NBER WORKING PAPER SERIES

STARVING (OR FATTENING) THE GOLDEN GOOSE?: GENERIC ENTRY AND THE INCENTIVES FOR EARLY-STAGE PHARMACEUTICAL INNOVATION

Lee Branstetter
Chirantan Chatterjee
Matthew J. Higgins

Working Paper 20532
http://www.nber.org/papers/w20532

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2014

We thank Tamer Abdelgawad, Iain Cockburn, Darren Filson, Carolin Haeussler, Bart Hamilton, Dietmar Harhoff, Sherry Knowles, Margaret Kyle, Jeff Macher, Joseph Mahoney, Alex Oettl, Ivan Png, Jerry Thursby, and Brian Wright as well as seminar participants at the University of Illinois, University of Passau, University of California Berkeley, Carnegie Mellon University, Ludwig Maxmilians University Munich, National University of Singapore, India Statistical Institute (Delhi), Georgia Institute of Technology and conference participants at the USPTO Conference on Patents, Entrepreneurship and Innovation (Washington, DC) and the Asia Pacific Innovation Conference (APIC), for valuable comments and discussions. Programming and research assistance by Jeremy Watson, Winston Yang and Suvojyoty Sahais is gratefully acknowledged. We also thank IMS Health Incorporated for their generous support and access to their data. The statements, findings, conclusions, views, and opinions contained and expressed herein are not necessarily those of IMS Health Incorporated or any of its affiliated or subsidiary entities. The statements, findings, conclusions, views, and opinions contained and expressed in this article are based in part on data obtained under license from the following IMS Health Incorporated or affiliate information service(s): IMS Midas™, IMS Lifecycle™, IMS National Disease and Therapeutic Index™, IMS National Prescription Audit™, IMS Health Incorporated or its affiliates. Higgins acknowledges funding from The Imlay Professorship. Chatterjee acknowledges IIM Bangalore for supporting his extended research visit to Georgia Tech. Higgins and Branstetter acknowledge funding from NSF SCISIP Grants #1064122 and #1360057. Authors are listed alphabetically and the usual disclaimers apply. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Starving (or Fattening) the Golden Goose?: Generic Entry and the Incentives for Early-Stage Pharmaceutical Innovation

Lee Branstetter, Chirantan Chatterjee, and Matthew J. Higgins

NBER Working Paper No. 20532
September 2014
JEL No. D2,L5,L51,L65,M2

ABSTRACT

Over the last decade, generic penetration in the U.S. pharmaceutical market has increased substantially, providing significant gains in consumer surplus. What impact has this rise in generic penetration had on the rate and direction of early stage pharmaceutical innovation? We explore this question using novel data sources and an empirical framework that models the flow of early-stage pharmaceutical innovations as a function of generic penetration, scientific opportunity, firm innovative capability, and additional controls. While the aggregate level of early-stage drug development activity has increased, our estimates suggest a sizable, robust, negative relationship between generic penetration and early-stage pharmaceutical research activity within therapeutic markets. A 10% increase in generic penetration is associated with a 7.9% decline in all early-stage innovations in the same therapeutic market. When we restrict our sample to first-in-class pharmaceutical innovations, we find that a 10% increase in generic penetration is associated with a 4.6% decline in early-stage innovations in the same market. Our estimated effects appear to vary across therapeutic classes in sensible ways, reflecting the differing degrees of substitution between generics and branded drugs in treating different diseases. Finally, we are able to document that with increasing generic penetration, firms in our sample are shifting their R&D activity to more biologic-based (large-molecule) products rather than chemical-based (small-molecule) products. We conclude by discussing the potential implications of our results for long-run welfare, policy, and innovation.

Lee Branstetter
Heinz College and
Department of Social and Decision Sciences
Carnegie Mellon University
Pittsburgh, PA 15213
and the Peterson Institute for
International Economics
and also NBER
branstet@andrew.cmu.edu

Chirantan Chatterjee
Department of Corporate Strategy & Policy
Indian Institute of Management Bangalore
chirantan.chatterjee@IIMB.ERNET.IN

Matthew J. Higgins
Scheller College of Business
Georgia Institute of Technology
800 West Peachtree Street
Atlanta, GA 30308
and NBER
matt.higgins@scheller.gatech.edu

1 Introduction

In his provocative paper, "The Health of Nations," Yale University economist William Nordhaus (1999) argues that the advances in human welfare generated by better medical science over the past half century have been equal in value to the consumption increases from all other sources put together. Victor Fuchs (1982) has suggested that most of the real improvement in human health generated over this period stems from modern medicine's expanding arsenal of pharmaceutical products. While documenting these claims in a way that meets modern evidentiary standards is challenging, the work of scholars such as Frank Lichtenberg (2001, 2004, 2007) has provided evidence suggesting that the gains from pharmaceutical innovation have been very large. In the long run, global investments in pharmaceutical research have proven to be very good ones.

These benefits have come with significant costs; pharmaceutical innovation is risky and expensive. These costs are passed on to consumers in the form of higher prices for branded pharmaceuticals. In recent years, prescription drug spending in the U.S. has exceeded $300 billion, an increase of $135 billion since 2001. Consumption of prescription drugs now accounts for approximately 12 percent of total health care spending (GAO, 2012). However, over this time period, generic products have accounted for an increasing share of prescription drug expenditures, saving consumers an estimated $1 trillion (GAO, 2012). Current regulation attempts to strike a balance between access to lower cost generics on the one hand and adequate incentives to promote pharmaceutical innovation on the other. While the rise in generic penetration has brought substantial benefits to consumers (Branstetter *et al*., 2013), some have argued that the regulatory "balance" has shifted so far in the direction of access to inexpensive drugs that it has undermined the incentives for new drug development (Higgins and Graham, 2009; Knowles, 2010). Such a shift could have strong implications, even for drug companies outside the United States, because the global industry relies disproportionately on the U.S. market as a source of its profits. Has the increase in generic entry affected pharmaceutical innovation? Our study attempts to address this question and quantify, for the first time, the impact of generic entry on early-stage drug development.

We start by constructing a novel and unique dataset that allows us to analyze this issue at the level of narrowly defined therapeutic areas. Instead of relying on patents as measures of innovation, we focus on early-stage drug development. While patenting is certainly important in the pharmaceutical industry, it can occur anytime throughout the drug development process, and it often occurs long before the actual therapeutic value of a compound has been demonstrated. Our outcome variable, on the other hand, allows us to measure what is actually happening in the early stages of the clinical development

process. We also utilize comprehensive data on branded and generic drug sales across all therapeutic categories in the U.S. market[1], obtained at the firm-product-year level, such that we can measure the differential exposure of individual firms to generic competition across these different therapeutic markets. Finally, we seek to control for changes in scientific opportunity by building a comprehensive database of citation-weighted scientific journal articles in the medical sciences and mapping them to our pharmaceutical product market categories.

Using these data, we find that the *aggregate* level of new drug development has *not* declined as generic penetration in the U.S. market has risen; the total number of new compounds (including both small and large molecules) in early stage development has risen over our sample period (Figure 1). However, rising generic competition has had a statistically and economically significant impact on *how* pharmaceutical product development is undertaken and *where* those efforts are focused. We show this by using an empirical framework that models the flow of early-stage pharmaceutical innovations as a function of generic entry and penetration, as well as scientific opportunity and challenges, firm innovative capability and a vector of additional controls. Using this framework, we document a negative and significant relationship between generic entry (penetration) and early-stage innovation at the ATC 2-digit therapeutic category level. The elasticity from our specification implies that a 10% increase in generic penetration in a particular market will *lower* early-stage innovations, in that same market, by 7.9%.

The interpretation that an increase in generic penetration within a market lowers early-stage innovation is strengthened by a series of alternative specifications and robustness checks. First, we limit our sample to a set of therapeutic categories where substitution between generics and branded products is limited for clinical reasons, and we find that our measured effect attenuates to the point of insignificance, as expected. Second, we show that our estimated effect is strongly negative for early-stage innovation, where it is possible to redirect R&D in response to market shifts, but much weaker for late-stage innovation, where firms have stronger incentives to deploy products that have survived the clinical trials process, even if generic competition is limiting the addressable market. Third, we show that our baseline effect is robust to inclusion of (therapeutic market * year) interaction terms that effectively remove all the unobserved market-specific effects that change in a common way across firms.

Finally, we consider the possibility that, within therapeutic markets, a *shift* is occurring out of chemical-based (small molecule) products and into biologic-based (large molecule) products. The regulatory mechanisms that have accelerated generic entry in chemical-based drugs do not extend to biologics. Additionally, the pathways by which biologic-based generics (known in the industry as

[1] We use the phrases *therapeutic area, therapeutic market, therapeutic category* and *product markets* interchangeably in this paper. In our empirical work, they correspond to 2-digit categories within the World Health Organization's Anatomical Therapeutic Chemical (ATC) classification system.

'biosimilars') could enter the U.S. market have yet to be finalized.[2] Exploiting this regulatory difference between chemical-and biologic-based innovations, we find a positive relationship between generic entry and a *shift* towards biologic-based products within therapeutic categories. As conjectured by Golec *et al*. (2010), this movement suggests that the nature of innovation taking place in the pharmaceutical industry is changing.

Is this shift in the direction and nature of drug development socially beneficial or socially harmful? At this stage in the research process, it is not yet possible to produce a definitive answer to this question. On the one hand, one could argue that current regulation is 'pushing' innovation toward therapeutic markets for which significant numbers of viable generics do not exist. In other words, R&D efforts and expenditures could be flowing to therapeutic areas that are relatively underserved, thereby generating welfare gains. On the other hand, our evidence of a significant movement in the data from development of chemical-based to biologic-based products may have important implications for the future, especially since biologics tend to be more expensive, on average, than chemical-based products. Until current regulatory challenges are resolved, these higher prices may persist for long periods of time. As the regulatory playing field tilts sharply in the direction of biologics, and firms respond rationally to the incentives they confront, we cannot rule out the possibility that recent efforts to balance access with incentives for innovation will give us cheaper drugs today, but more expensive drugs tomorrow.

The paper proceeds as follows. Section 2 provides a discussion of the U.S. regulatory environment in which pharmaceutical firms operate and a brief description of the rise in generic penetration. Section 3 reviews important features of the drug development process and discusses prior work on the potential impact of rising generic penetration on pharmaceutical innovation. Our empirical specification and data are outlined in Section 4. Results are presented in Section 5,and we conclude in Section 6.

2 The U.S. Regulatory Environment and the Rise of Generic Penetration

The current regulatory environment faced by pharmaceutical companies in the U.S. can be traced to the passage of the Drug Price Competition and Patent Term Restoration Act in 1984, informally known as "Hatch-Waxman" Act. One of the hallmarks of this legislation is the balance it tries to strike between access by consumers to inexpensive generic drugs, on the one hand, and the protection of adequate incentives for new drug development on the other. Hatch-Waxman allows expedited Food and Drug

[2] The FDA released non-binding draft guidance in May 2014 with an open comment period until August 2014. During our current sample period there was no pathway for biosimilars to enter the market. However, even when biosimilars are finally able to enter the U.S. market, large-molecule drugs will have a much longer period of guaranteed monopoly than small-molecule drugs. This difference could affect the economic incentives for developing these different types of drugs.

Administration (FDA) approval for generic entry while extending the life of pharmaceutical patents in order to compensate innovators who lost time on their "patent clocks" waiting for FDA approval (Grabowski, 2007).

When a pharmaceutical company submits a New Drug Application (NDA) to the FDA for approval, the law requires the company to identify all relevant patented technologies necessary to create the drug; these patents are subsequently listed in the FDA Orange Book.[3] Upon approval of a drug, the FDA will extend the life of the patents protecting the drug, in order to partly compensate the pharmaceutical firm for time lost on the "patent clock" during the FDA-mandated approval process (Grabowski, 2007).[4] In addition, the FDA will also grant each new approved product regulatory protection lasting for five years ("data exclusivity") that runs concurrently with patent protection.[5] During this data exclusivity period, regardless of the status of the underlying patent(s), no generic entry may occur. At the conclusion of data exclusivity branded products are protected only by their patents; this period running from the cessation of data exclusivity to the expiration of the patent(s) is commonly referred to as "market exclusivity." Figure 2 illustrates these overlapping periods of protection for branded drugs

Prior to the passage of Hatch-Waxman, generic manufacturers seeking to sell their products in the U.S. market had to demonstrate the safety and efficacy of their products by putting them through clinical trials. While the outcome of these trials lacked the uncertainty involved in the trials of an innovative new drug, the time and expense involved were a significant disincentive for generics manufacturers to put products on the market, since they could not charge a premium price to offset the costs of clinical trials. Before Hatch-Waxman, it is estimated that more than 150 products existed without any patent protection and without any generic entry (Mossinghoff, 1999). While Hatch-Waxman did not lessen the burden of the clinical trials process for branded pharmaceutical companies seeking approval for new drugs, it essentially eliminated the requirement for separate clinical trials for generic manufacturers. All generic manufacturers had to do was demonstrate "bioequivalence" with branded products by showing that the active ingredient in their product diffused into the human bloodstream in a manner similar to the original product.

---

[3] For biologic-based or "large molecule drugs" the initial application is a Biologics License Application. However, a similar requirement to disclose patents exists, and this disclosure also becomes a matter of public record.

[4] There are limits to this. Pharmaceutical firms cannot receive a patent extension of more than five years, nor are they entitled to patent extensions that give them effective patent life (post approval) of greater than 14 years.

[5] There are exceptions to the general rule of 5 years of data exclusivity. Drugs targeting small patient populations (known in the industry as "orphan drugs") receive 7 years of data exclusivity. Reformulations of existing drugs receive only 3 years of data exclusivity. New drugs that treat pediatric illnesses receive an additional 6 months of data exclusivity.

Hatch-Waxman provides four pathways (or "Paragraphs") a generic firm may follow in order to gain entry into a market (Figure 3). The process starts with the filing of an Abbreviated New Drug Application (ANDA) by a generic manufacturer with one of the four Paragraph certifications. A Paragraph I certification is one for which the originator firm has not filed patent information for its branded product. Paragraph II certification is used when the branded product's patent has already expired (*i.e.*, market exclusivity has already ended), and Paragraph III certification is invoked when the generic manufacturer notes that the patent on the branded product *will* expire on a certain date and that it seeks to enter only after patent expiry. The fourth certification, Paragraph IV, is used when the generic manufacturer claims that it does not infringe on a branded product's patents or the generic firm argues that the patents protecting the branded product are invalid.

A Paragraph IV certification can be filed with the FDA after the conclusion of data exclusivity anytime during the market exclusivity window.[6] If it is successful, a Paragraph IV challenge can significantly decrease the effective patent life of branded products, bringing generics to the market earlier than otherwise would be the case (Higgins and Graham, 2009; Grabowski and Kyle, 2007). It is important to emphasize, however, that the multiple avenues provided by Hatch-Waxman for generic entry (currently) apply only to chemical-based or small molecule drugs. Throughout our sample period, there was no legal mechanism (in the U.S. market) through which the manufacturer of a "biosimilar" (the industry term for the generic version of a branded biologic-based drug) could demonstrate that its substance was equivalent to the original drug.

With no way to establish bioequivalence, any generic version of a biologic-based drug would have to undergo separate clinical trials to receive FDA approval -- a barrier to entry so daunting that no biosimilar has yet been introduced in the U.S. market. This historical absence of an entry pathway for biosimilars reflects, in part, the nascent state of the biotech industry when Hatch-Waxman was passed, as well as the real scientific challenges of determining bioequivalence for biologic-based drugs, which are far more complicated than chemistry-based drugs and interact with human biophysical systems in ways that are not always perfectly understood. Under the Obama Administration, new legislation provided the legal basis for biosimilar entry, but that legislation guarantees biologic-based drugs 12 years of data exclusivity - a period of legal monopoly 2.4 times longer than that afforded to chemistry-based drugs. Furthermore, the enabling regulations that would permit biosimilar entry have yet to be finalized by the FDA. Europe already permits biosimilars, but generally requires limited clinical trials to confirm bioequivalence prior to approval. The high cost of these trials - even when they are limited in time and

[6] Generic manufacturers may file a Paragraph IV certification up to one year prior to the end of data exclusivity but the FDA may not act on it until the conclusion of data exclusivity.

scope relative to those required of innovator dugs -- will likely constrain generic entry in the biotech side of the pharmaceutical market for the foreseeable future.

While a starkly different statutory treatment of chemical-based and biologic-based drugs has been established in U.S. law since the passage of Hatch-Waxman, the practical impact of these very different regulatory regimes has significantly strengthened in recent years. Generic penetration at the end of the 1980s and in the early 1990s was constrained by a FDA scandal that temporarily slowed down the processing of new generic drug applications, and by an unusually productive era of new drug introductions by the branded drug companies that extended into the mid-1990s.[7] Since then, however, generic penetration has intensified sharply. This has been partly driven by the rising incidence of Paragraph IV challenges. For example, by the end of the 2000s, ANDA applications with Paragraph IV certifications accounted for more than 40% of all generic filings (Higgins and Graham, 2009; Berndt *et al*., 2007).

3 Pharmaceutical Innovation and Generic Entry

We began our paper with the claim advanced by Nordhaus (1999) that the advances in human welfare generated by better medical science over the past half century equal in value the consumption increases from all other sources put together. Nordhaus's claim is backed up by evidence documenting the extensive gains in longevity and other dimensions of human health over the period; multiplying these gains by even conservative estimates of the value of a "statistical life" result in very large numbers (*e.g.*, Murphy and Topel, 2006). The work of Lichtenberg (2001, 2004, 2007) and others has lent credence to Victor Fuchs' (1982) assertion that the most important driver of this improvement has been pharmaceutical innovation. Efforts to infer the welfare impact of pharmaceutical innovation using modern models of demand for differentiated products, such as Ellickson *et al*. (2001), Cleanthous (2002), and Dunn (2012), have also yielded large estimates. Coincident advances in nutrition, pollution abatement, diagnostic techniques, and the gradual decline of unhealthy behaviors like tobacco smoking make it difficult to determine exactly what fraction of the observed improvement in health outcomes is attributable to new drugs, but few would contest the unique importance and impact of pharmaceutical innovation. This implies that public policies affecting the rate and direction of pharmaceutical innovation also take on special importance.

---

[7] The FDA scandal was widely covered in the media at the time -- see, for example, *New York Times* (1989). Cockburn (2006) discusses shifts in the measured research productivity of the pharmaceutical industry. A large cohort of new and successful branded products entered in the marketplace in the 1980s and 1990s, limiting the market importance of generic competition. As this wave of products lost patent protection -- or was challenged under Paragraph IV -- and was not fully replaced by newly introduced branded products, the financial pressure generated by generic competition increased.

3.1 Pharmaceutical innovation: costs and controversies

Pharmaceutical innovation is not just important -- it also difficult, time-consuming, risky, and expensive. A comprehensive accounting of costs has to include expenditures on drug candidates that fail at some point in the process. Recent estimates by DiMasi and Grabowski (2012) suggest that these costs have risen as high as one billion dollars per approved drug, though these cost estimates have been subjected to considerable criticism and controversy. Previous studies have described the various stages of the drug development process, including DiMasi, Hansen, and Grabowski (1991, 2003), DiMasi and Grabowski (2012), and Mossinghoff (1999). This process is typically divided into the following phases: *pre-discovery*, *drug discovery*, *preclinical development*, and *clinical trials*.

In the *pre-discovery phase*, drug companies study the basic scientific research of other firms and public science institutions and conduct their own, as they seek to understand the fundamental biochemical mechanisms that underlie diseases and the kind of chemicals or proteins that might work to disrupt or reverse those mechanisms, curing the patient. *Drug discovery* begins when drug companies start identifying and testing specific compounds. In the early stages, it is common for companies to evaluate thousands of compounds, using chemical tests and other means, before focusing on a few hundred compounds in the *pre-clinical stage*. The preclinical stage involves more in-depth, focused, comprehensive testing of this smaller number of compounds, including tests of drugs in animals. The time that it takes a compound to move through the drug discovery and preclinical phases is generally 3-6 years.

When drug companies have identified compounds they wish to subject to *clinical trials* in human subjects, they submit an Investigational New Drug (IND) application to the FDA; this is legally required in order to move drug samples across state lines for the purposes of clinical testing. Firms must then move through three separate phases of clinical trials, each involving a larger number of human subjects. In Phase 1, a small group is tested to determine a safe dosage level and identify side effects. In Phase 2, the treatment is administered to a larger group, to determine effectiveness and also further evaluate its safety. In Phase 3, the treatment is administered to a still larger group and compared to commonly used treatments. When Phase 3 is successfully completed, the drug company submits a New Drug Application (NDA) to the FDA, including clinical trials results. The FDA evaluates this information before approving the drug. Once it is approved and sales begin, drug companies continue to do Phase 4 trials to acquire additional information on risks, benefits, and optimal use. DiMasi and Grabowski (2012) contend that only one drug obtains FDA approval for every 5 compounds that enter Phase 1, and it can take 6-7 years for a compound to move through all 3 phases. The total development cycle from discovery through approval can take, on average, nearly 12 years, and the distribution of approved drugs is characterized by

highly skewed returns. Pharmaceutical firms rely disproportionately on a small number of very successful products to maintain their financial viability.

Starting in the mid-1990s, however, the number of drug approvals fell sharply, even as industry R&D expenditures continued to increase. This led to an intense debate about the industry's research "productivity crisis" (Cockburn, 2006 and Scherer, 2010). The relatively low level of new product approvals persisted throughout our sample period and beyond. Experts disagree as to the causes or future persistence of this productivity slowdown. Nevertheless, it has created a rising level of concern (and financial stress) within the industry. Accelerating generic competition has been narrowing the profits of branded firms faster than successful new drug development has expanded them.

3.2 The rise of generic penetration and implications for pharmaceutical innovation

A number of recent studies have studied the intensification of generic competition in recent years and the impact of this shift on branded drug companies. We lack the space here to offer a comprehensive review of all the work in this domain, and, instead, cite selectively the work that is most relevant to our own analysis. Caves *et al.* (1991) offered an influential look at the early impact of Hatch-Waxman. More recent work includes Reiffen and Ward (2005), Saha *et al*. (2006), Grabowski (2007), Grabowski and Kyle (2007), and Berndt and Aitken (2010). Efforts to calculate the welfare impact of generic entry include Bokhari and Fournier (2013) and Branstetter *et al*. (2013). The latter study shows that the rising incidence of Paragraph-IV challenges has brought substantial gains to consumers. Hemphill and Sampat (2011, 2012) also focus on Paragraph-IV challenges, analyzing, among other things, which incumbent firms' patents tend to be challenged.

The possibility that rising generic penetration could undermine the incentives to undertake new drug development has been recognized in prior work. For example, Hughes *et al*. (2002) show in a theoretical model that providing greater access to a current stock of branded prescription drugs yields large benefits to existing customers. However, this access comes at a cost in terms of lost consumer benefits from reductions in the flow of future drugs. Other papers have also discussed this possibility, including Grabowski and Kyle (2007), Higgins and Graham (2009), Knowles (2010), and Panattoni (2011). This research stream has provided (mostly indirect or anecdotal) evidence suggesting that an intensification of generic competition has undermined incentives for R&D. However, to the best of our knowledge, no published study has yet provided direct econometric evidence demonstrating that generic

entry has caused a change in the rate or direction of new drug development.[8] The extent to which this occurs in practice remains an open question.

4 Empirical Models and Data

Previous research in this area has struggled with data limitations. We are fortunate to have access to a range of unique and comprehensive data sets that provide us with a useful degree of leverage over some of the econometric and measurement challenges we confront. Since we seek to measure the impact of rising generic penetration on drug development effort, it is especially important to have high-quality measures of pharmaceutical innovation and of exposure to generic competition. Our data allow us to track both variables by firm, by therapeutic category, and by year. The paragraphs below describe our data and our empirical approach.

4.1 Measuring and modeling pharmaceutical innovation

The regulatory structure imposed on the pharmaceutical industry makes early-stage product development relatively easy to track. Before obtaining approval to market a new drug, pharmaceutical firms must bring each prospective new product through a series of clinical trials, each one more comprehensive than the previous one. Because the introduction of new drugs is so important for the financial health of drug companies, the progress of new candidate drugs through the development "pipeline" is closely monitored, and commercial databases contain rich data on these candidates. We draw our measures of drug innovation from one such commercial database, Pharmaprojects. Not only is there nearly universal coverage of all candidate drugs being tested for eventual sale in the U.S. market, but we also know the chemical composition of the drug, the prospective disease targets, the therapeutic market in which it is likely to be sold, and the development history (some drugs are initially developed to fight one disease but then are discovered to have positive effects against others). The database also records information on product development suspensions and discontinuations as well as product withdrawals from the market after introduction. The richness of the data allows us, in principle, to examine the relationship between rising generic penetration and the emergence of new compounds through various stages of the drug development process across firms, therapeutic categories, and time.

However, attempts to assess this relationship confront a major challenge. At the same time that generic entry has been rising, the pharmaceutical industry has encountered a widely publicized "productivity crisis" (Cockburn, 2006). Although there has been no measured slowdown in aggregate early stage drug development, new drug approvals peaked in the mid-1990s and were stagnant or falling

[8] In related work, Budish, Roin, and Williams (2013) provide evidence that variation in effective patent life distorts incentives for investment in cancer drugs. This study does not consider the impact of rising generic competition

through the rest of our sample period. While this opinion is by no means universally held, there are some inside and outside the industry who suggest that this decline reflects an emerging exhaustion of research opportunities. In this view, the easy-to-discover drugs have already been introduced; and, the diseases that are now the focus of research effort are extremely complex and difficult to treat. To the extent that there really is a decline in research productivity, this could lead firms to ratchet back their drug development efforts, even in the absence of a growing generic threat to profitability. Our empirical challenge will be to assess the impact of increased generic entry on new pharmaceutical innovation while controlling, as best we can, for contemporaneous changes in research opportunities and other factors that might influence drug development.

We propose to do this using a regression specification that models innovation as a function of scientific opportunity and challenges, firm innovative capability, downstream co-specialized assets, expected market size, and generic entry, with a vector of additional controls:

$$I_{ijt} = \alpha_i + \alpha_j + \alpha_t + \beta_1 G_{ijt-1} + \beta_2 O_{jt-1} + \beta_3 V_{jt+2} + \beta_4 Z_{ijt-1} + \beta_5 D_{ijt-1} + \beta_6 P_{ijt-1} + \beta_7 SA_{ijt} + \beta_8 S_{it} + \varepsilon_{ijt} \quad (1)$$

where I_{ijt}, measures early-stage innovations by firm *i* in ATC 2-digit market *j* in time *t*. We define "early stage innovations" as the count of individual compounds in "preclinical development" or in Phase 1 clinical trials. If firms are responding to changes in the intensity of generic competition, changes in perceived scientific opportunity or changes in expected market opportunity, we would expect a measurable impact to show up at this stage. In contrast, drugs that have already moved on to Phase 2 or Phase 3 trials are likely to continue through the development process to the end, even if the firm plans to curtail or eliminate future research in that area in response to rising competition or diminished technological opportunity.[9] Because the outcome variable is a count variable, the statistical model employed in our regression should be one designed to handle count data. As such, we use fixed effects Poisson and negative binomial estimators (Hausman*et al*., 1984; Woolridge, 1999). Given that not all firms innovate in each therapeutic category in each year, it is possible that the data may contain zeros. Our count data models have the advantage of dealing with this outcome in a natural way.

The specification, as written, includes fixed effects for year (α_t), firm (α_i), and therapeutic (ATC) category (α_j). There are 13 years, 178 firms, and 126 ATC 2-digit categories in our data, and we run into convergence challenges when we seek to estimate our Poisson and negative binomial count data models

[9] We present empirical results later in the paper that are consistent with this view.

using the full set of year and ATC 2-digit dummies. In the results we will report below based on count data models, we get around this convergence challenge by estimating ATC 1-digit dummy variables and, in some specifications, we also include a paired fixed effect, interacting therapeutic market dummies with year dummies, (α_j*α_t). As a robustness check, we also run linear versions of our models with the full set of 2-digit ATC fixed effects as well as interaction terms of 2-digit ATC fixed effects and year dummies. These more robust specifications yield results consistent with those reported in the main text -- our measures of generic penetration are negative and statistically (and economically) significant.

We depend on the Pharmaprojects classification of drug candidates into the various ATC categories. Unfortunately, this data is most consistently reported only at the 2-digit level. Other key variables are available at a greater level of disaggregation (i.e., ATC 4-digit), but because we are seeking to relate these to innovative effort, we can disaggregate no further than the level of our innovation data. The regressions that are described below are therefore run at the level of firm-ATC 2 digit-year. Firms are included in our sample if they have at least one approved product and at least one early-stage innovation. This limitation excludes some smaller, research-intensive firms. We argue below that the bias introduced by this sample selection, to the extent that it exists, likely weakens our estimated results relative to what holds in reality.

4.2 Measuring generic penetration (G_{ijt-1} and $G_{jt-1}M_{ijt-1}$)

Hatch-Waxman laid out the modes by which generic manufacturers can enter chemical-based therapeutic markets. This entry leads to rapid deterioration in the sales of branded products (Saha *et al.*, 2006). However, the incidence of rising generic impact is quite uneven across therapeutic categories and time. Firms also differ in terms of their exposure to this competition. Fortunately, we are able to employ disaggregated data from the IMS MIDAS™ database. This database tracks the sales of nearly every pharmaceutical product sold in the U.S. by firm, product, and quarter, and the data are mapped to ATC categories at the 4-digit level. Our data is limited to the years 1998-2010, and this data restriction determines the time dimension of our study.

Fortunately, this window covers a period of intensifying generic competition. Within this period, we are able to determine the extent of generic penetration that firm *i* faces in therapeutic area *j* in time *t-1*. We define our first measure of generic penetration, G_{ijt-1}, as the sum of generic sales in therapeutic area *j* at time *t-1* divided by the sum of generic and firm *i* branded sales in therapeutic area *j* at time *t-1*. As a robustness check, we define a second measure of generic penetration, $G_{jt-1}M_{ijt-1}$, as the ratio of generic sales to total sales in therapeutic area *j* in time *t-1* multiplied by the ratio of branded sales by firm *i* in

therapeutic area *j* in year *t-1* divided by total branded sales of firm *i* in year *t-1*. We thus measure changes in the intensity of generic competition across therapeutic areas, and weight them by their relative importance in the branded sales portfolios of individual firms. A negative coefficient on either measure implies that as generic penetration in a therapeutic market increases, the flow of innovations decreases.

4.3 Measuring scientific opportunity (O_{jt-1})

In order to identify the effect of changes in generic competition on innovation, we must also effectively control for underlying scientific opportunities within each therapeutic market *j* at time *t-1*. Prior research has demonstrated the link between academic research and industrial R&D (*e.g.*, Mansfield, 1995; Gittelman and Kogut, 2003); these linkages are particularly strong in pharmaceuticals. Similar to Furman *et al*. (2006), we construct a bibliographic measure that captures publicly available academic research in the life sciences.

We start by merging data from IMS MIDAS™, our comprehensive database of pharmaceutical products, with the IMS NDTI™ database, which captures physician prescription behavior. This latter database identifies the diseases for which physicians are actually prescribing the drugs in MIDAS™.[10] IMS MIDAS™ is categorized by ATC codes and the IMS NDTI™ database is categorized by International Statistical Classification of Disease (ICD-9) diagnostic codes. Merging these two databases enabled us to generate a concordance between ICD-9 diagnostic codes and ATC product codes (at the 4-digit level). Next, we extracted the top 10 ICD-9 diagnostic codes for each ATC 4-digit category. These ICD-9 codes have unified keywords associated with them that were used as search terms in the National Library of Medicine's PUBMED database. This search yielded journal articles published between 1950 and 2010 relating to our various keywords that we were then able to map back to disaggregate ATC 4-digit categories. Ultimately, we identified a unique sample of 6.5 million journal articles. However, some journal articles were mapped to multiple ATC 4-digit categories, thereby yielding 20.9 million raw article counts. Next, we used the unique PMID identifiers for these articles to gather their forward citations from the year of publication to the end of 2010 in the SCOPUS Sciverse database. Our sample of 20.9 million articles generated over 345 million forward citations. Finally, since our unit of observation in a therapeutic market is at the ATC 2-digit level, we aggregate our annual, citation-weighted counts of journal articles up from the ATC 4-digit level to the ATC 2-digit level, take natural logs, and lag the stock by one year to create our control variable, O_{jt-1}.

[10] Because the IMS NDTI™ database is based on surveys of practicing physicians, it captures "off-label" prescribing behavior -- that is, the prescribing of medicines for diseases for which they are not officially approved by the FDA as treatments.

4.4 Expected future market size (V_{jt+2})

Innovative effort is likely to be driven, in part, by perceived market opportunities as well as scientific opportunity (Acemoglu and Linn, 2004). In other words, firms will develop more drug candidates in disease categories where the potential addressable market is larger. As a way of controlling for this, we average total sales from IMS MIDAS™ in therapeutic area *j* over year t, year t+1, and year t+2, measured in inflation-adjusted dollars, and insert this as a control, effectively presuming rational and accurate expectations on the part of companies developing new drugs. We do not assign any causal interpretation to the estimated coefficient on V_{jt+2}. We incorporate this measure solely as a way of controlling for the association in our data between perceived market size and R&D investment. As it turns out, the sign and significance of the coefficients on our measures of generic penetration are not sensitive to the inclusion or exclusion of this variable, but, for completeness, we will report specifications in which this measure is included.

4.5 Scientific challenges (Z_{ijt-1})

In contrast to scientific opportunities that may potentially "pull" firms *towards* a specific therapeutic market, we control for scientific challenges that may "push" firms *away* from a specific therapeutic market. Utilizing data from Pharmaprojects we identify all suspended, discontinued and withdrawn products across the entire research pipeline from pre-clinical candidates to approved products. Development can be ended and products pulled for a multitude of reasons many of which, at their most fundamental level, are due to some type of scientific challenge. For example, Merck pulled Vioxx® from the market due to negative side-effects, while the Alzheimer disease drug candidate *semagacestat* was discontinued by Eli Lilly in Phase III clinical trials after disappointing results. The failure of one or more leading products within a broader drug development program could indicate the presence of common or related flaws in the products that are still under development. This, in turn, could lead the firm to scale back, terminate, or redirect research and development efforts in response. Seeking to control for this, we define our proxy for the scientific challenges faced by the firm, Z_{ijt-1}, as the number of products suspended, discontinued or withdrawn by firm *i*, in therapeutic market *j* at time *t-1*.

4.6 Firm capabilities (D_{ijt-1} and P_{ijt-1}), marketing assets (SA_{ijt}), and firm size (S_{it})

Clearly, pharmaceutical companies differ in the drug development capabilities they have built over time. A given firm is more likely to introduce a new compound in a therapeutic category in which it already has substantial research expertise. In order to control for this persistence of firm-level capabilities

we use data from Pharmaprojects to create a three-year moving average of past drug introductions, D_{ijt-1}, by firm *i* in the same therapeutic market *j*. This three-year moving average is lagged one period, (*t-1*). In addition to controlling for past products, we also control for late-stage innovations within the product pipeline. Again, using data from Pharmaprojects we define P_{ijt-1} as the number of compounds under development by firm *i* that are in Phase 2 or Phase 3 clinical trials in therapeutic market *j* at time *t-1*.

Prior research has also documented the connection between downstream co-specialized assets and a strong commitment to research efforts within a particular therapeutic class (Teece, 1986; Chan *et al.*, 2007). The presence of these assets can create a 'lock-in' effect, influencing the allocation of research effort across therapeutic categories. Similar to Ceccagnoli *et al.* (2010), we control for the distribution of a firm's downstream co-specialized assets across therapeutic categories by including a ratio of promotions to product sales, SA_{ijt}, for firm *i* within therapeutic category *j* at time *t*. Promotions and product sales are collected from IMS MIDAS™ and the promotions activity includes detailing, journal advertising and direct-mail. Detailing is the direct promotion of products by pharmaceutical representatives to physicians. Finally, firm size can impact innovation rates. As such, we control for firm size with pharmaceutical sales by firm *i* in year *t*, S_{it}. Sales data was gathered from IMS MIDAS™ and natural logs were taken.

4.7 An empirical specification for measuring the shift into biotech drugs

Current regulation suggests an alternative approach to estimating the impact of generics on innovation. Chemistry-based pharmaceutical products become susceptible to Paragraph III generic entry after patent expiration (*i.e.*, end of market exclusivity). They also become susceptible to early generic entry via Paragraph IV challenges only five years after approval (*i.e.*, end of data exclusivity, Figure 2). The same legal frameworks do not (yet) provide a pathway for biosimilar entry after biologic patent expiration, nor is there the equivalent of a Paragraph IV challenge to biologic-based drugs. Furthermore, biologic-based products are now explicitly guaranteed 12 years of data exclusivity, so even if and when Paragraph IV-type challenges of biologic drugs become feasible, they will occur much later in the product life cycle. Clearly, this difference in regulation creates an incentive for pharmaceutical companies to favor biologic-based ("large molecule") therapies over chemical-based ("small molecule") therapies, even if the latter may be more effective in a purely therapeutic sense. This suggests an alternative specification:

$$CI_{ijt} - BI_{ijt} = \alpha_i + \alpha_t + \alpha_j + \beta_1 G_{ijt-1} + \beta_2 O_{jt-1} + \beta_3 Z_{ijt-1} + \beta_4 (CD_{ijt-1} - BD_{ijt-1}) + \beta_5 (CP_{ijt-1} - BP_{ijt-1}) + \beta_6 (CSA_{ijt} - BSA_{ijt}) + \beta_7 S_{it} + \varepsilon_{ijt} \quad (2)$$

Here, the dependent variable measures the difference between chemistry-based innovations and biologic-based innovations. Likewise, our controls for firm-specific development capability and market

presence are redefined to reflect relative capability in chemistry-based versus biologic-based development. Given these controls, we would not expect generic penetration (G_{ijt-1}) to have an impact on the choice of technology – unless firms' research choices are being affected by the prospect of generic competition.

4.7 Difference in early-stage innovation ($CI_{ijt} - BI_{ijt}$)

If current regulation is in fact causing biologic-based innovation to be preferred to chemical-based innovation, then we need to modify our innovation measure in order to capture this change. Using the *Origin of Material* field within Pharmaprojects we are able to sort early-stage innovation (I_{ijt}) into either biologic-based (BI_{ijt}) or chemical-based (CI_{ijt}) innovation. In operationalizing Equation (2), the dependent variable is the difference between these two types of innovation, $CI_{ijt} - BI_{ijt}$. A negative coefficient on a right-hand side (RHS) variable (such as G_{ijt-1}) would imply that as that variable increased the difference ($CI_{ijt} - BI_{ijt}$) would decline. In other words, BI_{ijt} is greater than CI_{ijt} or the flow of biologic-based innovations exceeds the flow of chemical-based innovations.[11]

It is possible for firm i to have more biologic-based innovations than chemical-based innovations in therapeutic market j at time t. In this case, our difference variable ($CI_{ijt} - BI_{ijt}$) will become negative, preventing us from using count data models. We therefore create a new variable, $cat(CI_{ijt} - BI_{ijt})$, that equals 1, 2 and 3 if ($CI_{ijt} - BI_{ijt}$) is negative, zero or positive, respectively. This reclassification allows us to use an ordered logit specification (Hausman *et al.*, 1992). Again, a negative coefficient on an independent variable would imply that as that variable increased, the dependent variable, $cat(CI_{ijt} - BI_{ijt})$, will decline. In this case the difference, ($CI_{ijt} - BI_{ijt}$), will become negative and the interpretation is the same as above.

For our specification in Equation 2 we can use the *Origin of Material* field within Pharmaprojects to decompose our measure of late-stage innovations, P_{ijt-1}, past drug introductions, D_{ijt-1}, and our measure of scientific challenges, Z_{ijt-1}, faced by firm i in therapeutic market j, into their chemical-based (CP_{ijt-1}, CD_{ijt-1} and CZ_{ijt-1}, respectively) and biologics-based (BP_{ijt-1}, BD_{ijt-1} and BZ_{ijt-1}, respectively) components. We can also decompose our ratio of promotions to product sales, SA_{ijt}, for firm i within therapeutic market j at time t, into its chemical-based (CSA_{ijt}) and biologic-based (BSA_{ijt}) components. Empirically, in Table 6 we create the variables $diff(P_{ijt-1})$, $diff(Z_{ijt-1})$, $diff(D_{ijt-1})$, and $diff(SA_{ijt})$ defined as the difference

[11] As a robustness check for this specification we also employ a SUR model (Table 7). Results are consistent between our various specifications and will be discussed more fully in Section 5. We thank Ivan Png for this suggestion.

between the chemical- and biologic-based components: ($CP_{ijt-1} - BP_{ijt-1}$), ($CZ_{ijt-1} - BZ_{ijt-1}$), ($CD_{ijt-1} - BD_{ijt-1}$) and ($CSA_{ijt} - BSA_{ijt}$), respectively.

5 Empirical Results

5.1 Descriptive statistics

Descriptive statistics for our variables are presented in Table 1. Our dependent variable, I_{ijt}, captures early-stage innovation and varies between 0 and 36 for firm *i*, in therapeutic market *j*, at time *t*. While our firms had, on average, 0.78 early-stage innovations within a therapeutic market at time *t*, it should be remembered that not every firm has an early-stage innovation in every therapeutic market in each year. If we focus solely on therapeutic categories with activity, then the average increases to 2.12 early-stage innovations. Firms in the top quartile of firm size had, on average, 3.07 innovations within therapeutic market *j* at time *t*, as compared to 1.45 innovations for the smallest quartile firms. ATC N, focusing on the nervous system, had the largest number of innovations, while ATC P, which focuses on anti-parasitic products, had the lowest number of innovations. The relative contribution to total innovations of each broad (ATC 1-digit) therapeutic category over our sample period is displayed in Figure 4.

Inspection of the raw data shows that, in the aggregate, there has been no decline in early-stage innovation over our sample period, even as the level of generic penetration has risen and the number of approved drugs has fallen (Figure 1). This suggests that generics have had limited impact on the overall aggregate rate of early-stage innovation. However, we find strong evidence that generics have had a statistically and economically significant impact on where development activity is concentrated and how it is undertaken.

Our baseline measure of generic penetration, G_{ijt-1}, has a mean value of 54% and a median value of just over 80%. Over our sample period, measured generic penetration increased significantly. Our measure of technological opportunity, O_{jt-1}, measured by the logarithm of the stock of citation weighted articles in year *t-1* for therapeutic market *j*, varied between 0 and 17.9, with an average of 8.09. This translates into an average value of approximately 4.35 million citations for each therapeutic market *j* in each year. Over our sample period the greatest technological opportunity existed in ATC categories N5 (psycholeptics) and N6 (psychoanaleptics). Our measure of technological challenges, Z_{ijt-1}, had an average value of 0.05. The number of challenges varied between 0.26 and 6 with the greatest technical challenges experienced in ATC T2, which includes various recombinant-based products, such as interferon.

On average, our firms had a lagged three-year moving average of 0.24 recently introduced products (D_{ijt-1}) and 0.09 drug candidates in the latest stages of product development (P_{ijt-1}) in therapeutic market *j* at time *t-1*. Our control for downstream co-specialized assets, SA_{ijt}, the ratio of promotions to sales for firm *i* in therapeutic market *j* at time *t*, averaged 45%. This suggests firms are making significant downstream investments in therapeutic areas in which they operate (and plan to operate).

5.2 Impact of generic entry on the flow of innovation

Do changes in generic penetration have an effect on the flow of early-stage drug innovation? We estimate Equation 1 with a fixed effects Poisson specification (Table 2). We also present results using a fixed-effects negative binomial specification (Table 3). The dependent variable in all specifications is I_{ijt} , or the count of firm *i* innovations in therapeutic market *j* at time *t*. Model 1 in both tables (Table 2 and Table 3) presents a baseline regression with our market size measure and firm-level control variables, including our measures of new product introductions, late stage product development, advertising, sales, and firm, year, and therapeutic market fixed effects (estimated at the ATC 1-digit level). Model 2 in each table adds controls for scientific opportunity (O_{jt-1}) and scientific challenges (Z_{ijt-1}). Finally, in Models 3, 4, and 5, we include our baseline measure of generic penetration along with differing sets of fixed effects. Model 3 includes just firm and year fixed effects; Model 4 adds therapeutic area fixed effects, while Model 5 includes an interaction between the year and therapeutic market fixed effects. This interaction, we argue, controls for unobserved changes in a particular therapeutic market in a specific year. Model 6 replaces our baseline measure of generic penetration, G_{ijt-1}, with our alternative measure, $G_{jt-1}M_{ijt-1}$. The results presented in this table are obtained using clustered standard errors at the firm level. When we cluster our standard errors at the therapeutic area level (ATC), we obtain results qualitatively similar to those shown here.[12]

Across all specifications and models we find negative and statistically significant coefficient estimates for our measures of generic penetration. This negative relationship suggests that, at the firm level, increases in generic penetration are related to decreases in the flow of early stage innovation *in that therapeutic area*. Taking the coefficient from our complete negative binomial specification (Model 5, Table 3) as our baseline estimate, we calculate an elasticity equal to -0.79. In other words, a 10% increase in generic penetration experienced by a firm in a particular market is related to a 7.9% decrease in early-stage innovation by that firm in that market. To our knowledge this is the first empirical evidence that documents the effect of generic penetration in the U.S. market on early-stage pharmaceutical innovation.

[12] We replicate Tables 2 and 3 with clustered standard errors at the therapeutic area level in Appendix Tables A1 and A2.

If fewer candidates are entering a given therapeutic pipeline within a given firm, then fewer drugs will eventually come out.

Generic penetration into a market is clearly harmful for branded producers. From a social welfare perspective, however, the interpretation is more nuanced. If the presence of viable generics in a market rises, our results indicate that innovation will decrease in that market.[13] However, the stability of early-stage drug development effort at the aggregate level, suggests that much of the decline in innovation within markets facing a high degree of generic competition is offset by increased innovative effort elsewhere. Indeed Pammolli *et al*. (2011) argues that one of the reasons R&D productivity has declined has been a shift into areas with unmet therapeutic needs, which also have higher risks of failure. Our results are consistent with the view that drug development is shifting out of therapeutic areas facing more intense generic competition and into domains facing less generic competition. Furthermore, our results provide one possible explanation for why this shift is occurring. In essence, Hatch-Waxman, by providing mechanisms of entry for generics, creates conditions under which the pharmaceutical industry redirects R&D efforts to markets less served by generics.

If R&D efforts are shifting across therapeutic areas, this can have significant future consequences, with a net impact on social welfare that is difficult to calculate. On the one hand, if the therapeutic category that is seeing research expenditures leave has a different success probability than the therapeutic category to which expenditures are flowing, this could have consequences for the net flow of innovation (either increasing or decreasing). On the other hand, new product development in a domain with few (or no) existing effective therapies may have greater social value than similar development in an area with a broad range of existing effective therapies, even if the R&D success probabilities are lower in the domain with few therapies. In this paper, we do not take a strong stand on the ultimate welfare consequences of this shift. Instead, we seek to document its existence and magnitude. The welfare consequences of the shift remain the focus of ongoing research.

[13] In theory, generics should be perfect substitutes for branded drugs since they are bioequivalent. Cleanthous (2002) shows that the data do not support this relationship and suggests this is the result of 'spurious product differentiation', which he defines as arising "…when consumers perceive physically identical products to differ in quality." Recent evidence, however, may suggest that consumer perceptions have merit and while drugs may be bioequivalent, they may indeed differ in quality. Several articles appeared in the April 17, 2007 edition of the prestigious journal *Neurology* discussing the high incidence of break-through seizures with generic anti-epileptics. Insurance companies such as Blue Cross Blue Shield of Georgia allow pediatric customers to stay on branded anti-epileptic medications even though generics are available. Differences across generics for the same brand have also been reported. We are not suggesting all generics have problems but it appears in some instances where the therapeutic window is very narrow these perceptions may have some merit.

Turning to our controls for scientific opportunity (O_{jt-1}) and scientific challenges (Z_{ijt-1}), we find that both positively and significantly influence the flow of early-stage innovation. Using a similar approach in the creation of their scientific opportunity variable, Furman *et al*. (2005) find a positive relationship with pharmaceutical patenting. Our results take this one step further and document a relationship with actual early-stage drug development. Much of the basic science research that is captured in our variable takes place in academic settings; as such this finding is broadly consistent with past work documenting the role of academic research in industrial innovation (*e.g.*, Mansfield, 1995; Cohen *et al*., 2002).

Interestingly, while our findings are consistent with our *a priori* beliefs with respect to scientific opportunity, the same cannot be said with respect to scientific challenges. Our initial beliefs were that scientific opportunity might serve as a mechanism to 'pull' innovative effort into a particular area, while challenges might serve as a mechanism to 'push' innovation away from it. That would imply a negative coefficient on our challenges variable -- but the coefficient is positive and significant at conventional levels. One interpretation of this positive coefficient is that that failures can serve as a learning mechanism for future endeavors (Chiou *et al.*, 2012). Statin drugs, which today are one of the largest selling therapeutic areas, had a difficult beginning in 1978, with the unsuccessful launch of Mevacor®. Over time, however, the industry worked through these difficulties as new technologies led to the five types of statin-molecules currently sold in U.S. A different, but related interpretation of the positive coefficient is simply that firms with a significant research commitment to a particular therapeutic category are more likely to have a few failures along the way, and the positive coefficient on our proxy for scientific challenges simply picks up that effect.

The sign of the coefficient on our expected market opportunity variable, V_{jt+2}, is not uniform across all specifications, but tends to be negative and statistically significant in the more complete specifications. Does this result imply that firms systematically invest less in product development in the more lucrative markets? We do not believe so. While we stress that this variable is included solely as a control, and we do not place any strong causal interpretation on the estimated coefficient, we believe that the negative coefficient is consistent with the fact that larger product markets, in terms of total sales, are also more crowded product markets, in terms of the number and range of existing therapies. Rather than increase drug development in a market that is already well served by a range of existing products, firms may elect to direct their efforts to less developed markets.

We control for firms' research capabilities by using their past innovative output in a particular therapeutic market, as measured by the lagged count of products in late-stage product development, P_{ijt-1},

and the lagged count of new product introductions, D_{ijt-1}. As expected, both are positively and significantly related to the flow of early-stage innovations. To some extent, these variables may pick up part of the positive impact of market size on investment in new products - new product introductions and late stage development may both be more likely in larger markets - and the impact of these variables, which vary over firms as well as across markets and time, may help explain the negative coefficient on V_{jt+2}. Across the two baseline specifications (Model 5, Table 2 and Table 3), our measure for firm size, S_{it}, is mostly positive and significant. This result should not be interpreted as necessarily indicating a positive relationship between firm size and 'innovation,' as our dependent variable is a simple count of early-stage pipeline products; we make no distinction between internally generated and externally acquired products.[14] Finally, our measure of marketing intensity or downstream co-specialized assets, SA_{ijt}, is statistically indistinguishable from zero.

5.3 The (nonexistent) impact of generic competition on innovation in markets with limited substitution between branded and generic products

Most prescription health plans in the U.S. allow for the use of branded products until generics become available. In most cases patients will be given the generic by retail pharmacies unless the prescription is written "Dispense as Written" or if the patient explicitly asks for a branded drug (in which case there is usually a much higher co-payment). More recently, however, insurance firms have begun to actively engage in "cross-molecular" substitution. For example, let's assume there are three branded products in a particular market, *Drug A*, *Drug B* and *Drug C*, sold by three different pharmaceutical firms. Each branded product has a different chemical composition (i.e., a different molecule), and uses a different biochemical pathway to address the underlying illness. Then, a generic for *Drug B* enters the market. To save money, insurance companies can encourage patients taking *Drug A* or *Drug C* to switch to *Generic B*. While insurance firms cannot force patients to move (they request patients take a letter from the insurance company to their doctor) they can entice them with lower (or no) copayments for *Generic B*.

Since physicians, not patients or insurance companies, write prescriptions, these financial incentives will only shift drug consumption to the generic products if physicians also consent to the change. However, in many therapeutic markets, practicing physicians have long regarded different drugs, based on different molecules and utilizing different biochemical pathways to attack the disease, as equally effective therapies for the underlying illness. In such cases, physicians will often consent to the insurance companies preferred change, especially if it saves their patients money. We refer to this possibility of

[14] See Cohen (2010) for an extensive discussion of the literature analyzing the relationship between firm size and innovation.

substitution across drugs and molecules within a therapeutic category in response to emerging price differentials as that category's degree of *cross-molecular substitution* (Branstetter *et al*., 2013). Where cross-molecular substitution is high, the implications for branded products can be quite profound. In such markets, the emergence of a generic equivalent to *any* branded product can affect the revenue streams of *all* branded products, leading to wide-ranging declines in revenues and profits.

The extent of these impacts will vary across therapeutic categories, depending on the degree of cross-molecular substitution within that category. For example, based on conversations with practicing physicians, we would expect higher substitution in therapeutic areas such as anti-infectives, hypertension and allergies and lower substitution in markets such as depression and epilepsy. In general, the complexity and sensitivity of the human brain and the complicated nature of neurological disorders work to strictly limit the degree of cross-molecular substitution in drugs that treat neurological and psychiatric disorders. They even limit the degree to which practitioners are willing to use allegedly "bioequivalent" generic versions of the branded drug. When practitioners find a good match between a drug treatment and a patient in these domains, they are often reluctant to switch to a cheaper generic. For example, multiple articles in the April 17, 2007 edition of the prestigious journal *Neurology* discussed the high incidence of break-through seizures associated with transition to the use of generic anti-epileptics. These concerns and the associated costs of break through seizures led some insurance companies, such as BlueCross Blue Shield of Georgia, to allow pediatric customers to stay on branded anti-epileptic medications even though a generic was available (Branstetter *et al*., 2013).

Economic intuition suggests that if a class of branded drugs was less susceptible to cross-molecular substitution and generic competition, then we might expect to see a muted innovation response to rising generic competition in that particular market. Focusing on the markets that include anti-epileptics, anti-depressants, and anti-psychotics, we indeed see this in our results (Table 4). Increases in generic penetration, whether measured by G_{ijt-1} or $G_{jt-1}M_{ijt-1}$, do not appear to have any statistically significant effect on early-stage innovation in these therapeutic areas. This suggests that there are markets for which direct substitution to a generic may be problematic, cross-molecular substitution is low, and as a result the effect on early-stage innovation is less of a concern.[15]

*5.4 The impact of generic penetration on "novel" drug development, the impact on late-stage product development, and robustness to the (ATC2*year) fixed effects*

[15] As a robustness check, through consultations with practicing physicians we identified markets that they deemed 'high CMS', namely the anti-infective markets J01-J04. When we replicate the findings in Table 4 for these high CMS markets the coefficients on both G_{ijt-1} and $G_{jt-1}M_{ijt-1}$ are negative and significant at the 1% level. This is what we would expect to see in high CMS markets.

Our baseline measure of innovation, I_{ijt}, which is the count of products in early-stage development, does not discriminate between pharmaceutical products that are the first in their class and those that come much later in the history of a therapeutic area. This reflects, in part, the difficulty of drawing a clear or meaningful line between "truly innovative" drugs and "me-too" drugs. The history of the industry provides several examples in which the first products in a class had significant shortcomings or side effects -- and the real breakthroughs in terms of therapeutic efficacy came several product introductions later.[16] Even when new products are "merely" recombinations or reformulations of existing active ingredients, the new products can often provide significant therapeutic benefits to certain categories of patients.

Despite these realities, critics of the pharmaceutical industry have accused branded firms of responding to generic entry or the threat of generic entry by coming up with branded "innovations" that are not true innovations, but merely minor modifications of earlier branded products. If the negative impact of rising generic entry on new product introductions, identified in our regressions, were limited to incremental innovations with little or no therapeutic value, that would have different policy implications from an effect that extended to the most novel compounds and drugs.

The Pharmaprojects database includes a variable that grades each product under development in terms of its novelty -- the most novel compounds are ones that are first in their class. We do not accept the proposition that all compounds without this "novel" designation have limited therapeutic value. For the reasons discussed above, we believe the Pharmaprojects designation excludes a large number of socially useful new product introductions. Nevertheless, the designation allows us to introduce a useful robustness check that may address the concerns of those who are convinced that only pharmaceutical product introductions that satisfy a strict definition of novelty are socially useful. In Models 1 and 2 in Table 5, we present the results of regressions in which we replace our comprehensive count of drugs in early-stage development with a count of only "novel" drugs in early-stage development, as defined by Pharmaprojects. In fixed effects negative binomial regressions, the coefficients on both measures of generic penetration are negative and statistically significant, indicating that rising generic penetration is associated with a statistically significant decline in the rate of introduction of "novel" products. The elasticity from Model 1 implies that a 10% increase in generic penetration in a particular market will *lower* early-stage *novel* innovations, in that same market, by 4.6%. Put another way, our results are not driven by a crowding out of purely incremental inventions or reformulations.

[16] See Arcidiacono *et al*. (2013).

We carefully defined innovation as early-stage product development - counts of compounds in preclinical development or Phase 1 clinical trials. This choice was informed by conversations with industry practitioners. As compounds move through the costly, expensive, and risky clinical trials process, they require ever-higher levels of investment by the firm. A drug that has survived Phase 2 and Phase 3 clinical trials is likely to be introduced, even if generic penetration is rising sharply in a way that might lead to a throttling back of early-stage research in that therapeutic area. If our view of the drug development process, and the responsiveness of its different stages to changes in generic penetration, is correct, then it suggests a further robustness check. Drugs at the later stages of the development process should be significantly less responsive to our measures of generic penetration than our measures of earlier stage innovations.[17]

Following this logic, in Models 3 and 4 of Table 5, we redefine our dependent variable, *Late Stage* I_{ijt}, as a count of firm *i*'s products in Phase 2 or Phase 3 trials in market *j* at time *t*. In these models, we find that neither measure for generic penetration, G_{ijt-1} or $G_{jt-1}M_{ijt-1}$, is significantly correlated with late-stage product innovation.[18] This is in line with our expectations, and strengthens our interpretation of the results using measures of early-stage product development.

Our final robustness check seeks to incorporate ATC 2-digit fixed effects and the interactions of ATC 2-digit market fixed effects and year fixed effects into the specification. This is not feasible in the fixed effects negative binomial models -- attempts to estimate these nonlinear specifications with so many fixed effects fail to reach convergence. However, it is possible to incorporate ATC 2-digit fixed effects and even the interaction of ATC 2-digit fixed effects with year effects into a linear specification of Equation (1). The results for the full specification, including ATC2 and (Year*ATC2) interaction fixed effects, are shown in Models 5 and 6 of Table 5. We view this an especially strong test of the hypothesis that an increase in generic penetration is associated with a decline in innovative activity, because all of the factors associated with an ATC 2-digit market that vary over time in a common way across firms are swept out with the interaction terms. Despite this, and despite the imperfect fit between the count data in our dependent variable and the statistical assumptions undergirding our linear specification, we still find that both measures of generic penetration are negatively associated with early-stage drug development, and these effects are statistically significant at conventional levels.[19] The elasticity from Model 5 implies

[17] We thank Jeff Macher of Georgetown University for suggesting this robustness check.
[18] In these regressions, our dependent variable is identical to P_{ijt-1}, so we omit this variable from our set of control variables.
[19] Recall that our unit of observation is at the firm-market-year where market *j* is defined at the ATC2 level. As such, in these models the loss of significance on O_{jt-1} and V_{jt+2} should not be surprising given the ATC2 fixed effect and the (Year*ATC2) interaction fixed effect.

that a 10% increase in generic penetration in a particular market will *lower* early-stage innovations, in that same market, by 4.1%.

5.5 Are generics driving a switch to biologics-based drug development?

Other researchers have conjectured that declining revenues associated with small-molecule (chemical-based) products are increasingly motivating firms to switch to large-molecule (biologic-based) products (Wong, 2009; Golec *et al*, 2010). As we have noted above, such a shift could have mixed consequences for future drug development. Biologics are more expensive than chemical-based products, on average (Aitken *et al*., 2009; Trusheim *et al*., 2010), and final regulatory guidance that provides a pathway for biosimilar entry in the U.S. market has yet to be issued. This means that even after the patents of a branded biologic-based product expire, there is no pathway for generic entry without full clinical trials. If uptake between the two types of products over their entire product lifecycle remains similar, then a shift from chemistry-based to biologic-based drugs could imply that, all else equal, the percent of overall health care expenditures spent on pharmaceuticals will increase.[20]

In order to consider whether a shift to biologic-based products may be occurring as a consequence of rising generic penetration, we estimate the specification described in Equation 2. The dependent variable in this specification is the difference between early-stage chemical-based innovations and early-stage biologic-based innovations. As constructed, this variable can now take on negative values, which prevents us from using count data models. Instead, we create a variable, $cat(CI_{ijt}\text{-}BI_{ijt})$, that equals 1, 2 and 3 if the difference ($CI_{ijt} - BI_{ijt}$) is negative, zero, or positive, respectively, and we estimate Equation 2 using an ordered logit model (Models 1-3, Table 6). Model 1 in Table 6 reports our complete baseline specification with all fixed effects except the (Year*ATC) interaction. Models 2 and 3 include the full set of fixed effects across both measures of generic penetration. For comparative purposes, we also report results from a linear regression on the same dependent variable with our standard set of controls and fixed effects. These results are reported in Models 4 and 5. Across all specifications our measure of generic penetration is negatively and significantly related to the difference in types of early-stage innovations. This suggests that as generic penetration increases, our dependent variable, $cat(CI_{ijt}\text{-}BI_{ijt})$, declines which, in turn, implies that the difference, ($CI_{ijt} - BI_{ijt}$) is decreasing. In other words as generic penetration increases, the flow of biologic-based innovations is greater than the flow of chemical-

[20] Even after final guidance is issued for biosimilars in the U.S., as long as the data exclusivity period remains at 12 years there will still be a significant difference between the regulatory incentives for biologic-based drugs and chemical-based drugs. As of this writing, data exclusivity for chemical-based drugs is 5 years (with additional extensions available for pediatric use, orphan designation and reformulations.)

based innovations for firm i, in market j, at time t. It appears that pharmaceutical firms are responding to generic competition by shifting to biologics, where they do not face similar competitive constraints.

The coefficient on O_{jt-1} remains positive and statistically significant, as in earlier regressions. Our controls for firm capabilities also remain consistent. The difference in chemical-based and biologic-based approved late-stage innovations, ($CP_{ijt-1} - BP_{ijt-1}$), and approved products, ($CD_{ijt-1} - BD_{ijt-1}$), is positive and significant, as expected. In other words, if a firm has more chemical-based products relative to biologic-based products then the flow of chemical-based early-stage innovations relative to biologic-based innovations is greater. Not only do pharmaceutical firms continue to develop products within the same therapeutic category but they also appear to continue to develop products of the same type.

Table 7 provides results from a third specification -- one in which two separate linear models predicting chemical-based product innovations and biologic-based production innovations, respectively, are run as a system, using the seemingly unrelated regressions (SUR) approach. In Model 1, with our primary measure of generic penetration, G_{ijt-1}, we can see that is negatively associated with chemical-based innovation, but positively associated with biologic-based innovation, and both relationships are significant at the conventional threshold levels. We noted earlier in the paper that our sample is limited to firms with at least one approved product and at least one candidate drug in early-stage development. This sampling restriction excludes some small, research-intensive firms. However, these smaller entities are overwhelmingly focused on biologic drug development. We strongly believe their inclusion in our empirical analysis would, if anything, significantly strengthen the general tenor of our findings, especially those concerning the shift out of chemical-based drugs and into biologic-based drugs.

6 Conclusion

For many years, researchers and industry observers have conjectured that rising generic penetration might have an impact on the rate and direction of pharmaceutical innovation. Using a new combination of data sets, we are able to estimate the effects of rising generic penetration on early-stage pharmaceutical innovation. While the overall level of early-stage drug development has continued to increase, generics have had a statistically and economically significant impact on *where* that development activity is concentrated and *how* it is done. In the full sample, we find that, as our baseline measure of generic penetration increases by 10% within a therapeutic market, we observe a decrease of 7.9% in early-stage innovation in that market. This implies that drug development activity is moving out of markets where generic competition is increasing and into domains where it is relatively less intense.

Our preferred interpretation of this relationship, namely that a rise in generic penetration leads to a decline in drug development in that market, is strengthened by the finding that this relationship varies across therapeutic areas in ways that conform to our prior expectations. In earlier work (Branstetter *et al*., 2013), we pointed out that the degree of substitution between generics and branded products can vary substantially across therapeutic areas. In markets where the substitution possibilities between generics and branded drugs are more limited, changes in generic penetration could be expected to have a weaker impact on innovation. This is indeed what we observe when we focus on three markets containing drugs that treat neurological and psychiatric disorders, where clinicians are sometimes reluctant to move away from a good "match" between a patient and a drug, even when a cheaper generic alternative (of a different drug) becomes available. In these markets, we find no statistically significant effect of generics on early-stage innovations. However, in markets with high levels of cross-molecular substitution we see the opposite.

In a similar manner, we would expect the measured negative correlation between rising generic penetration and new drug development to be strong and significant for early-stage drug development, where it is still feasible to redirect research efforts, but much weaker in late-stage drug development, where candidate drugs have already proved their safety and efficacy in a series of increasingly expensive and stringent clinical trials and are generally introduced even if the market is known to be limited by increasing generic competition. We find exactly this pattern in the data, providing further support for our preferred interpretation of the statistical relationship. The robustness of our results is also confirmed when we limit our sample to drugs candidates designated as "novel" by our Pharmaprojects database. These drugs are first in their class. This shows that our results are not driven by generic competition simply pushing out "me-to" drugs or reformulation/recombinations of existing therapies. For better or worse, the rise in generic penetration is associated with a decline in novel drug development. The elasticity from our results implies that a 10% increase in generic penetration in a particular market will lower early-stage novel drug development, in that same market, by 4.6%.

We also note that, in a linear specification, the negative relationship between drug development and rising generic penetration is robust to the inclusion of a full set of ATC 2-digit dummy variables and the interaction of these dummy variables with our year dummy variables. In this specification, where all the unobserved factors impacting a 2-digit therapeutic area over time in a common way across firms are effectively removed, the key empirical relationship remains negative, strong, and statistically robust, regardless of how we measure it. The elasticity from our results implies that a 10% increase in generic penetration in a particular market will lower early-stage innovation, in that same market, by 4.1%.

Finally, we also consider the economic incentives created by regulation to shift, within therapeutic markets, from chemical-based to biologic-based products. Currently, data exclusivity is much longer for biologic-based products, and the regulatory pathway to market for biosimilars has yet to be finalized. We conjecture that as chemical-based products are pressured by generics, pharmaceutical firms will begin to change the nature of their innovation by shifting to biologics. This is indeed what we observe. Increases in generic penetration in market *j* appear to lead to an increase in the relative amount of biologic-based drug development. As generic penetration in market *j* rises, firms do not appear to be abandoning market *j* completely, but rather changing the *nature* of the innovation they pursue.

We have shown that the rise of generic competition is reshaping the locus of drug development activity. Is this a good thing? In this paper, we have refrained from taking a strong stand on the welfare impact of this shift. The data we would need to determine this are not yet available, and, at this point, we can only speculate on the sign of the ultimate welfare impact. On the positive side, one can argue that social welfare is enhanced when pharmaceutical firms are induced to shift development efforts away from markets where a broad range of effective and cheap generic therapies already exist to ones with far fewer treatment options. This can be true even if the probabilities of research success are lower in the domains into which research effort is being pushed, because the social returns to expanding the range of treatment options is so relatively high. Even an increasing shift to more expensive biologic-based drugs may be beneficial in the long run if much of the potential for further advance in small-molecule drugs has already been exhausted. However, it is equally easy - and for us, equally plausible - to imagine a less positive outcome. Rising generic competition could be eliminating the development of new drugs that have all the benefits of existing therapies without the side effects. Such new drugs would have social value, even in markets with an extensive range of existing therapies. The less explored domains into which the pharmaceutical industry's small-molecule developments are being pushed may yield little or no success. Such pessimism would be consistent with much of the discussion of the pharmaceutical industry's longstanding "productivity crisis." Finally, by tilting the regulatory playing field so heavily against small-molecule drug development and in favor of biologics, we may be inducing the global industry to give up on the former domain that has done so much to advance global health through the provision of cheap, relatively simple, effective drugs long before the potential benefits of further research have been exhausted.[21]

[21] In fact, many industry insiders believe that there are hundreds of small molecule compounds with as yet undiscovered therapeutic benefits. Because the patents on these compounds expired long ago, there is no mechanism by which a branded pharmaceutical company could appropriate the returns from R&D into these new therapeutic benefits. This line of thinking raises the possibility that there is a gold mine of potentially high-return research

The first step toward a more definitive conclusion about the welfare impact of the shift in drug development would be the creation of a map that locates the various therapeutic categories in terms of their proximity in technology space. It is reasonable that firms pressed by rising generic competition would seek to redeploy their R&D resources in domains that are not wholly dissimilar from the ones in which they have been working. Despite decades of high-quality empirical research on the pharmaceutical industry, no researchers have yet created such a mapping. With such data at hand, we could begin to explore not just the declines in drug development that have been induced by generic competition, but the increases in development in technologically proximate markets. These data would also facilitate the comparison of research success probabilities in the domains where drug development effort is declining and ones in which it is increasing.

Even with such data at hand, assessment of the full welfare impact of the recent shift will require strong assumptions that allow us to sketch out the counterfactual distribution of research effort that would have existed in the absence of the recent rise in generic competition. Despite this, we believe the effort is not just worthwhile, but necessary. Whether the effect was intended or not, the rise of generics in the U.S. market is significantly reshaping the pattern of global drug development efforts. We need to know if this is pushing that pattern closer to or further away from the social optimum. As is usually the case in economic research, much remains to be done.

---

projects that are currently inaccessible to the global pharmaceutical industry. Meanwhile, the existing regulatory regime induces them to spend billions on extremely complex, large-molecule therapies whose full interaction with the human body is imperfectly understood, and where the rate of failure in clinical trials is correspondingly high. See Higgins *et al*. (2014) and Roin (2014) for an explication of this idea and potential policy solutions.

References

Acemoglu, D. and Linn, J., 2004. Market size in innovation: Theory and evidence from the pharmaceutical industry. The Quarterly Journal of Economics 119 (3), 1049-1090.

Aitken, M., Berndt, E., Cutler, D., 2009. Prescription drug spending trends in the U.S.: Looking beyond the turning point. Health Affairs 28(1), 151-160.

Arcidiacono, P., Ellickson, P. B., Landry, P., Ridley, D. B., 2013. Pharmaceutical followers. International Journal of Industrial Organization *31*(5), 538-553.

Berndt, E., Mortimer, R., Bhattacharjya, A., Parece, A., Tuttle, E., 2007. Authorized generic drugs, price competition and consumers' welfare. Health Affairs 26(3), 790-799.

Berndt, E., Aitken, M., 2010. Brand loyalty, generic entry, and price competition in pharmaceuticals in the quarter century after the 1984 Hatch-Waxman legislation. NBER Working Paper No. 16431.

Bokhari, F. A., Fournier, G. M. 2013. Entry in the ADHD drugs market: Welfare impact of generics and me-too's. The Journal of Industrial Economics 61(2), 339-392.

Branstetter, L., Chatterjee, C., Higgins, M. 2013. Regulation and welfare: Evidence from Paragraph IV generic entry in the pharmaceutical industry. Revision of NBER working paper 17188.

Budish, E., Roin, B., Williams, H. 2013. Do fixed patent terms distort innovation? Evidence from cancer clinical trials. NBER Working Paper 19340.

Caves, R., Whinston, M., Hurwitz, M., 1991. Patent expiration, entry and competition in the U.S. pharmaceutical industry. Brookings Papers on Economic Activity: Microeconomics, 1-48.

Ceccagnoli, M., Graham, J., S., Higgins, J., M., Lee, J. 2010. Productivity and the role of complementary assets in firms' demand for technology innovations. Industrial and Corporate Change, 19(3), 839-869.

Chan, T., Nickerson, A., J, Owan, H. 2007. Strategic management of R&D pipelines with cospecialized investments and technology markets. Management Science 53, 667–682.

Chiou, J., Magazzini, L., Pammoli, F., Riccaboni, M., 2012.The value of failure in pharmaceutical R&D. Unpublished working paper, IMT Lucca EIC Working Paper Series 01.

Cleanthous, P., 2002. Patient welfare implications of innovation in the U.S. antidepressant market. Unpublished working paper, New York University.

Cockburn, I. M. 2006. Is the pharmaceutical industry in a productivity crisis? in Innovation Policy and the Economy, Jaffe, A. B., Lerner, J. & Stern, S. (eds), 7(1), 1–32, The MIT Press.

Cohen, W., 2010. Fifty Years of Empirical Studies of Innovative Activity and Performance. In Brownyn H. Hall and Nathan Rosenberg (ed.), Handbook in Economics of Innovation, North-Holland.

Cohen, W., Nelson, R., Walsh, J., 2002. Links and impacts: The influence of public research on industrial R&D. Management Science 48(1), 1 – 23.

Congressional Budget Office, 1998. How increased competition from generic drugs has affected process and returns in the pharmaceutical industry. Congressional Budget Office, The Congress of the United Sates, http://www.cbo.gov.

DiMasi, J., Hansen, R., Grabowski, H., 1991. Cost of innovation in the pharmaceutical industry. Journal of Health Economics 10(2), 151-185.

DiMasi, J., Hansen, R., Grabowski, H., 2003. The price of innovation: new estimation of drug development costs. Journal of Health Economics 22, 151-185.

DiMasi, J. A., Grabowski, H. G., 2012. R&D costs and returns to new drug development: a review of the evidence. Tufts Center for the Study of Drug Development. March/April CSDD Impact Report. The Oxford handbook of the economics of the biopharmaceutical industry. Oxford University Press, Oxford, 21-46.

Dunn, A. 2012. Drug innovations and welfare measures computed from market demand: The case of anti-cholesterol drugs. American Economic Journal: Applied Economics 4(3), 167-189.

Ellickson, P., Stern, S., Trajtenberg, M. 2001. Patient welfare and patient compliance -An empirical framework for measuring the benefits from pharmaceutical innovation. In Medical Care Output and Productivity (pp. 539-564). University of Chicago Press.

Federal Trade Commission, 2002. Generic entry prior to patent expiration: FTC study. U.S. Government Printing Office, Washington, D.C., July.

Fuchs, Victor, 1982. Who Shall Live? Health, Economics, and Social Choice. Wiley, New York.

Furman, J.L., Kyle, K., M. Cockburn, I. Henderson, M., R. 2006. Public and private spillovers, location, and the productivity of pharmaceutical research. NBER Working Paper Number 12509.

General Accounting Office, 2012. Research on Savings from Generic Drug Use. GAO-12-371R.http://www.gao.gov/assets/590/588064.pdf

Gittelman, M., Kogut, B., 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. Management Science 49 (4), 366–382.

Golec, J., Hegde, S., Vernon., M. J., 2010.Pharmaceutical R&D spending and threats of price regulation. Journal of Financial and Quantitative Analysis 45, 239-264.

Grabowski, H., 2007. Competition between generic and branded drugs. In Pharmaceutical Innovation: Incentives, Competition, and Cost-Benefit Analysis in International Perspective, Sloan, F., Hsieh, C. (eds.), Cambridge University Press: Cambridge, MA.

Grabowski, H., Kyle, M., 2007. Generic competition and market exclusivity periods in pharmaceuticals. Managerial and Decision Economics 28, 491-502.

Hausman, J., Hall, H. B., Griliches, Z. 1984. Econometric models for count data with an application to the patents-R&D relationship. Econometrica 52, 909-938.

Hausman, J., Lo, A., MacKinlay, C. 1992. An ordered probit analysis of transaction stock prices. Journal of Financial Economics 34, 83-104.

Hemphill, S., Sampat, B., 2011. When do generics challenge drug patents. Journal of Empirical Legal Studies 8(4), 619-649.

Hemphill, C., Sampat, B., 2012. Evergreening, patent challenges, and effective market life in pharmaceuticals. Journal of Health Economics, 327-341.

Higgins, M.J., Graham, S., 2009. Balancing innovation and access: patent challenges tip the scales. Science 326, 370-372.

Higgins, M.J., Thursby, J., Thursby, M., 2014. Bench-to-bench bottlenecks in translation. Science Translational Medicine 6, 250fs32.

Hughes, J., Moore, M., Snyder, E., 2002. Napsterizing pharmaceuticals: Access, innovation and consumer welfare. NBER Working Paper No. 9229.

Knowles, S., 2010. Fixing the legal framework for pharmaceutical research. Science 327, 1083-1084.

Lichtenberg, F. R. 2001. The benefits and costs of newer drugs: Evidence from the 1996 medical expenditure panel survey. NBER Working paper 8147.

Lichtenberg, F. R. 2004. Public policy and innovation in the U.S. pharmaceutical industry. In: Holtz-Eakin, D., Rosen, H. (eds), Entrepreneurship and Public Policy. MIT Press, Cambridge, MA.

Lichtenberg, F. R. 2007. The impact of new drug launches on U.S. longevity and medical expenditure. American Economic Review 97, 438-441.

Mansfield, E. 1995. Academic research underlying industrial innovation: Sources, characteristics and financing. Review of Economics and Statistics 77, 55-65.

Mossinghoff, G., 1999. Overview of Hatch-Waxman act and its impact on the drug development process. Food and Drug Law Journal 54(2), 187-194.

New York Times, 1989. The Generic Drug Scandal. October 2, 1989.

Nordhaus, D., William. 1999. The health of nations: The contribution of improved health to living standards. NBER working paper 8818, accessed at http://nordhaus.econ.yale.edu/health_nber_w8818.pdf

Pammolli, F., Magazzini, L., Riccaboni, M., 2011.The productivity crisis in pharmaceutical R&D. Nature Reviews Drug Discovery 10, 428-438.

Panattoni, L., 2011. The effect of Paragraph IV decisions and generic entry before patent expiration on brand pharmaceutical firms. Journal of Health Economics 30, 126-145.

Reiffen, D., Ward, M., 2005. Generic drug industry dynamics. Review of Economics and Statistics 87, 37-49.

Roin, B., 2014. Solving the problem of new uses. Michigan State Law Review, *forthcoming*. http://ssrn.com/abstract=2337821.

Saha, A., Grabowski, H., Birnbaum, H., Greenberg, P., Bizen, O., 2006. Generic competition in the US pharmaceutical industry. International Journal of the Economics of Business 13, 15-38

Scherer, F. M., 2010. Pharmaceutical Innovation. In Brownyn H. Hall and Nathan Rosenberg (ed.), Handbook in Economics of Innovation, North-Holland.

Teece, D.J. 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy, Research Policy 15, 286-305.

Trusheim, M., Aitken, M., Berndt, E., 2010. Characterizing markets for biopharmaceutical innovations: Do biologics differ from small molecules? Forum for Health Economics and Policy 13(1), 1-43.

Wong, G., 2009. Biotech scientists bank on big pharma's biologics push. Nature Biotechnology 27, 293-295.

Wooldridge, J., 1999. Distribution-free estimation of some nonlinear panel data models. Journal of Econometrics 90(1), 77-97.

FIGURE 1. EARLY-STAGE INNOVATIONS, 1998-2010. This figure tracks the aggregate flow of early-stage pharmaceutical innovations, defined as the annual count of compounds at the preclinical stage or in Phase 1 clinical trials. We provide annual aggregate counts for our sample firms (solid line) and for the entire population (dotted line) of compounds contained in the Pharmaprojects database. Over our time period, 1998-2010, the number of early-stage innovations, including both small- and large-molecules, has increased. Our sample closely tracks the population, with differences being explained by our sample restrictions. Recall, firms must have at least one approved product and one early-stage innovation in order to be incorporated into our sample.

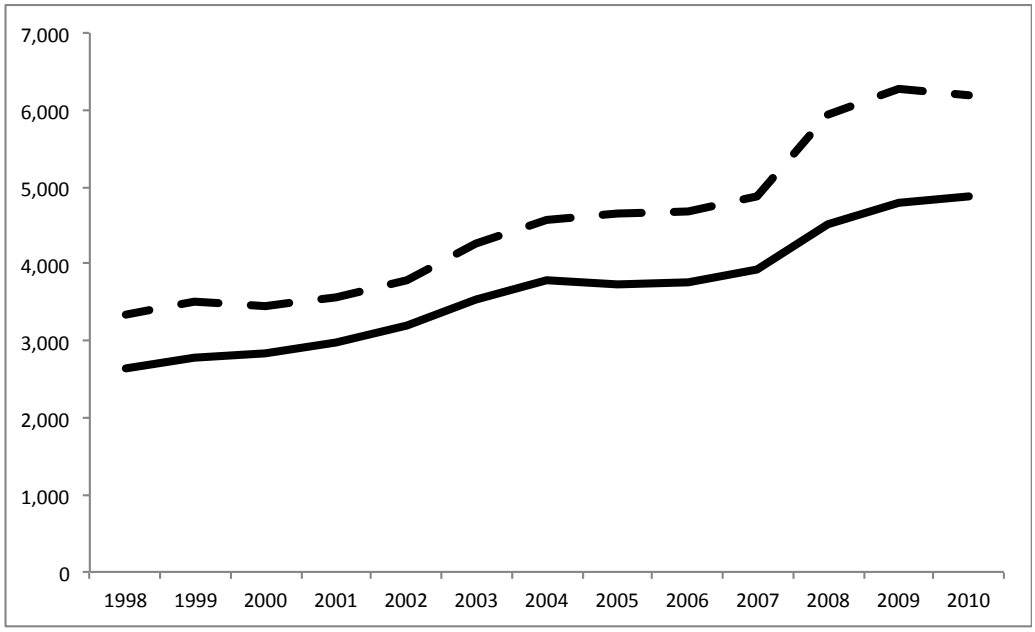

FIGURE 2. EXCLUSIVITIES AND INNOVATION IN PHARMACEUTICALS. This figure illustrates the two types of protection conferred on new drugs. When a new drug is approved by the FDA, the first five year period (seven years for orphan drugs and 5 ½ years for pediatric drugs) carries with it a regulatory protection called 'data exclusivity' that runs concurrent with underlying patent protection. Data exclusivity protects the underlying clinical data. The FDA will approve no generics during data exclusivity, regardless of whether the innovator has patents in force. At the conclusion of data exclusivity a drug is protected only by its patents until they expire or are successfully challenged, a period termed 'market exclusivity'. Para-IV challenges occur only during the market exclusivity period. Note that patents are generally applied for and granted well before drugs are approved by the FDA.

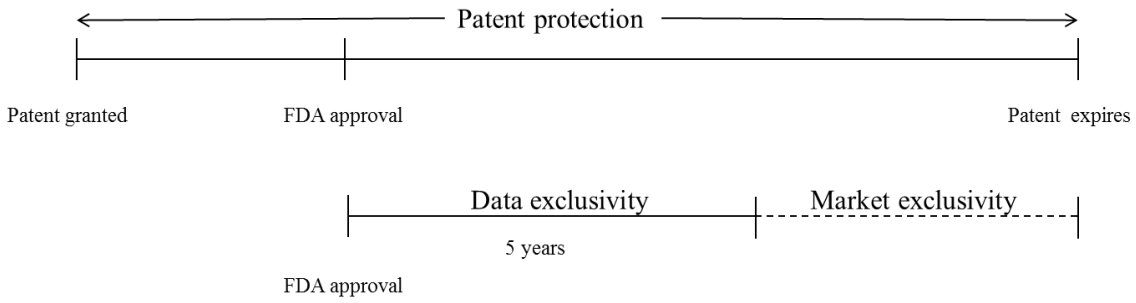

FIGURE 3. ANDA PATENT CERTIFICATION OPTIONS FOR GENERIC MANUFACTURERS.
Generic entry in the U.S. market can occur in through one of four regulatory pathways, each denoted by the number of the paragraph in the Hatch-Waxman Act that defines them. Paragraph I, Paragraph II, and Paragraph III are used by generic manufacturers for drugs whose patents are either not listed in the FDA Orange Book or for those patents that have expired (or will expire). Paragraph IV is the only pathway that facilitates generic entry before the expiry of patents. Source: FTC (2002).

Certification	Requirement	Outcome
Paragraph I:	Required patent information has not been filed	FDA may approve ANDA immediately; one or more generic applicant may enter
Paragraph II:	Patent has expired	FDA may approve ANDA immediately; one or more generic applicant may enter
Paragraph III:	Patent has not yet expired but will expire on a particular day	FDA may approve ANDA effective on the date the patent expires; one or more generic applicant may enter
Paragraph IV:	Patent is invalid or is non-infringed by generic applicant	Generic applicant provides notice to the patent holder and NDA filer. Entry by the first-filer may or may not occur.

FIGURE 4. RELATIVE CONTRIBUTION TO TOTAL INNOVATIONS ACROSS THERAPEUTIC CATEGORIES

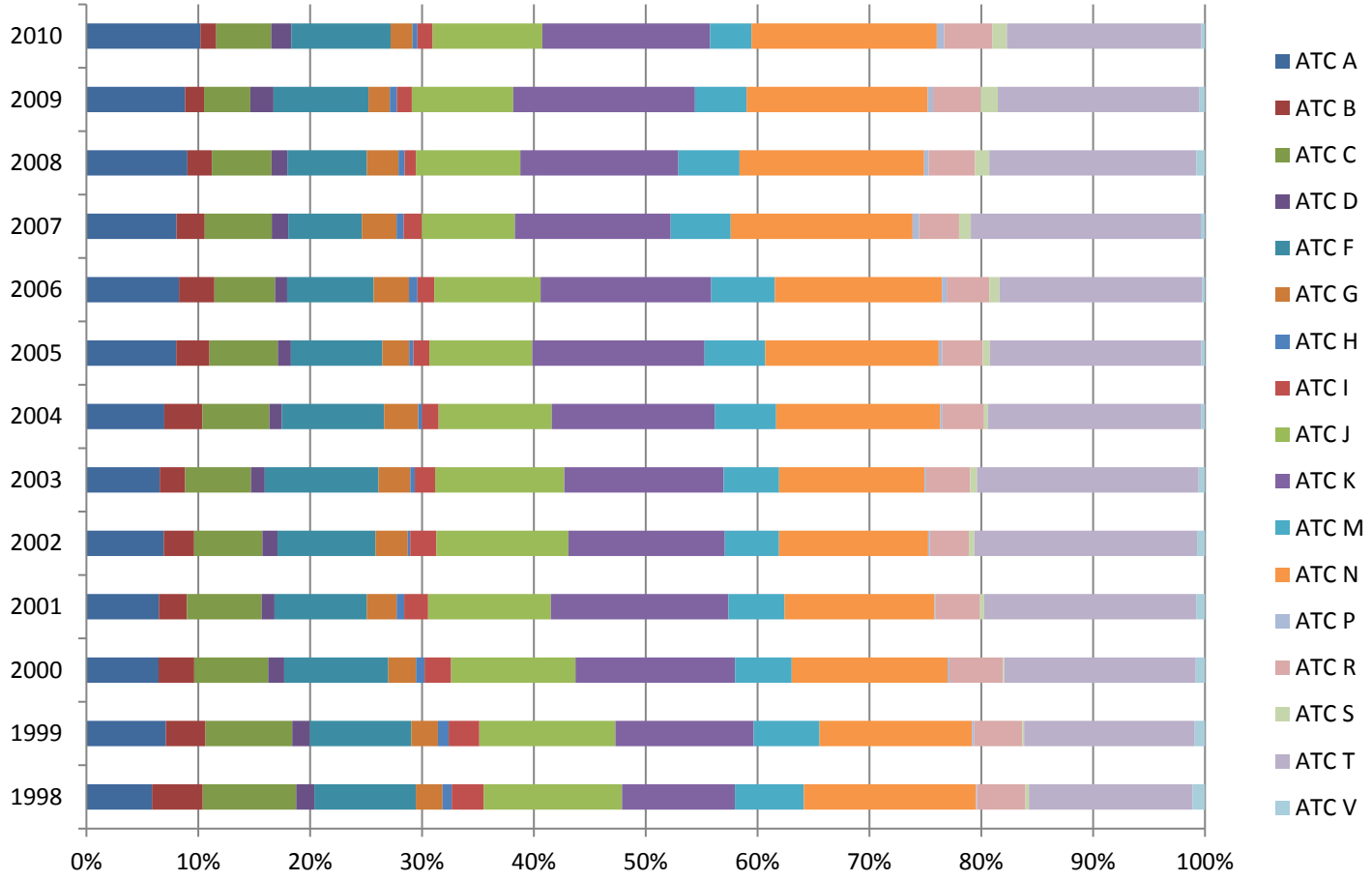

TABLE 1. VARIABLE DEFINITION AND DESCRIPTIVE STATISTICS

VARIABLES	DEFINITION	SOURCE	OBS	MEAN	S. DEV.	MIN	MAX
I_{ijt}	Early stage innovations: Count of early stage pipeline (pre-clinical + phase 1) at *i, j, t* level.	Pharmaprojects	29,486	0.78	1.81	0	36
G_{ijt-1}	Generic penetration, baseline measure: Ratio of generic sales to sum of focal firm and generic sales at *i, j, t-1* level.	IMS MIDAS™	29,486	0.54	.46	0	1
$G_{jt-1}M_{ijt-1}$	Generic penetration alternative: Ratio of generic sales to total sales at *j, t-1* level multiplied by the ratio of firm *i* sales in *j* at time *t-1* to total firm *i* sales in all markets at time *t-1*.	IMS MIDAS™	29,486	0.54	.46	0	1
V_{jt+2}	Expected market value: Logarithm of total sales in market *j* averaged over years *t*, *t+1*, and *t+2*.	IMS MIDAS™	29,486	13.05	3.82	0	16.79
O_{jt-1}	Technological opportunity: Logarithm of stock of citation-weighted articles in year *t-1* for j^{th} therapeutic market.	IMS NDTI™& MIDAS™, PubMed and SCOPUSSciverse	29,486	8.09	7.30	0	17.9
Z_{ijt-1}	Technological challenges: Counts of suspended or discontinued pipeline projects and withdrawn approved products at *i, j, t-1* level.	Pharmaprojects	29,486	0.05	0.26	0	6
D_{ijt-1}	Firm innovative capability I: Moving average of product introductions in *t-1, t-2, t-3* at the *i, j, t-1* level.	Pharmaprojects	29,486	0.24	1.01	0	25.67
P_{ijt-1}	Firm innovative capability II: Count of Phase II and Phase III products at the *i, j, t-1* level.	Pharmaprojects	29,486	0.09	0.35	0	6
SA_{ijt}	Downstream co-specialized assets: Ratio of promotions at the *i,j, t* level and total pharmaceutical sales at the *i, j, t* level.	IMS MIDAS™	29,486	0.45	19.36	0	2225
S_{it}	Firm size: Logarithm of total pharmaceutical sales at the *i, t* level.	IMS MIDAS™	29,486	12.64	4.45	0	17.23

TABLE 2. FLOW OF INNOVATION: POISSON REGRESSION

	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5	MODEL 6
	I_{ijt}	I_{ijt}	I_{ijt}	I_{ijt}	I_{ijt}	I_{ijt}
G_{ijt-1}			-1.638***	-1.368***	-1.354***	
			(0.116)	(0.135)	(0.135)	
$G_{jt-1}M_{ijt-1}$						-5.446***
						(1.199)
O_{jt-1}		0.023***	0.005	0.045***	0.046***	0.026***
		(0.005)	(0.004)	(0.004)	(0.004)	(0.005)
Z_{ijt-1}		0.383***	0.447***	0.355***	0.357***	0.388***
		(0.031)	(0.029)	(0.026)	(0.024)	(0.027)
V_{jt+2}	-0.046***	-0.074***	0.007***	-0.071***	-0.074***	-0.083***
	(0.012)	(0.011)	(0.005)	(0.011)	(0.011)	(0.012)
D_{ijt-1}	0.110***	0.123***	0.088***	0.110***	0.114***	0.128***
	(0.010)	(0.0106)	(0.013)	(0.009)	(0.008)	(0.009)
P_{ijt-1}	0.244***	0.134***	0.241***	0.127***	0.136***	0.141***
	(0.044)	(0.036)	(0.050)	(0.033)	(0.035)	(0.037)
SA_{ijt}	-0.002	-0.002	-0.001	-0.001	-0.001	-0.003
	(0.003)	(0.003)	(0.002)	(0.001)	(0.001)	(0.003)
S_{it}	0.010*	0.012**	0.019**	0.020***	0.020**	0.013**
	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)	(0.006)
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
ATC1 FE	Y	Y	N	Y	Y	Y
(Year*ATC1) FE	N	N	N	N	Y	Y
N	29,486	29,486	29,486	29,486	29,486	29,486

Notes: Coefficients and clustered standard errors (by firm) are reported. ATC is defined as anatomical therapeutic chemical classification. *** p<0.01, ** p<0.05, * p<0.10

TABLE 3. FLOW OF INNOVATION: NEGATIVE BINOMIAL REGRESSION

	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5	MODEL 6
	I_{ijt}	I_{ijt}	I_{ijt}	I_{ijt}	I_{ijt}	I_{ijt}
G_{ijt-1}			-1.745***	-1.471***	-1.458***	
			(0.104)	(0.113)	(0.113)	
$G_{jt-1}M_{ijt-1}$						-6.422***
						(1.076)
O_{jt-1}		0.017***	0.015***	0.044***	0.044***	0.019***
		(0.005)	(0.004)	(0.004)	(0.004)	(0.005)
Z_{ijt-1}		0.649***	0.674***	0.568***	0.566***	0.644***
		(0.032)	(0.038)	(0.031)	(0.031)	(0.032)
V_{jt+2}	-0.026**	-0.057***	-0.002	-0.054***	-0.056***	-0.065***
	(0.011)	(0.011)	(0.005)	(0.010)	(0.010)	(0.012)
D_{iit-1}	0.229***	0.213***	0.207***	0.179***	0.178***	0.214***
	(0.032)	(0.030)	(0.036)	(0.025)	(0.025)	(0.029)
P_{ijt-1}	0.200***	0.122***	0.276***	0.101**	0.102***	0.125***
	(0.041)	(0.042)	(0.055)	(0.039)	(0.040)	(0.043)
SA_{iit}	-0.005	-0.004	-0.001	-0.002	-0.002	-0.005
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)
S_{it}	0.001	0.005	0.011	0.014*	0.014*	0.007
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.007)
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
ATC1 FE	Y	Y	N	Y	Y	Y
(Year*ATC1) FE	N	N	N	N	Y	Y
N	29,486	29,486	29,486	29,486	29,486	29,486

Notes: Coefficients and clustered standard errors (by firm) are reported. ATC is defined as anatomical therapeutic chemical classification. *** p<0.01, ** p<0.05, * p<0.10

TABLE 4. FLOW OF INNOVATION, LOW CMS MARKETS: NEGATIVE BINOMIAL REGRESSIONS

	MODEL 1 I_{ijt}	MODEL 2 I_{ijt}	MODEL 3 I_{ijt}	MODEL 4 I_{ijt}	MODEL 5 I_{ijt}
G_{ijt-1}			-0.146 (0.271)		
$G_{jt-1}M_{ijt-1}$				-4.967 (8.108)	2.619 (3.523)
O_{jt-1}		0.376** (0.145)	0.383*** (0.147)	0.370*** (0.131)	0.370*** (0.145)
Z_{ijt-1}		0.266** (0.075)	0.265*** (0.076)	0.575*** (0.114)	0.263*** (0.077)
V_{jt+2}	0.880*** (0.201)	-0.217 (0.280)	-0.222 (0.283)	-0.056 (0.261)	-0.166 (0.279)
D_{ijt-1}	0.169*** (0.029)	0.132*** (0.03)	0.125*** (0.028)	0.443*** (0.124)	0.127*** (0.031)
P_{ijt-1}	-0.009 (0.108)	-0.056 (0.102)	-0.046 (0.101)	0.209 (0.174)	-0.064 (0.106)
SA_{ijt}	-0.279 (0.567)	-0.107 (0.513)	-0.025 (0.503)	-0.245 (0.265)	-0.108 (0.511)
S_{it}	0.002 (0.028)	0.017 (0.025)	0.022 (0.027)	0.052* (0.031)	0.014 (0.023)
Firm FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	N	Y
N	1,577	1,577	1,577	1,577	1,577

Notes: Coefficients and clustered standard errors (by firm) are reported. Low cross-molecular substitution (CMS) markets include three neurology markets: N3 (anti-epileptics), N5 (psycholeptics) and N6 (psychoanaleptics). Since these three sub-markets are included in ATC N, we exclude 1-digit ATC fixed effects from these specifications. *** p<0.01, ** p<0.05, * p<0.10

TABLE 5. FLOW OF INNOVATION: ROBUSTNESS CHECKS

	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5	MODEL 6
	NEG BN	NEG BN	NEG BN	NEG BN	OLS	OLS
	Novel I_{ijt}	*Novel* I_{ijt}	*Late Stage* I_{ijt}	*Late Stage* I_{ijt}	I_{ijt}	I_{ijt}
G_{ijt-1}	-0.821***		0.175		-0.583***	
	(0.023)		(0.176)		(0.071)	
$G_{jt-1}M_{ijt-1}$		-6.222***		0.497		-0.909***
		(2.034)		(2.685)		(0.267)
O_{jt-1}	0.048***	0.036***	-0.001	0.011		
	(0.011)	(0.013)	(0.010)	(0.010)		
Z_{ijt-1}	0.355***	0.406***	0.319***	0.476***	1.505***	1.532***
	(0.055)	(0.056)	(0.029)	(0.035)	(0.179)	(0.174)
V_{jt+2}	-0.168***	-0.185***	-0.004	-0.0155		
	(0.032)	(0.043)	(0.016)	(0.021)		
D_{ijt-1}	0.097***	0.120***	0.120***	0.195***	0.364***	0.371***
	(0.018)	(0.021)	(0.020)	(0.027)	(0.042)	(0.042)
P_{ijt-1}	0.036	0.049			0.478***	0.501***
	(0.071)	(0.010)			(0.126)	(0.125)
SA_{ijt}	0.001	-0.002	0.013	-0.015**	-0.001	-0.001
	(0.002)	(0.003)	(0.016)	(0.007)	(0.001)	(0.001)
S_{it}	0.004	-0.004	-0.012	-0.016	0.009	0.010
	(0.010)	(0.009)	(0.009)	(0.011)	(0.011)	(0.010)
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
ATC1 FE	Y	Y	Y	Y	N	N
ATC2 FE	N	N	N	N	Y	Y
(Year*ATC1) FE	Y	Y	Y	Y	N	N
(Year*ATC2) FE	N	N	N	N	Y	Y
N	29,486	29,486	29,486	29,486	29,486	29,486

Notes: Models 1 – 4 report results from negative binomial fixed effects regressions and include 1-digit ATC fixed effects. Models 5 and 6 report results from OLS regressions with 2-digit ATC fixed effects and year*ATC2 interaction terms (Y*). Coefficients and clustered standard errors (by firm) are reported. ATC is defined as anatomical therapeutic chemical classification. *** p<0.01, ** p<0.05, * p<0.10

TABLE 6. SHIFT FROM CHEMICAL-BASED TO BIOLOGIC-BASED DRUGS

	MODEL 1	**MODEL 2**	**MODEL 3**	**MODEL 4**	**MODEL 5**
	OLOGIT	OLOGIT	OLOGIT	OLS	OLS
	$cat(CI_{ijt}\text{-}BI_{ijt})$	$cat(CI_{ijt}\text{-}BI_{ijt})$	$cat(CI_{ijt}\text{-}BI_{ijt})$	$cat(CI_{ijt}\text{-}BI_{ijt})$	$cat(CI_{ijt}\text{-}BI_{ijt})$
G_{ijt-1}	-1.818***	-1.824***		-0.332***	
	(0.048)	(0.048)		(0.021)	
$G_{jt-1}M_{ijt-1}$			-3.880***		-0.713***
			(0.294)		(0.117)
O_{jt-1}	0.029***	0.032***	0.003	0.006***	0.002
	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)
diff(Z_{ijt-1})	2.426***	2.445***	2.50***	0.174***	0.194***
	(0.203)	(0.204)	(0.193)	(0.029)	(0.031)
V_{jt+2}	-0.056***	-0.065***	-0.087***	-0.013***	-0.017***
	(0.014)	(0.014)	(0.014)	(0.004)	(0.004)
diff(D_{ijt-1})	0.778***	0.781***	0.787***	0.063***	0.064***
	(0.056)	(0.057)	(0.057)	(0.010)	(0.011)
diff(P_{ijt-1})	1.871***	1.877***	2.006***	0.209***	0.226***
	(0.065)	(0.066)	(0.061)	(0.018)	(0.019)
diff(SA_{ijt})	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
S_{it}	0.020*	0.023**	0.022**	0.002	0.002
	(0.011)	(0.011)	(0.011)	(0.002)	(0.003)
Firm FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
ATC1 FE	Y	Y	Y	Y	Y
(Year*ATC1) FE	N	Y	Y	Y	Y
N	29,486	29,486	29,486	29,486	29,486

Notes: Models 1-3 report results from ordered logit regressions. Models 4 and 5 report results from OLS regressions. Coefficients and clustered standard errors (by firm) are reported. ATC is defined as anatomical therapeutic chemical classification. *** p<0.01, ** p<0.05, * p<0.10.

TABLE 7. SHIFT FROM CHEMICAL-BASED TO BIOLOGIC-BASED DRUGS: SUR REGRESSION

MODEL 1

cat(CI_{ijt}-BI_{ijt})

	CI_{ijt}	BI_{ijt}
G_{ijt-1}	-0.587***	0.042***
	(0.021)	(0.017)
O_{jt-1}	0.023***	-0.001
	(0.001)	(0.001)
Z_{ijt-1}	1.307***	0.404***
	(0.028)	(0.023)
V_{jt+2}	-0.046***	-0.023***
	(0.005)	(0.004)
D_{ijt-1}	0.209***	0.169***
	(0.007)	(0.006)
P_{ijt-1}	0.130***	0.494***
	(0.021)	(0.018)
SA_{ijt}	-0.001	0.001
	(0.001)	(0.001)
S_{it}	0.008**	0.001
	(0.004)	(0.003)
Firm FE	Y	Y
Year FE	Y	Y
ATC1 FE	Y	Y
(Year*ATC1) FE	Y	Y
N	29,486	29,486

Notes: Coefficients and clustered standard errors (by firm) are reported. ATC is defined as anatomical therapeutic chemical classification. *** p<0.01, ** p<0.05, * p<0.1

APPENDIX TABLE A1. FLOW OF INNOVATION: POISSON REGRESSION

	MODEL 1 I_{ijt}	MODEL 2 I_{ijt}	MODEL 3 I_{ijt}	MODEL 4 I_{ijt}	MODEL 5 I_{ijt}	MODEL 6 I_{ijt}
G_{ijt-1}			-1.638***	-1.368***	-1.354***	
			(0.304)	(0.251)	(0.249)	
$G_{jt-1}M_{ijt-1}$						-5.446***
						(1.283)
O_{jt-1}		0.0234	0.00495	0.0453***	0.0456***	0.0256
		(0.0174)	(0.0171)	(0.0131)	(0.0132)	(0.0170)
Z_{ijt-1}		0.383***	0.447***	0.355***	0.357***	0.388***
		(0.104)	(0.0890)	(0.0882)	(0.0893)	(0.106)
D_{ijt-1}	0.110***	0.123***	0.0880***	0.110***	0.114***	0.128***
	(0.0209)	(0.0175)	(0.0223)	(0.0163)	(0.0175)	(0.0198)
P_{ijt-1}	0.244***	0.134*	0.241**	0.127*	0.136*	0.141**
	(0.0767)	(0.0720)	(0.0978)	(0.0721)	(0.0720)	(0.0702)
SA_{ijt}	-0.00239	-0.00239	-0.000699	-0.000624	-0.000572	-0.00256
	(0.00433)	(0.00422)	(0.00211)	(0.00165)	(0.00167)	(0.00468)
S_{it}	0.0101**	0.0124**	0.0192**	0.0202***	0.0199***	0.0134**
	(0.00399)	(0.00517)	(0.00753)	(0.00616)	(0.00622)	(0.00560)
V_{jt+2}	-0.0455	-0.0740	0.00734	-0.0710	-0.0744	-0.0829
	(0.0604)	(0.0618)	(0.0242)	(0.0569)	(0.0585)	(0.0632)
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
ATC1 FE	Y	Y	N	Y	Y	Y
(ATC1*Year) FE	N	N	N	N	Y	Y
N	29,486	29,486	29,486	29,486	29,486	29,486

Notes: Coefficients and clustered standard errors (by ATC) are reported. ATC is defined as anatomical therapeutic chemical classification. *** p<0.01, ** p<0.05, * p<0.10

APPENDIX TABLE A2. FLOW OF INNOVATION: NEGATIVE BINOMIAL REGRESSION

	MODEL 1 I_{ijt}	MODEL 2 I_{ijt}	MODEL 3 I_{ijt}	MODEL 4 I_{ijt}	MODEL 5 I_{ijt}	MODEL 6 I_{ijt}
G_{ijt-1}			-1.745***	-1.471***	-1.458***	
			(0.265)	(0.230)	(0.230)	
$G_{jt-1}M_{ijt-1}$						-6.422***
						(1.299)
O_{jt-1}		0.0167	0.0146	0.0437***	0.0436***	0.0188
		(0.0158)	(0.0150)	(0.0116)	(0.0119)	(0.0156)
Z_{ijt-1}		0.649***	0.674***	0.568***	0.566***	0.644***
		(0.0976)	(0.0733)	(0.0786)	(0.0796)	(0.0966)
D_{ijt-1}	0.229***	0.213***	0.207***	0.179***	0.178***	0.214***
	(0.0568)	(0.0526)	(0.0583)	(0.0433)	(0.0425)	(0.0523)
P_{ijt-1}	0.200*	0.122	0.276**	0.101	0.102	0.125
	(0.102)	(0.0953)	(0.130)	(0.0975)	(0.0959)	(0.0933)
SA_{ijt}	-0.00463	-0.00434	-0.00139	-0.00158	-0.00150	-0.00459
	(0.00499)	(0.00487)	(0.00237)	(0.00242)	(0.00241)	(0.00523)
S_{it}	0.000849	0.00470	0.0114*	0.0143**	0.0140**	0.00698
	(0.00420)	(0.00428)	(0.00677)	(0.00610)	(0.00619)	(0.00446)
V_{jt+2}	-0.0264	-0.0566	-0.00226	-0.0543	-0.0563	-0.0647
	(0.0572)	(0.0508)	(0.0237)	(0.0474)	(0.0518)	(0.0552)
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
ATC1 FE	Y	Y	N	Y	Y	Y
(ATC1*Year) FE	N	N	N	N	Y	Y
N	29,486	29,486	29,486	29,486	29,486	29,486

Notes: Coefficients and clustered standard errors (by ATC) are reported. ATC is defined as anatomical therapeutic chemical classification. *** p<0.01, ** p<0.05, * p<0.10