15\}$ years of market exclusivity remain consistent with the three empirical signatures of a missing market. For each cohort, the probability that a drug is re-approved for a new use increases sharply in the years following initial approval (marked with a dotted grey line), peaks soon thereafter, then declines. In each case, the probability of re-approval for a new use falls to zero within two years of exclusivity expiration. Unsurprisingly, given the comparatively small samples, these trends are noisier than those in Figures 2, 3, and 4, but each is consistent with the idea that incentives for investment are high but declining

[105] Note that our data record the precise dates of initial approval, market exclusivity expiration, and re-approval, so we are able to construct these periods precisely.

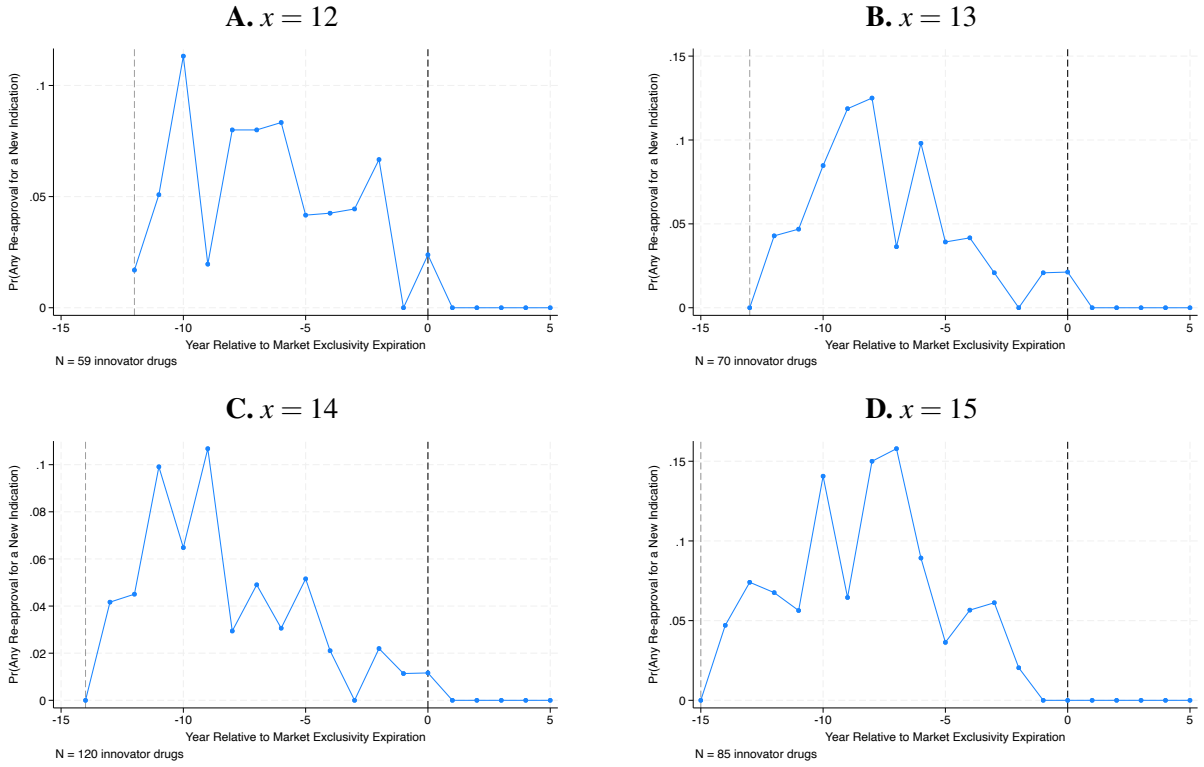
Figure A3: Fractional Exclusivity Periods



Notes: Figure plots the probability of re-approval for a new use for the sample of innovator drugs defined in Section 4.5 in each of six periods relative to market exclusivity expiration ($N = 990$). 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 split each drug's period of market exclusivity into five periods of equal length and plot the probability of re-approval in each of these periods; we consider trends in a sixth period following exclusivity expiration, set to the same length.

prior to exclusivity expiration and zero afterward.

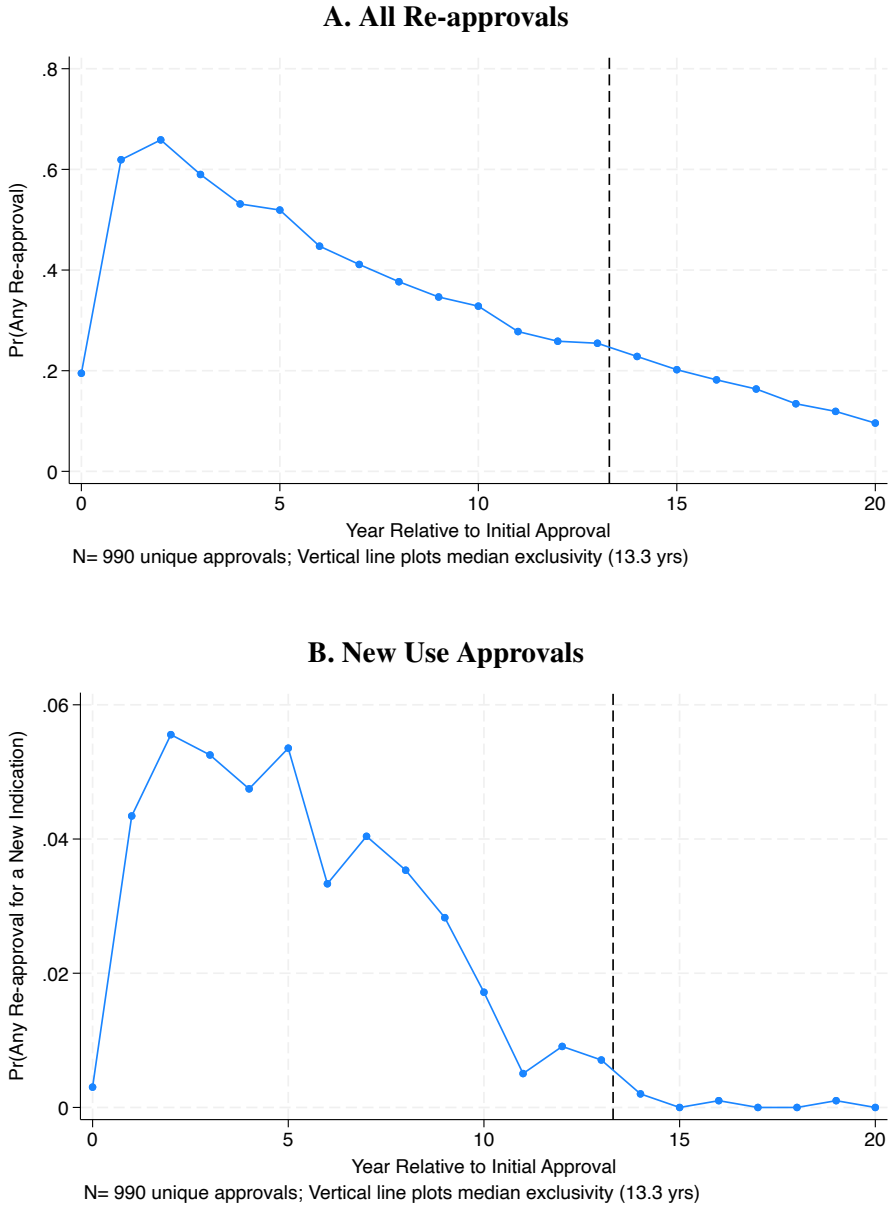
Figure A4: Exclusivity Cohorts



Notes: Panels plot the probability of re-approval for a new use for samples of innovator drugs with exactly x years of market exclusivity, where $x = \{12, 13, 14, 15\}$. Trends are relative to market exclusivity expiration, 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 select $x = 12, 13, 14, 15$, which capture the period around median market exclusivity (13.3 years).

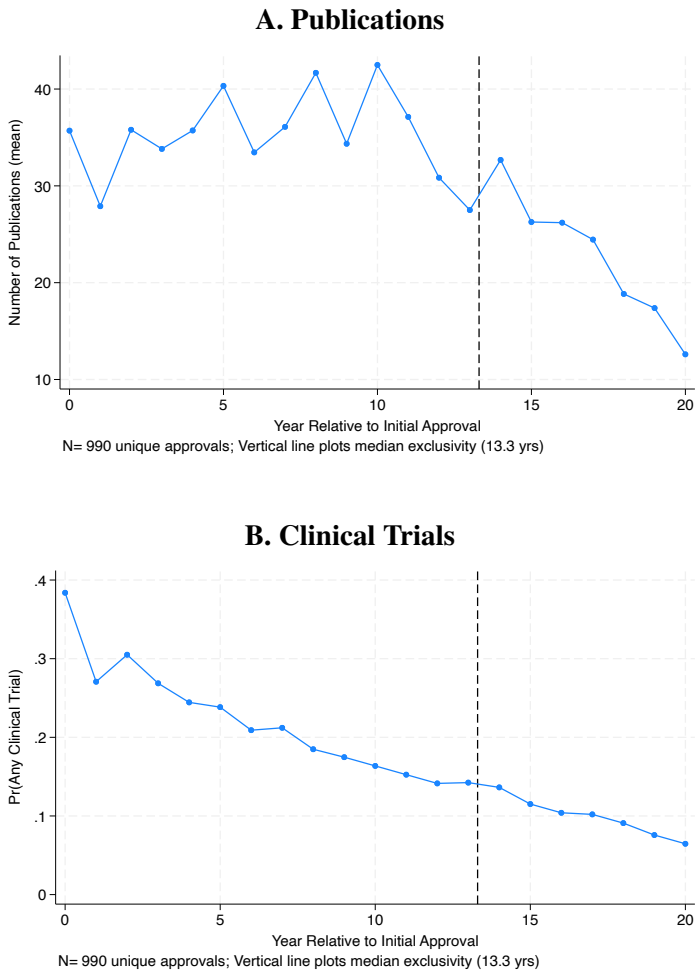
B.2 Year relative to initial approval

Figure A5: FDA Approvals Relative to Initial Approval



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 ($N = 990$).

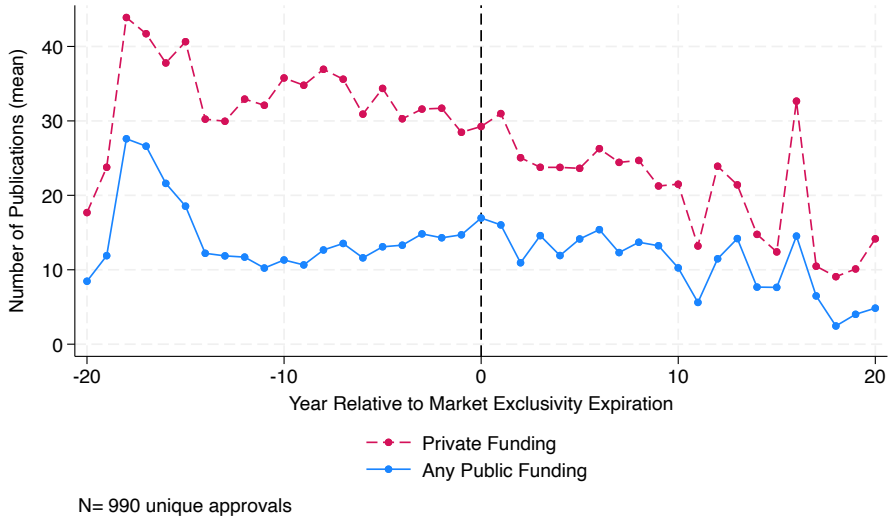
Figure A6: Research Investments Relative to Initial Approval



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.

B.3 Scientific Publications

Figure A7: Scientific Publications Relative to Exclusivity Expiration

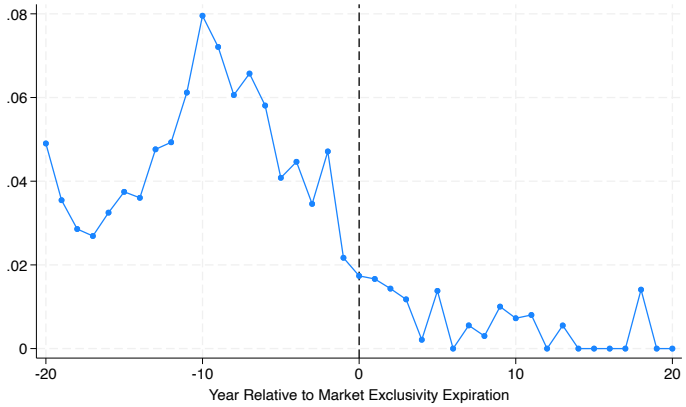


Notes: Figure plots the average number of scientific publications published in each relative year for the sample of innovator drugs defined in Section 4.5 ($N = 990$). Publication counts are plotted separately for publications that acknowledge receiving any public research support (blue solid line) and publications that do not (red 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.

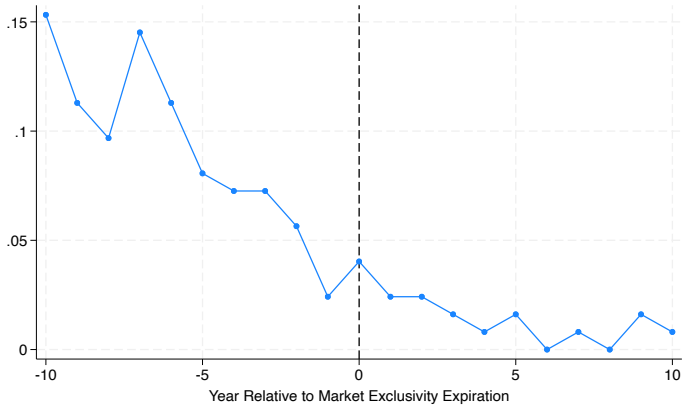
B.4 New use approvals associated with any NDA

Figure A8: Probability of Re-Approval (All NDAs)

A. Pooled Sample



B. Balanced Panel (10 years)



Notes: Panel A plots the probability of re-approval for a new use for all active moieties associated with the sample of innovator drugs defined in Section 4.5 ($N = 990$). Panel B is a ten-year balanced panel ($N = 124$). Trends are relative to market exclusivity expiration for the innovator NDA, 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.

B.5 Proof of Proposition 3

As shown in the main text, the worst case loss from setting Π too low is $p_{\text{comm}}V - c = \frac{c}{\Pi}V - c$, while the worst case loss from setting Π too high is $(V-c)-(V-\Pi) = \Pi - c$. Hence we seek to choose Π to solve

$$\min_{\Pi} \max \left[\frac{c}{\Pi}V - c, \Pi - c \right] \tag{11}$$

Observe that the worst case loss from underpaying is strictly decreasing in Π (and is infinite at $\Pi = 0$) while the worst case loss from overpaying is strictly increasing in Π (and is negative at $\Pi = 0$). Additionally, both worst cases are continuous in Π. Therefore, the solution to (11) equates

the two worst case losses:

$$\frac{c}{\Pi^*}V - c = \Pi^* - c.$$

Algebra then yields

$$\Pi^* = \sqrt{cV},$$

as claimed. □