THE LAW AND ECONOMICS OF GENERIC DRUG REGULATION

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ECONOMICS
AND THE COMMITTEE OF GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Christopher Scott Hemphill
August 2010
I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Timothy Bresnahan, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Petra Moser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

A Polinsky

Approved for the Stanford University Committee on Graduate Studies.

Patricia J. Gumport, Vice Provost Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.
Abstract

This dissertation examines the law and economics of generic drug entry, and the problems that arise from specific U.S. regulatory arrangements that govern innovation and competition in the market for patented pharmaceuticals. As Chapter 1 explains, competitive entry by generic drug makers is limited by both patents and industry-specific regulation, which together provide the means for brand-name drug makers to avoid competition and thereby recoup large investments in research, development, and testing. At the same time, the complex rules of the Hatch-Waxman Act furnish a pathway by which generic drug makers may challenge the validity or scope of brand-name patents, with a view to entering the market with a competing product prior to patent expiration.

The subsequent chapters examine several aspects of the competitive interaction between brand-name and generic drug makers. Chapter 2 analyzes settlements of patent litigation between brand-name and generic drug makers, in which the brand-name firm pays the generic firm in exchange for delayed market entry. Such pay-for-delay settlements are an important, unresolved question in U.S. antitrust policy.

The analysis reveals that the pay-for-delay settlement problem is more severe than has been commonly understood. Several specific features of the Act—in particular, a 180-day bounty granted to certain generic drug makers as an incentive to pursue pre-expiration entry—widen the potential for anticompetitive harm from pay-for-delay settlements, compared to the usual understanding. In addition, I show that settlements are “innovation inefficient” as a means of providing profits and hence ex ante innovation incentives to brand-name drug makers. To the extent that Congress
established a preferred tradeoff between innovation and competition when it passed the Act, settlements that implement a different, less competition-protective tradeoff are particularly problematic from an antitrust standpoint.

Chapter 3 synthesizes available public information about pay-for-delay settlements in order to offer a new account of the extent and evolution of settlement practice. The analysis draws upon a novel dataset of 143 such settlements. The analysis uncovers an evolution in the means by which a brand-name firm can pay a generic firm to delay entry, including a variety of complex “side deals” by which a brand-name firm can compensate a generic firm in a disguised fashion. It also reveals several novel forms of regulatory avoidance. The analysis in the chapter suggests that, as a matter of institutional choice, an expert agency is in a relatively good position to conduct the aggregate analysis needed to identify an optimal antitrust rule.

Chapter 4 examines the co-evolution of increased brand-name patenting and increased generic pre-expiration challenges. It draws upon a second novel dataset of drug approvals, applications, patents, and other drug characteristics. Its first contribution is to chart the growth of patent portfolios and pre-expiration challenges. Over time, patenting has increased, measured by the number of patents per drug and the length of the nominal patent term. During the same period, challenges have increased as well, and drugs are challenged sooner, relative to brand-name approval. The analysis shows that brand-name sales, a proxy for the profitability of the drug, have a positive effect on the likelihood of generic challenge, consistent with the view that patents that later prove to be valuable receive greater ex post scrutiny. The likelihood of challenge also varies by patent type and timing of expiration. Conditional on sales and other drug
characteristics, drugs with weaker patents, particularly those that expire later than a drug’s basic compound patent, face a significantly higher likelihood of challenge. Though the welfare implications of Hatch-Waxman patent challenge provisions are complicated, these results suggest these challenges serve a useful purpose, in promoting scrutiny of low quality and late-expiring patents.
Acknowledgments

In preparing this dissertation I have incurred many debts. Tim Bresnahan and Mitch Polinsky supported the project during its long gestation. Laura Hemphill, Bert Huang, Mark Lemley, Petra Moser, and Richard Posner provided particularly helpful comments on the individual papers. My colleagues at Columbia, including Harold Edgar, Elizabeth Emens, Robert Ferguson, Victor Goldberg, Harvey Goldschmid, Jeffrey Gordon, Zohar Goshen, Michael Heller, Bert Huang, Avery Katz, Clarisa Long, Thomas Merrill, Gillian Metzger, Edward Morrison, Alex Raskolnikov, Dan Richman, Robert Scott, and Susan Sturm, offered a terrific sounding board. Bhaven Sampat has been an ideal collaborator and co-author of the empirical project discussed in Chapter 4. Experts at the Food and Drug Administration (FDA) and Federal Trade Commission (FTC), including Brad Albert, Donald Beers, Elizabeth Dickinson, Don Hare, Mike Kades, and Markus Meier, spent many patient hours explaining various aspects of generic drug regulation to me.

The Handler, Kapnick, and Olin Foundations, the Program in Antitrust and Intellectual Property at Stanford Law School, and Columbia’s Julius Silver Program in Law, Science and Technology supported this project at various stages. The Blanche and Irwin Lerner Center for Pharmaceutical Management Studies at Rutgers Business Center, where I am affiliated as a non-residential fellow, furnished sales data. An ISERP seed grant at Columbia funded the purchase of additional sales data from IMS Health. Joshua Sharfstein, Nancy Sager, and Beth Noveck assisted in securing FDA data under a Freedom of Information Act request. Greg Rosston, John Shoven, and the
Stanford Institute for Economic Policy Research provided a superb setting in which to complete the project.


On Chapter 3, Jonathan Baker, Ben Beaton, Mike Carrier, Glenn Cohen, Tom Cotter, Dan Crane, Rebecca Eisenberg, Einer Elhauge, John Golden, Laura Hemphill, Keith Hylton, Margaret Lemos, Christopher Leslie, Tracy Lewis, Anup Malani, Michael Meurer, Barak Richman, Ben Roin, David Rosenberg, Carl Shapiro, Scott Stern, Jeannie Suk, Talha Syed, Willard Tom, Phil Weiser, and Christopher Yoo; audiences at Boston University, Columbia, Harvard, Stanford, and the Universities of Bonn, Colorado-Boulder, East Anglia, Michigan, Minnesota, North Carolina, and
Pennsylvania; and participants at the annual meeting of the American Law and Economics Association all provided helpful discussion and comments.

On Chapter 4, Joe Farrell, Howard Shelanski, and participants in workshops at Columbia’s Mailman School of Public Health, ETH Zurich, the FTC, Pfizer’s annual PEPC conference, Stanford Law School, the University of Toronto, and Yale Law School provided helpful comments. Janine Beach, Gerald Brant, Melanie Brown, Jessie Cheng, Akshay Deoras, Doug Geyser, Heng Gong, Matt Huppert, Jeffrey Kessler, Andrea Lee, Kevin Lin, Tejas Narechania, Teddy Nemeroff, Sam Salganik, Rodrigo Santos, Sannu Shrestha, Yan Shurin, and Kyle Zipes provided excellent research assistance at various stages of the project.

Substantial parts of Chapters 2 and 3 have been published (Hemphill 2006, 2009a) and also provided the basis for testimony before the Subcommittee on Commerce, Trade, and Consumer Protection of the House Committee on Energy and Commerce (Hemphill 2007, 2009b). I have consulted with the FTC on the antitrust issues raised by brand-generic patent settlements. Views or errors in this manuscript are mine alone.
# Table of Contents

Abstract ........................................................................................................................................ iv  
Acknowledgments .................................................................................................................. vii  
List of Tables ...................................................................................................................... xi  
List of Figures ...................................................................................................................... xii  
Chapter 1: An Introduction to Generic Drug Regulation ......................................................... 1  
Chapter 2: The Regulatory Economics of Pay-for-Delay Agreements ............................... 13  
Chapter 3: A Public Survey of Pay-for-Delay Agreements .................................................. 79  
Chapter 4: Patent Portfolios and Generic Drug Challenges ................................................ 139  
Figures and Tables .............................................................................................................. 172  
Appendix: Selected Agreements Between Drug Makers ..................................................... 192  
References ............................................................................................................................. 218  
Cases Cited ............................................................................................................................ 235
List of Tables

Table 1: Antitrust Challenges to Pay-for-Delay Settlements .................................................... 180
Table 2: Availability of Full Text of Settlement Agreements .................................................. 182
Table 3: Typology of Settlements ............................................................................................ 183
Table 4: Settlements with Monetary Payment ........................................................................ 183
Table 5: Settlements with Retained Exclusivity ...................................................................... 184
Table 6: Pay-for-Delay Settlements Summarized .................................................................... 184
Table 7: Description of Variables ............................................................................................ 185
Table 8: Summary of Variables ............................................................................................... 186
Table 9: Data Structure of Duration Analysis ........................................................................... 186
Table 10: Logit Model of Hazard of Challenge—Baseline and Patent Measures ..................... 187
Table 11: Logit Model of Hazard of Challenge—Alternative Patent Quality Measures ........... 188
Table 12: Logit Model of Hazard of Challenge—Non-Parametric Sales ................................ 189
Table 13: Logit Model of Hazard of Challenge—Lagged Sales ............................................... 190
Table 14: Logit Model of Hazard of Challenge—Future Sales ............................................... 191
List of Figures

Figure 1: Share of Drugs with Pre-Expiration Challenge by Sales Quintile........................................ 172
Figure 2: Average Relative Price of Prozac Before and After Generic Entry ............................. 172
Figure 3: Approval Timeline for a New Drug .............................................................................. 173
Figure 4: The Effect of a Guaranteed Bounty .............................................................................. 173
Figure 5: Patents Per Drug ............................................................................................................ 174
Figure 6: Share of Drugs with Non-AI Patent ............................................................................. 175
Figure 7: Years of Nominal Patent Life ....................................................................................... 176
Figure 8: Share of Drugs with Paragraph IV Challenge ............................................................. 177
Figure 9: Share of Drugs with Early Paragraph IV Challenge.................................................... 178
Figure 10: Predicted Values of Hazard of Challenge by Drug Type .......................................... 179
Figure 11: Predicted Values of Hazard of Challenge by Patent Portfolio .................................... 179
Chapter 1: An Introduction to Generic Drug Regulation

Introduction

This dissertation examines a set of problems that arise when makers of patented brand-name drugs face potential competition from manufacturers of so-called “generic” drugs. The problems are a consequence of incentives set up by a highly complex, industry-specific regulatory regime, the Drug Price Competition and Patent Term Restoration Act of 1984, commonly known as the Hatch-Waxman Act. The starting point of this thesis is that the central challenge in answering questions like these is to understand the specific elements of the regulatory scheme and the economic incentives created by those elements. From that first step, one can more readily identify the opportunity for and scope of the resulting collusive or exclusionary behavior, the likelihood and size of welfare losses, and the nature of an optimal regulatory response. The essays in this dissertation reflect an attempt to help resolve concrete problems in a very important segment of the economy: the development and distribution of prescription drugs. At the same time, they may offer a template of use in resolving other questions of competition and regulation.

In approaching this task, questions of economics and law are deeply intertwined. I hope to persuade economists (and some lawyers!) that the regulatory details here are not merely details, but at the very core of the incentives facing individual drug makers. Skipping the specifics misses the point of what the drug makers are attempting to accomplish, the size of the resulting economic effects, and the likeliest routes to deterrence or remediation. What is required, then, is a sophisticated understanding of the regulatory setting, combined with a nuanced view of economic effects.
This chapter sets the stage for the analysis conducted in subsequent chapters. Part I briefly introduces the economic role of patents in promoting innovation by brand-name drug makers. Part II explains a particular form of generic drug competition—“Paragraph IV” challenges—in which a generic drug maker seeks to compete with a brand-name drug maker prior to the expiration of patents protecting the brand-name drug. Part III introduces the subsequent chapters and provides a further discussion of the mode of regulatory economics pursued here.

I. Patents as an Instrument of Innovation

A new brand-name drug is, in essence, an information good. Once its formula is understood, it is relatively straightforward and cheap for others to manufacture it without incurring similar research and development costs. In general, drug companies, compared to innovators in other industries, cannot as easily rely upon a head start, complementary assets, and scale of production as means to preserve profits. Nor can a drug maker easily keep the chemical formula secret.

For blockbuster drugs as with blockbuster films, the ability to legally exclude rivals from offering a copy preserves the return on a massive initial investment. Legal protection is not the whole story, to be sure. For example, so-called “biologics” derived from living sources are relatively difficult to make and replicate, providing their manufacturers with an additional source of protection. Moreover, as observed in Chapter 4, some drugs are easier to copy than others. But legal protection is central for the great majority of brand-name drugs. Economic theory predicts that the expectation of profits from new discoveries will induce investment in research, development, and
testing. The available empirical evidence suggests that higher drug profits are indeed correlated with greater research and development efforts (Giaccotto et al. 2005).

The industry’s reliance upon patents, as one legal instrument for preserving appropriability, is unusually great. Large-scale surveys of research and development employees have indicated that patents are unimportant for appropriating returns from research and development in most industries (Levin et al. 1987; Cohen et al. 2000). Pharmaceuticals are the most important exception. This mode of support for drug research, development, and testing is augmented by the work of government and academic researchers.

New drugs are thought to result from an unusually simple technology of innovation. This industry possess relatively few indicia of cumulative innovation, in which progress is incremental, and the set of potential innovators widely dispersed (Lessig 2001; Scotchmer 2004). As a general matter, pharmaceuticals have been associated with the normative case for patents that are “broad, stand alone, and confer almost total control over subsequent uses of the product” (Burk and Lemley 2003). Thus, studying drug invention and competition avoids certain complexities in patent policy that affect the study of other industries. That is not to say that one product is always associated with a single patent. Increasingly, as shown in Chapter 4, multiple patents are filed for a single product.

Sequential innovation does play a role in drug development, but not in the way usually understood. The usual point is that when an innovation developed elsewhere is itself the raw material for further invention, strong, multiple rights of exclusion can lead to underuse (Heller 1998, 2008; Shapiro 2001). In theory, underuse concerns may play
some role in the overlapping field of biotechnology, where patented upstream research tools can produce downstream underuse (Heller and Eisenberg 1998; Rai 2001), but there is little evidence so far that innovation has been hampered in practice (Walsh et al. 2003). In generic drug development, sequential innovation takes a different form. The generic drug maker is the later innovator, and seeks to produce a close substitute to its predecessor. Chapter 2 examines the regulatory arrangements that encourage the development of substitutes.

II. Regulation as a Second Instrument of Innovation—and Entry

Drug innovation and competition is more complicated than a patent-focused account might initially suggest. In addition to the patent regime, a complex, industry-specific regulatory scheme affects the degree of appropriation and competition to which a brand-name drug maker is subject. This scheme comprises the Food, Drug, and Cosmetic Act of 1938, the Hatch-Waxman Act passed in 1984, and a set of further statutory changes contained in the Medicare Modernization Act (MMA) of 2003. In addition to the statutes, entry is controlled by Food and Drug Administration (FDA) regulations interpreting and implementing these statutes, and also by federal courts, which have provided extensive rulings on various controversial elements of the statutes and regulations. Taken together, these elements of regulation are as important as patents in determining the duration and scope of protection for a drug. They are, however, much less understood.

This chapter introduces the scheme. Under the Food, Drug, and Cosmetic Act, a brand-name firm must demonstrate that a new drug is safe and effective before the FDA
will approve it for marketing. Making that demonstration as part of a so-called New Drug Application (NDA) is a lengthy, expensive process, consuming years and many millions of dollars to conduct the necessary clinical trials (DiMasi et al. 2003).

Once the brand-name firm places a patented drug on the market, a generic firm may seek to market a competing, “therapeutically equivalent” version of the same drug by filing an Abbreviated New Drug Application, or ANDA, with the FDA. The generic firm must demonstrate that the rate and extent of absorption of the active ingredient (AI) in its proposed product are the same as in the brand-name drug. Establishing equivalence is not trivial but is much less expensive than NDA clinical trials, requiring an outlay on the order of $1 million (FDA 2003b). In addition, the generic firm must show that the products contain the same conditions of use, route of administration, dosage form, strength, and labeling.

For many drugs, the generic firm seeks entry prior to patent expiration. Most patents covering a drug—on its active ingredient, the composition of the drug, or a method of use—are listed by the brand-name firm in an FDA document titled “Approved Drug Products with Therapeutic Equivalence Evaluations,” more commonly known as the Orange Book. The Orange Book also lists, as the formal title suggests, any approved therapeutically equivalent generic drugs.

A generic firm’s ANDA must contain one of four certifications described in § 355(j)(2)(A)(vii). (Unless otherwise noted, statutory references are to Title 21 of the United States Code.) A “Paragraph I” certification seeks immediate approval on the grounds that no patents have been listed in the Orange Book. A Paragraph II certification asserts that any relevant patents have expired. A Paragraph III certification
concedes that one or more applicable patents have not expired, and that approval is not sought until patent expiration. In that case, the FDA delays ANDA approval until patent expiration.

Generic firms seeking entry prior to patent expiration file a Paragraph IV certification. A Paragraph IV certification asserts that one or more applicable patents are invalid or not infringed by the proposed generic product. In this dissertation, I focus on ANDAs with Paragraph IV certifications (ANDA-IVs). If the filing is successful, the generic firm can launch a competing product without repeating the costly safety and efficacy studies that the FDA requires as a condition of brand-name approval.

The assertion contained in a Paragraph IV certification can take several forms. For example, a generic firm might argue that the patent is invalid because it was procured inequitably, that it is inherently anticipated by the prior art, or that the drug’s initial testing violates the public use bar. Alternatively, the firm might contend that it has devised a noninfringing bioequivalent form of the drug—for example, a different crystalline structure of the same AI, or a different way to accomplish some desirable time-release feature of the innovator’s drug. Note that a generic product can be bioequivalent to its brand-name counterpart, as required for a successful ANDA application, and yet not infringe the brand-name firm’s patents. The filing of such an ANDA is an act of patent infringement. In response to the ANDA, the brand-name firm may and frequently does file a patent infringement suit to establish validity and infringement.

Most ANDAs are filed under Paragraph I, II, and III (FTC 2002). Nevertheless, ANDAs with Paragraph IV certifications play an important role in encouraging early
generic competition. Between 1984 and 2000, generic firms filed Paragraph IV challenges for 104 drugs (FTC 2002). Eighty drugs were challenged between January 2001 and June 2003 (Muris 2003). And, as Chapter 4 discusses, that number has continued to grow.

Challenges have become the norm for the top-selling drugs. Comparing the ten best-selling drugs of 2000 (listed in Pear 2001) to the FDA’s published list of drugs receiving Paragraph IV challenges (FDA 2010) reveals that all but one attracted a challenge. The remaining drug was ineligible for a Paragraph IV challenge because it had no Orange Book-listed patents in the first place. The bestsellers of 2005 (listed in Herper 2006) show much the same result. Of the top 20 drugs, six are biologic drugs not subject to the Hatch-Waxman regime. All but one of the remaining fourteen drugs have faced pre-expiration patent challenges.

Clearly, the size of the market opportunity is one reason why generic firms launch challenges. This point is demonstrated vividly in Figure 1. The figure reports pre-expiration challenges for new molecular entities—roughly speaking, drugs with a novel AI—approved between 1996 and 2002. Eighty-five percent of drugs in the top quintile of sales receive a challenge within three years of becoming eligible for such a challenge. As sales drop, the likelihood of challenge does too. In the bottom quintile of sales, there are no challenges. Chapter 4 examines what other factors matter, in addition to sales.

These challenges have important effects upon early generic entry. The FTC studied challenges initiated between 1992 and 2000 involving 104 drugs (FTC 2002). Of the fifty-nine drugs whose challenges were neither pending nor settled at the end of
the study period, the innovator declined to sue with respect to twenty-nine, effectively permitting rapid generic entry. The generic firm won in another twenty-two cases. The innovator won in the remaining eight cases. These figures ignore two cases in which the patent expired before the litigation was resolved, and one in which an NDA was withdrawn before the litigation was resolved.

ANDA challenges have led to pre-expiration competition for many major drugs. Of the ten best sellers from 2000, at least four—Paxil, Prilosec, Prozac, and Zocor—have experienced pre-expiration competition (MDIS 2002; Generic Line 2003; Greene 2006). When a generic firm enters the market, prices fall, often dramatically. Prozac provides an example. Figure 2, drawn from Frank and Hartman 2009, shows relative prices on a quantity-adjusted basis, both before and after generic entry in August 2001. After generic entry, prices fell by about a third over the subsequent three years. With multiple generic entrants, prices fall even more.

For this reason, pre-expiration challenges have become a central issue in debates about drug innovation and competition. Some commentators argue that Paragraph IV challenges are necessary to clear away patents, increasingly asserted by brand-name pharmaceutical firms, that are of questionable validity (Engelberg 1999), as a route to low-price access to drugs. On the other hand, Higgins and Graham 2009 and others have suggested that these challenges have a significant negative impact on effective patent life and research incentives, and contribute to a perceived reduction in industry innovation.
III. Plan and Approach of the Dissertation

The dissertation is organized as follows. Chapter 2 provides a theoretical analysis of a situation that frequently arises in pre-expiration challenges, in which the associated patent infringement litigation goes seriously awry. The parties, rather than adjudicating the merits of the patent suit, settle, and do so in a way that reduces consumer welfare. These settlements arising from Paragraph IV challenges are the most important unresolved question in U.S. antitrust policy, measured by the amount of commerce affected and repeated high-level judicial attention.

As Chapter 2 explains, the settlements are the subject of a large economic and legal literature. The literature, however, has ignored the economic effect of important, specific elements of the regulatory scheme. Recognizing these elements and their impact on drug maker incentives provides a new perspective on which settlements are anticompetitive. The analysis demonstrates that the settlement problem is much worse than is commonly believed, even by advocates of antitrust liability. Central to this analysis is the role of an industry-specific bounty awarded to certain generic drug makers as an incentive to pursue Paragraph IV challenges. The bounty greatly widens the scope of feasible collusion.

Although economists and legal scholars have debated extensively which settlements are anticompetitive and why, they have given little attention to a range of empirical questions. For example, how many problematic settlements are there? How can we go about identifying a problematic settlement in the first place? How have these settlements changed over time?

Chapter 3 turns to an empirical examination of these questions. It relies upon a
novel dataset of brand-generic settlements drawn from publicly available sources. The analysis demonstrates the difficulty of getting things right in practice, given both the inevitable incompleteness of information about settlements, and ongoing changes in the practice of regulated firms. The analysis highlights certain advantages that agencies possess in identifying and deterring anticompetitive conduct.

Chapter 4 shifts the focus from the resolution of Paragraph IV challenges to an inquiry into their determinants, a question that has not been subjected to systematic empirical study. This chapter, which is joint work with Bhaven Sampat, draws upon a second novel dataset of drug approvals, applications, patents, and other drug characteristics. Its first contribution is to trace the co-evolution of brand-name and generic drug maker strategies. Brand-name drug makers are increasingly building sizable patent portfolios with respect to a single drug. Generic drug makers, meanwhile, are increasingly filing Paragraph IV challenges against individual drugs, and doing so earlier in the product’s life cycle. The analysis draws upon a second novel dataset relating the incidence of Paragraph IV challenges to the timing of drug approval, market size and other characteristics of the drug, and the patent portfolio developed by the brand-name drug maker and faced by the generic firm. The analysis uncovers evidence that the addition of “weak” patents—that is, patents on aspects of the drug other than the AI—with late expiration dates is causing part of the rise in Paragraph IV challenges.

This chapter also provides a second perspective on the institutional choice discussion introduced in Chapter 3. The decision of a generic drug maker to launch a Paragraph IV challenge is a setting in which ex post analysis through individualized scrutiny seems to be working fairly well, at least as judged by which patents are
challenged and when, for it is the weakest patents on the most important drugs that are disproportionately challenged.

This dissertation employs a mix of analytical techniques, including theory, econometrics, and focused case studies. It is a work of law and economics, in the sense that it provides an economic analysis of legal rules and institutions, with a view to identifying the optimal rule or optimal decision-maker (for an introduction, see Polinsky 2003). Beyond that, the goal throughout is to remain deeply engaged with the specific elements of regulation in a particular industry. In what follows, we will examine a valuable, industry-specific “bounty” granted to certain generic drug makers that challenge brand-name drug patents; explore the differential effects of noninfringement and invalidity claims; and assess the difference in economic incentives for makers of highly novel new drugs—so-called “new chemical entities”—and new and improved versions of existing therapies. These legal distinctions are highly relevant to the incentives of drug makers and the outcome of regulation.

The approach to competition policy pursued here is unusual. The research agenda is driven on the one hand by work that cuts across many industries—for example, that of industrial economists to understand the effects of a particular practice, and efforts by legal scholars to reconcile antitrust and intellectual property law—and on the other hand, by lawyers and economists focused on the proper resolution of a specific case (for a representative collection, see Kwoka and White 2004). Largely left out of that agenda is sustained attention to the incentives set up by regulation in a particular industry. Such an approach is far from revolutionary—attention to specific regulatory features has played an important role in the work of Paul Joskow, Alfred Kahn, and
Michael Scherer, among others—though it seems to have fallen out of fashion. This dissertation reflects an effort to reorient attention to that intermediate level of analysis. Doing so entails some loss in generality, but the hope is that the resulting insights justify that tradeoff.
Chapter 2: The Regulatory Economics of Pay-for-Delay Agreements

Introduction

To what extent do legislative enactments shape the scope of antitrust liability? The answer is not purely a matter of antitrust law. Antitrust’s basic law, the Sherman Act, takes a famously broad approach in its two major liability-setting provisions. Section 1 purports to condemn “[e]very contract, combination . . . , or conspiracy, in restraint of trade”; section 2 forbids a firm to “monopolize.” These provisions do not much constrain antitrust enforcement agencies or courts. Subsequent interpretation has narrowed the scope of section 1 to unreasonable restraints and given content to the ill-defined concept of “monopolization.” A law referred to as “the Magna Carta of free enterprise” (Topco 1972) can hardly be expected to determine the results of particular cases. Instead, enacted antitrust law is generally understood to grant agencies and courts a broad license to develop policy in an incremental fashion (Elhauge 2002; Eskridge and Ferejohn 2001; Manning 2003; Merrill 1985; Farber and McDonnell 2005).

That license has limits, for two other kinds of regulatory law address firm conduct within the ambit of antitrust. One important and familiar source is intellectual property law, particularly patent law. Accounts of the intersection between antitrust and patent law emphasize the conflict in means between the two (Hovenkamp et al. 2009). The usual account of antitrust law emphasizes allocative efficiency, avoidance of the distortion that results when consumers’ unwillingness to pay high prices diverts them to less desirable substitutes (Posner 2001). That concern produces a preference for low prices, though only down to a point, since prices below marginal cost also harm allocative efficiency.
The instrumental case for patent law, by contrast, depends upon high prices as a means to reward and thereby encourage innovation, a source of dynamic efficiency. Because many competitive practices both distort allocation and provide a dynamic benefit, the conflict in means between antitrust and intellectual property can be stark. A substantial literature seeks an optimal reconciliation between these competing values by encouraging innovation without sacrificing too much consumer access (Barton 1997; Baxter 1966; Kaplow 1984).

Intellectual property law, however, is not the only kind of regulatory enactment that affects antitrust decisionmaking. This chapter isolates and examines a second overlap between antitrust and regulatory law, the ways in which an *industry-specific regulatory regime* alters the contours of antitrust enforcement. A particular regulatory regime sets the boundaries of feasible anticompetitive conduct. At the same time, it embodies a specific congressional judgment about the proper balance between competition and innovation in an industry, and hence provides content to what we might think of as special-purpose antitrust law. Both effects shape antitrust enforcement in often subtle ways. Identifying the impact of an industry-specific regulatory regime in a particular context requires careful, sustained attention to the principal features of the relevant regulatory scheme. That general project, though difficult, is also necessary to identify the boundaries of permissible competitive conduct in regulated industries as diverse as telecommunications, financial services, and—the primary focus of the present analysis—pharmaceuticals.

In this chapter, these questions are addressed in the specific context of “pay-for-delay” settlements in the pharmaceutical industry. Settlements result from a generic
drug maker’s effort to market a competing version of a patented brand-name drug, using the Paragraph IV procedure introduced in Chapter 1. The brand-name firm responds with a patent infringement suit that claims its product is protected by one or more patents, and the generic firm counters that the patent is invalid or not infringed by the proposed generic product. The brand-name firm, rather than take a chance that the generic firm might win that argument in court, thereby ending its exclusivity on the product, settles the litigation by paying the generic firm to abandon the challenge and delay entry. Does this agreement violate antitrust law?

This question is the most important unresolved issue in U.S. antitrust policy, measured both by economic importance and repeated high-level judicial attention. Successful pre-expiration challenges reallocate billions of dollars from producers to consumers. Settlements in 2008 alone included Lipitor (more than $7 billion in annual U.S. sales) and Nexium (more than $3 billion). The importance and difficulty of the question has prompted the Supreme Court to seek the Solicitor General’s views three times since 2004. Over the past twelve years, the FTC has attacked settlements involving six drugs as violations of antitrust law, two of which are pending as of May 2010. Private antitrust suits have been brought, ignoring follow-on suits, for four additional drugs. These challenges are listed in Table 1.

The issue remains unsettled. Thus far, courts have taken a relatively sympathetic, albeit highly uneven, stance toward pay-for-delay settlements. As shown in Table 1, three circuits have rejected antitrust liability, although one of these, the Second Circuit, has recently suggested it might reconsider this position in an en banc review by the whole circuit (Arkansas Carpenters 2010). One circuit has imposed
liability, albeit in an unusual “interim” settlement case. In addition, one district court recently recognized possible anticompetitive effects under the specific facts of one set of settlements, while appearing to adopt some of the reasoning of those circuits hostile to antitrust liability.

A large literature is focused upon the pay-for-delay settlement problem in light of its economic importance and continued uncertainty about the proper legal treatment of such settlements. Commentators have framed the question as part of a wider debate about the intersection of patent and antitrust, and frequently seek to resolve these cases at that level of generality. However, this perspective is incomplete. Existing analyses, though attentive to the antitrust-patent intersection, have overlooked the importance of the antitrust–regulated industry intersection.

This chapter fills that gap by examining in detail the industry-specific regulatory scheme introduced in Chapter 1. The Hatch-Waxman Act and related FDA regulations alter the prospect for anticompetitive conduct by regulated parties. An important feature of the regime is a large incentive to litigate the validity and scope of an innovator’s patents, a “bounty” worth hundreds of millions of dollars to a generic drug maker. The bounty has an unusual form: under certain circumstances, the generic firm enjoys the 180-day exclusive right to market a generic version of the drug in competition with the innovator. This is effectively a duopoly during that period before other generic firms are permitted to enter the market.

The availability of the bounty is limited in an important respect. As explained in Part II, only the first generic firm to challenge an innovator’s patents has any prospect of earning the bounty. Because no other firm has a similar opportunity, buying off the
first challenger is an effective means to head off the most potent threat to entry. Previous accounts have neglected this effect, and courts have misperceived the availability of the bounty, resulting upon occasion in serious error.

In addition, the bounty can provide a means, generally overlooked, for the innovator to compensate a generic firm. A settlement that guarantees the bounty to a generic firm can provide a disguised payment for delay, making possible an allocative harm even where little or no cash changes hands. This mechanism of payment has been underappreciated in legal analyses and academic commentary.

Explicating the economic incentives set up by the regulatory scheme has a second payoff. One prominent legal theory of antitrust enforcement states that where the conduct concerns an intellectual property holder, antitrust concerns are deferred in favor of the ex ante innovation incentives central to intellectual property policy (for a critical review, see Kaplow 1984). If that type of argument is accurate as a statement of law, then its logic should also apply to other types of regulatory law that govern the balance between dynamic and static welfare. In other words, taking that argument seriously requires attention not only to intellectual property, but to the decisions about dynamic and static welfare contained in industry-specific regulation.

A close examination of that decision, as implemented in the Hatch-Waxman Act, reveals that it imposes a tax on certain pharmaceutical innovators, but a tax whose effect is highly uneven. For some innovators, a different set of industry-specific features comes to the fore—a series of distinctive protections for innovators that serve to delay entry by a generic firm. These features effectively subsidize certain pharmaceutical innovations. Economic analysis of these industry-specific features is an important input
This chapter proceeds in three parts. Part I describes the pay-for-delay settlement problem and disagreement about its resolution among enforcement agencies, courts, and commentators. Part II explains the means by which the industry-specific regulation of pharmaceuticals alters the scope of anticompetitive activity by regulated parties. Part III assesses the economic judgment, made by Congress and embodied in the Hatch-Waxman Act, about the proper balance of competition and innovation, and shows how this judgment undermines certain arguments against antitrust liability. The conclusion discusses the utility gained by understanding other antitrust problems through the lens of regulatory design.

I. The Pay-for-Delay Settlement Problem

A pay-for-delay settlement of a pre-expiration patent challenge follows the basic pattern discussed in the introduction, although, as Chapter 3 explains, individual settlements offer many variations on the theme. In the course of settling patent litigation, the generic firm abstains from entry, the innovator agrees to make a payment to the generic firm, and the parties agree to dismiss the patent suit. The agreement may also provide for limited pre-expiration entry.

A large literature—more than thirty articles or book chapters, not including student law review notes, as of 2009—considers the competitive effects and proper antitrust treatment of pay-for-delay settlements. Many commentators note the static welfare loss that results when a brand-name firm makes a payment to the generic firm, and the generic firm delays entry. Economic modeling has shown formally that
settlements that include a cash payment from the patentee to the infringer provide consumers with less welfare, on average, than seeing the litigation to completion (Bulow 2004; Shapiro 2003a, 2003b).

Assessing this welfare loss is complex. In an ordinary market, setting a price above marginal cost produces an allocative distortion and accompanying welfare loss for consumers because consumers who value the good above its marginal cost, but below the prevailing price, are deflected to less-desired substitutes. To the extent that public and private insurance secures the purchase of a drug, this distortion is reduced, though it is not eliminated (as insurance is incomplete). Moreover, the higher price produces new distortions (and hence inefficiency) in the decisionmaking process of the insurance provider, through decisions to charge higher premiums and not to reimburse drugs whose value exceeds their marginal cost.

This conclusion is a special case in a much larger literature. Contracting at the expense of nonparties is a general problem (Segal 1999). In an agreement between competitors, consumers are the relevant nonparties. Despite consumers’ aggregate economic interest—the short-run consumer gain from lower prices exceeds producers’ reduced profits—collective action problems present an obstacle to paying off producers who (unless legally constrained) will act at the consumers’ expense. If transaction costs were low enough, consumers could band together and make a large fixed payment in exchange for marginal-cost pricing, either by contracting with or owning the producer (Hansmann 1996). These options seldom are available.

The static welfare loss has motivated an enormous antitrust enforcement effort, summarized in Table 1, including challenges to settlements on ten drugs. A rival’s
effort to remove a patent-based barrier to entry, like a price cut, provides an indirect allocative benefit in the course of a private pursuit of profit, and an agreement that reduces this benefit is a traditional object of prohibition under the Sherman Act. The conclusion that this loss gives rise to an antitrust violation depends upon acceptance of the view, on which these models are premised, that consumers are entitled as a matter of antitrust law to the average benefits of litigation. Part III offers an industry-specific defense of that premise.

Although many pay-for-delay agreements have attracted antitrust attention, many others have not. Chapter 3 discusses a large set of recent settlements that have not received antitrust scrutiny. Some older settlements have escaped antitrust scrutiny too, including a settlement involving the first blockbuster drug, Zantac. At the time of the settlement, Zantac was the world’s best-selling prescription medicine, with annual U.S. sales of about $2 billion, and removing the risk of early generic entry appears to have conferred upon Glaxo a multibillion-dollar benefit, judging from the multibillion-dollar increase in Glaxo market valuation immediately following settlement (Reguly 1995). In exchange, Glaxo, Zantac’s manufacturer, appears to have paid the generic firm more than $100 million to delay entry. The amount was not disclosed by the parties, but can be inferred by reverse engineering disguised information in FTC 2002, a comprehensive analysis of the first few years of settlement practice. The Appendix describes this and other settlements in detail, and the method used to infer the size of the payment. In many cases, including Zantac, the payments were not limited to cash. Chapter 3 describes the variety of forms that compensation can take.

Antitrust enforcement under the Sherman Act is not the only way to address the
pay-for-delay problem. Congress could pass new legislation banning certain settlements. (As of May 2010, the latest effort to do so fell short of enactment, having been removed from the omnibus health care bill shortly before its passage.) Alternatively, antitrust could be avoided altogether. For example, Congress could modify or eliminate the 180-day exclusivity period, particularly for settling parties, or provide a means and incentive for drug purchasers, including the government, to challenge pharmaceutical patents. Such changes could address the incentives that give rise to the pay-for-delay settlement problem in the first place. As an alternative, settlements could be challenged at the moment they are reached by requiring the court conducting the patent infringement case to approve the settlement using procedures akin to those employed in class actions to prevent collusive settlements. Private “objectors” or the FTC could be recruited to try to persuade the court that the settlement ought to be rejected. Putting aside the question of political feasibility, however, such changes might not determine the legal status of the many settlements that have already been reached. Thus, antitrust law is a necessary component of any complete resolution of the pay-for-delay issue.

The allocative harm from settlement has motivated commentators to propose a variety of tests for identifying anticompetitive settlements and a range of procedural responses if a troubling settlement is identified. Creighton and Sher 2009 summarize several of the leading positions. For example, Leffler and Leffler 2004 take the view that all settlements that combine payment with delayed entry are per se violations of antitrust law. Others commentators propose a presumption of illegality (Department of Justice 2009). Still others have suggested that the matter should be judged through a
more detailed examination of the strength of the patent, compared to the details of the settlement (Department of Justice 2007). The stronger the patent, the less troubling a long delay in entry would be.

A common theme in many proposals has been a safe harbor for certain settlements thought not to raise any competitive concern. There is a widespread sense in the literature (for example, Hovenkamp et al. 2005; Brodley and O’Rourke 2002) that a mere agreement on entry dates—that is, an agreement that divides up the remaining term into a period of brand exclusivity, followed by a period of competition—raises no anticompetitive concern. To this end, the FTC has provided a safe harbor for agreements that set an entry date but include no cash payment from the innovator to the generic firm (Schering 2003; FTC 2004; FTC 2006). Economic modeling of generic drug entry commonly accepts the same underlying view (Leonard and Mortimer 2005; Shapiro 2003a).

Some commentators have argued that such settlements can bring significant benefits, and therefore some or even all such settlements should be permitted. First, numerous commentators (Blair and Cotter 2002; Cotter 2003; Crane 2004; Crane 2002; Langenfeld & Li 2003) have noted that a lenient policy toward settlement increases the brand-name firm’s profits, thereby preserving and improving the incentive to innovate. As legal support for this economic argument, they point to the Patent Act. The Patent Act, they argue, takes a highly innovation-protective stance regarding the proper tradeoff between innovation and consumer access, to which antitrust law should conform.

The potential scope of this argument is extremely broad: *Any* practice currently
prohibited by antitrust law, as practiced by innovators seeking to increase their profits, could be defended upon this ground. Even simple price fixing could be excused. In general, antitrust lacks any such exemption for collusive behavior. Part III considers to what extent such arguments should count in a proper antitrust analysis.

In addition to its effect on brand-name firms, a lenient standard might also encourage generic drug makers to file challenges. As Judge Richard Posner noted in a case, *Asahi Glass* 2003, concerning the antitrust treatment of certain pharmaceutical agreements, restrictions on an infringer’s opportunity to settle affect its incentives: “A ban on reverse-payment settlements would reduce the incentive to challenge patents by reducing the challenger’s settlement options should he be sued for infringement . . . .” That case was not about a pay-for-delay settlement, but the quoted dictum, and its conclusion that limiting such settlements “might well be thought anticompetitive,” has proved influential among some courts that have considered pay-for-delay settlements. These analyses tend to neglect the point that a challenge motivated simply by the prospect of lucrative settlement is not very valuable from a static welfare standpoint.

Second, Willig and Bigelow 2004 and others have constructed models under which, for certain parameter values, settlement is procompetitive. Under certain assumptions about risk aversion, asymmetric information, liquidity constraints, or future innovation, removing the risk of competition can be desirable. Some of these models rely upon the dynamic incentive for brand-name innovation or generic challenge, while others show benefits even if only static welfare is considered.

One source of static benefit is that settlements conserve litigation expense and benefit parties who are in the best position to arrange their own affairs. Judicial
opinions permitting pay-for-delay settlements frequently rely upon the view that the benefits of settlement weigh against antitrust liability, echoing the Supreme Court’s statement, expressed more than a century ago, that settling patent litigation is “a legitimate and desirable result in itself” (Bement 1902). Partly this result simply reflects a judicial reflex in favor of settlement. This reflex may be unusually acute due to the highly technical nature of pharmaceutical patent cases, which many federal judges prefer to avoid. Settlement also saves litigation costs, which can be quite substantial. The American Intellectual Property Law Association has estimated a median expense of $4.5 million for patent litigation with more than $25 million at risk (AIPLA 2005). Drug makers can therefore expect to save millions of dollars each, at least where the case has not progressed very far. The innovator is likely to save more, because it would likely spend more, as it has more at stake in the case. However, the amounts involved are dwarfed by the much larger effects on consumer welfare for all but the least important drugs.

Third, as Schildkraut 2004 and others have noted, the underlying economic structure of a pay-for-delay settlement generalizes beyond the particular cases under consideration. The pharmaceutical industry settlements that have received so much attention are merely the most visible and dramatic examples of this economic structure. Suppose, for example, that a patentee sues an alleged infringer who has entered the market, and the alleged infringer later agrees to exit the market, in exchange for which the patentee waives a claim to accrued damages. This agreement matches the basic pay-for-delay structure: a conferral of value that heads off litigation that, if the alleged infringer won, would increase consumer access. Although there is no cash payment, the
alleged infringer’s prior entry makes forgiveness of accrued damages a source of compensation by the incumbent.

Nor is the waiver a necessary component of the deal; the essential problem is unchanged if the alleged infringer exits and pays the patentee a sum less than the value of the patentee’s infringement claim. To see this point, take a setting for which a damage-plus-waiver agreement is the settlement outcome, and increase the amount of damages accrued, so that the alleged infringer must now make a payment to satisfy the patentee. In this case, too, the settlement likely brings less expected consumer benefit than taking litigation to conclusion.

Because it is far from clear that, as a general matter, consumers are entitled to the expected outcome of the avoided litigation, imposing liability for pharmaceutical pay-for-delay settlements introduces the specter of antitrust liability in a wide range of cases in which settlement imposes negative externalities upon consumers. Part III takes up this challenge, providing an economic and legal basis to impose antitrust liability upon drug makers, without necessarily extending that judgment to other industries.

Some of the arguments made on behalf of settlement are quite difficult to credit. For example, McDonald 2003 asserts that the suppressed competition is not cognizable as an antitrust violation because the competition is merely probabilistic. That objection ignores the fact that the suppressed entry subject to antitrust regulation is almost always probabilistic. As the D.C. Circuit in Microsoft 2001 explained, “[I]t would be inimical to the purpose of the Sherman Act to allow monopolists free reign to squash nascent, albeit unproven, competitors at will . . . .”

A final, particularly doubtful argument emphasizes the need for the present
study. This argument relies upon the role of pharmaceutical regulation in altering the incentives of the parties, compared to the usual incentives of patentees and infringers. In particular, some courts have seized upon the fact that a generic firm has a strong incentive to challenge an innovator but faces little risk. The generic firm’s infringement is by certification rather than entry—indeed, entry is barred by an automatic stay of FDA approval, discussed in Part III—so the generic firm is not subject to large damages if it loses the suit. That is not to say that the generic firm has *nothing* at risk, for if it loses the suit, its investment in proving bioequivalence and in litigation will have been wasted.

Whereas a settlement of litigation in which entry had already occurred might include a payment from the infringer to the patentee, a settlement in the present context, if settlement with cash changing hands is to occur at all, must necessarily include a payment from the patentee to the infringer. From this, some courts have concluded that pay-for-delay settlements are a “natural by-product” of the incentives set up by the Hatch-Waxman Act (*Schering-Plough* 2005; *Tamoxifen* 2006).

These courts are right to recognize the importance of the regulatory regime, but judicial treatments reflect deep confusion about the implications of that regime. True, paying for delay is “natural,” in the sense that the result is not unexpected given the incentives of the parties; the parties, if not legally constrained, will prefer pay-for-delay settlement to litigation. But as Hovenkamp et al. 2003 explain, that fact in no way *justifies* payments for delay. No doubt many government actions make price-fixing easier. Just think of a Defense Department policy to reduce its number of suppliers, with the possible effect of less competition among the bidders (Biesecker 2002). Such an
action, however, provides no necessary protective coloration to oligopolists who subsequently choose to collude. To understand the effects of the regulatory regime requires a deeper examination of the incentives it creates.

II. Regulatory Design and Allocative Harm

As noted in Chapter 1, the pharmaceutical industry is most commonly associated with the simplest model of the patent system. But in fact, in defining the incentives of pharmaceutical innovators, the regulatory scheme reflects a number of idiosyncratic choices.

The differences start with the most basic: the term length of protection. Pharmaceutical innovations enjoy longer-lasting protection than innovations in other industries, which partly offsets the time consumed by clinical trials. Under 35 U.S.C. § 156, the drug maker receives a one-year extension for every two years spent in clinical trials, plus the time spent in post-trial FDA approval, subject to the limitations that the extension may not exceed five years or leave a remainder exceeding fourteen years. The effective term is extended by another six months if the drug maker performs tests to evaluate the drug’s pediatric health benefits. And certain drugs treating “rare diseases or conditions” are outside even this highly modified scheme; they receive sui generis seven-year exclusivity under the Orphan Drug Act (Anand 2005).

The Hatch-Waxman bounty—a special bounty granted to certain generic drug makers that file ANDAs with Paragraph IV certifications—is another major difference. This Part explains how that feature of the regulatory arrangement widens the prospect for allocative distortion, relative to the usual patent regime. It does so, first, by ensuring
that a pay-for-delay settlement is (if legal) an attractive and feasible proposition for the innovator and generic firm. Second, the ability of an innovator to guarantee a bounty to a generic firm, an opportunity unavailable under litigation, is a significant noncash means to pay for delay.

A. The 180-Day Bounty

Under § 355(j)(5)(B)(iv), the first generic firm to file an ANDA is entitled, upon FDA approval, to a 180-day exclusive right to market a generic version in competition with the brand-name firm, effectively creating a duopoly during that period. (The “duopoly” characterization ignores the effect of authorized generics, discussed below.) The legal form of the exclusivity is a delay in FDA approval of any other firm’s ANDA with a Paragraph IV certification. Winning a patent suit is one route to exclusivity. For example, if an innovator’s generic rival secures a judgment that the relevant patents are invalid or not infringed, the FDA may approve the generic firm’s ANDA, freeing the firm to market its competing generic version, protected initially by the exclusivity period.

Winning a suit is not the only route to exclusivity. Exclusivity merely requires FDA approval of the first filer, which can be secured without litigation if the innovator declines to sue the first filer, as may occur if the innovator’s patent is very likely invalid or not infringed. For a time, the FDA resisted this reading of the statutory text, insisting instead upon a “successful defense” before granting exclusivity, but the agency abandoned the interpretation after its judicial rejection (FDA 1998). This result likely did not come as a great surprise to regulated parties. The end of the successful defense
requirement was a death foretold by Judge Harold Greene, who made clear in 1989 that the FDA’s argument was weak as a textual matter (Inwood Laboratories 1989).

The reward provided by the bounty is valuable, worth several hundred million dollars to a generic firm that successfully challenges the patents on a major drug. For example, Apotex reportedly earned between $150 million and $200 million from the exclusivity period on Paxil, a blockbuster antidepressant (Apotex 2004). The bounty thus provides a substantial inducement to challenge drug patents. A bounty-hunting generic firm will go on the attack if the drug is very valuable or the innovator’s patents very weak—likely invalid or not infringed—or both. (For a further examination of this decision, see Chapter 4.) With respect to very valuable drugs, the challenge is justified even if the ex ante likelihood of success is low. The more valuable the drug, the lower the threshold probability of success necessary to justify a challenge.

A generic firm can justify a challenge with just a one-in-five chance of success, provided that the innovator’s sales range in the hundreds of millions of dollars; the level of sales for a best-selling drug likely justifies a challenge with a prospect of success of just one percent. For a back-of-the-envelope calculation, suppose that a generic firm can expect fifty-percent market penetration during one-half of a year of protected duopoly, with a profit margin of two-thirds, and no profits otherwise. If entry has a probability $p$ of success, the innovator’s annual sales are $S$, and the generic firm’s entry expense is $10$ million, then its expected profits are $pS/6–$10 million. The generic firm breaks even provided that $pS > $60 million. Thus a drug with $300$ million in sales supports a challenge that is twenty percent likely to succeed. A drug with $6$ billion in sales supports a challenge that is one percent likely to succeed. It is therefore no surprise that
so many of the best-selling drugs have attracted challenges.

B. How the Singular Availability of the Bounty Facilitates Payment for Delay

The specific form of the bounty affects the feasibility of payments for delay. To see why, note first that a pay-for-delay agreement must satisfy two conditions to make practical sense for the parties. The first condition is a *gain from trade*: The patentee loses more under early entry than the alleged infringer gains. This condition is likely to be satisfied where the new entrant serves exactly the same market as the incumbent, for total duopoly profits are normally less than monopoly profits. In the limiting case, duopolists jointly achieve the same profit-maximizing price and quantity of a monopolist. In some settings, however, entry rather than deferral may lead to higher total producer profits, as when the entrant has superior access to a market, a unique means to price discriminate, or lower costs. Where entry increases total profits, the entrant can pay the incumbent for permission to enter (if it lacks an entitlement to do so) or, if licensing is unavailable, simply enter and then pay damages, provided they are not too high.

Competition between innovators and generic drug makers satisfies the gain-from-trade condition (Leonard and Mortimer 2005). Consider, for example, a generic firm’s challenge with respect to some blockbuster drug. Without entry, the brand-name drug maker might expect to earn, say, $10 billion in profits from U.S. sales during the drug’s remaining patent life. (Assume, for example, five years of remaining patent protection, $2 billion in U.S. profits per year, and a discount rate offset by profit growth.) If it loses a patent challenge, then it and the successful generic firm would
share duopoly profits for 180 days, with small profits thereafter once additional firms entered the market. In that event, $1 billion might be a plausible estimate of each firm’s profits. After all, typically, the innovator retains price-insensitive customers and may even raise prices somewhat, while the generic firm sells at a roughly thirty-percent discount (Morgan Stanley 2004; Grabowski and Vernon 1992).

If the parties reach a settlement ending the dispute and no other generic firm initiates a challenge, the joint gain from an entry-preventing agreement is $8 billion—the innovator’s $10 billion no-entry profit, less the $2 billion jointly earned under entry. If the two share the joint gain equally and invalidation is certain, the innovator would pay the rival $5 billion to induce the rival to abandon its suit. (After paying the settlement fee, the innovator would retain $5 billion in profits, a $4 billion improvement upon entry. The rival would enjoy a $5 billion profit, once again a $4 billion ($5 billion – $1 billion) improvement upon entry.) Purchasers would lose the $8 billion that is transferred to producers instead, plus billions more in deadweight loss from the resulting allocative distortion. If invalidation is uncertain, the stakes are lowered accordingly; a twenty-five percent chance of invalidation makes the expected gain from trade $2 billion, implying an equal-sharing payment of $1.25 billion.

The foregoing assumes an equal-sharing approach (Rubinstein 1982), which is customary for these analyses. It can be doubted, however, whether the generic firm’s $1 billion gain under competition ought to be considered as part of the alternative to settlement (the “threat point”) within an alternating-offers game such as Rubinstein’s (Sutton 1986). If the $1 billion is treated instead as an outside option rather than a threat
point, the relevant gain is $9 billion, and the payment $4.5 billion.¹ Not only does an agreement benefit the generic firm compared to its expected return from litigation (otherwise the generic firm would not agree), but in fact the generic firm does even better than it would have, had it won the suit.

The second general condition is that the settlement must offer an effective means to delay entry. If there are many potential challengers, and paying one merely attracts others, a payoff does little good. Even a cursory review of the mechanisms for generic competition, however, suggests that this condition will be satisfied in the pharmaceutical context. A firm must file an ANDA to be eligible for a settlement. The ANDA contains a demonstration by the generic firm that its proposed product is bioequivalent to the innovator’s drug, and that the firm is capable of making the proposed product. The challenge process requires a detailed description of the basis for belief of invalidity or noninfringement for each relevant patent of the innovator. To be a credible threat to the innovator, a generic firm must undertake these expenses (one generic firm cannot free-ride on another’s showing of bioequivalence) and be prepared to see the suit to conclusion. The number of firms capable of such action is limited.

It might appear that a large number of well-funded entities could credibly threaten to initiate challenges in order to extract payoffs, but multiple factors limit this possibility in practice. First, their very number would make it pointless to pay off just one of them. Second, the credibility of such a threat is undermined by the technical requirements involved in actually filing an ANDA, though the parties might be able to contract around this difficulty. Third, without the filing of a challenge, it is more

¹ The exact result depends on the bargaining model employed. Exploring the difference in outcome, depending on the model used, is a fruitful subject for future research.
difficult to establish that the resulting agreement is in settlement of litigation.

Moreover, the generic firms are not identically situated. The firms have differing views about the likelihood of success in a particular challenge, different information about the infirmities of an innovator’s patents, differing abilities to make a bioequivalent version of the drug, and different speeds in developing a noninfringing alternative, as well as different estimates of the drug’s future profitability. As a result, firms will have different incentives to bring a challenge. As evidence for this, it was not until 2003 (nineteen years after the establishment of the regulatory regime) that the FDA issued guidelines to deal with multiple filings on the same day (FDA 2003).

Once the first generic firm files an ANDA-IV, moreover, a sharp difference in incentives emerges between that ANDA-IV filer and all other generic firms, because only the first filer is eligible for the exclusivity period. Even if the first filer loses, withdraws, or settles, a subsequent filer does not become eligible for the bounty. (Whether a subsequent filer becomes eligible for FDA approval, a distinct issue, is discussed below.) FDA regulations issued in 1994 make clear that only the first-filed ANDA potentially delays the approval of subsequently filed ANDAs by operation of the 180-day exclusivity period, an interpretation revisited and endorsed once again in 1999 (FDA 1999). Although antitrust analyses of settlement have generally missed this point, two legal analyses of the Hatch-Waxman Act, Engelberg 1999 and Eisenberg 2001, do correctly identify this effect in passing.

The first-filer-only interpretation is not the only plausible interpretation of the relevant statutory provision, but it is a defensible one. The relevant provision states that if “a previous application has been submitted,” a subsequent filer must wait until 180
days after the “first commercial marketing of the drug under the previous application” or a favorable court decision, whichever is earlier. In essence, the FDA concluded that the only “previous” application that triggers the delay is a first application. The alternative interpretation is that any previous application can be a source of delay, not just the first. The FDA considered and rejected the alternative interpretation; though it did not explain its reasoning in detail, it did state that in the case where the first filer withdrew its application, its preferred interpretation was consistent with a goal of “encouraging prompt challenges.” A related policy justification is that having the first filer as a single “champion” encourages a potential challenger to file an ANDA as early as possible. Moreover, the reference to the previous application suggests contemplation of only a single previous filer, which supports the FDA view.

Likely the FDA also recognized that the alternative reading can produce anomalous results. If not only a first filer but also a second filer can be a “previous applicant,” then the 180-day period, as enjoyed by a second filer, would not restrict the approval of a first filer (from the first filer’s point of view, the second filer is not a “previous applicant” under any interpretation), making the subsequent filer’s exclusivity into an entitlement of an oddly truncated sort. It is possible that innovators and generic firms had doubts about the correctness of the FDA’s interpretation, but provided that they attached at least some probability to its correctness, the analytical point—that later filers have much smaller incentives—holds. In 2003, Congress codified the FDA’s interpretation when it passed the MMA.

The singular availability of the bounty has been underappreciated in the economic and policy analyses of settlement. Most cases and commentary ignore or blur
the difference between a successful first filer, which receives exclusivity, and a filer that is first to win a challenge, which may not receive exclusivity. One federal appellate case (*Tamoxifen* 2006) that rejected antitrust liability for a pay-for-delay settlement provides a useful illustration. There, the panel majority relied upon the erroneous view that bounty eligibility *does* cede to other filers. According to the majority, the innovator’s settlement agreement with the first filer, by neutralizing the competitive threat of the first filer, “opened the [relevant] patent to immediate challenge by other potential generic manufacturers, which did indeed follow—spurred by the additional incentive (at the time) of potentially securing the 180-day exclusivity period available upon a victory in a subsequent infringement lawsuit.” The majority apparently believed that, at least during the period of the FDA’s successful defense interpretation (that is what the panel means by “at the time”), exclusivity eligibility ceded to a later filer.

As a result, the court mistakenly attributed a practically nonexistent incentive to subsequent filers. That this error was apparently not challenged when first made in the district court, briefed or corrected during the appeals process, or noted by the panel’s dissenting opinion, demonstrates that the singular availability of the bounty, and its significance for antitrust analysis, is poorly understood. In 2010, the Second Circuit acknowledged this error in its earlier opinion (*Arkansas Carpenters* 2010), relying upon the argument discussed above.

The court’s more recent analysis made clear that this mistake is not merely technical, including this as one of several reasons why the full court might want to reconsider the pay-for-delay question en banc. The error was important to correct because a correct understanding of the exclusivity period is necessary to a proper
understanding of generic firm incentives.

Generic firms other than the first filer will lag behind in the approval process, if they have bothered to file at all; they will also be less motivated to initiate or vigorously pursue a challenge. The subsequent filers’ return on a challenge, aside from being smaller, depends upon the outcome of the first filer’s suit (and possible settlement), providing a strategic motivation to slow down until that uncertainty is reduced. Another possible difference among generic firms is that one filer may have a claim that it is uniquely able to exploit. The subsequent filer retains some incentive even without the exclusivity period, particularly as winning may provide a head start in marketing. However, each filer benefits from favorable judgments in the others’ suits, reducing the benefits from aggressive pursuit. A further complication is that a subsequent filer sometimes has an incentive for speed that the first filer lacks. The first filer receives the exclusivity whether it proceeds quickly or slowly (although the value of the exclusivity may decline over time); a subsequent filer receives a proportionately larger fraction of the rewards of normal generic entry by securing entry earlier.

It is therefore inaccurate to assert, as Judge Posner did in the *Asahi Glass* case, that “[i]n a reverse-payment case, the settlement leaves the competitive situation unchanged from before the defendant tried to enter the market.” The settlement does secure an important change in the competitive situation: it removes from consideration the most motivated challenger, and the one closest to introducing competition. Similarly, although it may be correct in a literal sense that a settlement “clear[s] the field,” as the Second Circuit concluded in its earlier opinion, the implication is very different from that drawn by the court: The most vigorous challenger has been removed
from the field, thereby removing an important source of early competition.

Settling with the firm that is closest to introducing competition and has the greatest incentive to do so is a highly profitable opportunity, even if subsequent filers remain free to secure FDA approval. In addition, the entry of subsequent filers can be blocked entirely in some instances, due to a statutory bottleneck created by the Hatch-Waxman regime. This approval bottleneck is related to, but distinct from, the first-filer-only restriction on the 180-day bounty.

As already noted, the 180-day exclusivity period operates by delaying FDA approval of a later-filing generic firm’s ANDA-IV. In particular, the statute requires that a later-filed ANDA-IV not be approved until 180 days after the first filer’s initiation of commercial marketing or a court determination of invalidity or noninfringement, whichever comes first. A settlement between the first ANDA-IV filer and the innovator removes an opportunity for commercial marketing or a court determination. Without the occurrence of either triggering event, the later ANDA-IV filer is stuck, for the FDA lacks authority to approve the application.

The resulting delay is frequently emphasized in discussions of the pharmaceutical regime (Hovenkamp et al. 2003; Brodley and O’Rourke 2002). The degree of delay should not be overstated, however, since the block is incomplete. If a later ANDA filer wins a favorable court decision, that decision triggers the exclusivity period—that is, the first filer’s exclusivity period. The subsequent ANDA filer could enter 180 days later. If the innovator declined to sue the later filer, as often happens, the generic firm would be obliged to file a declaratory judgment action. A further complexity arises when there are no subsequent filers to be blocked. That, however,
does not necessarily imply that there is no harm, since would-be filers may have been deterred by the futility of filing in light of the fact or likelihood of a blocking settlement.

Nor is the bottleneck a pervasive feature of pay-for-delay settlements, for two reasons. First, the bottleneck applies only to settlements reached during a limited time period. The bottleneck did not arise until the demise of the successful defense requirement, for under that interpretation a pending suit between an innovator and first ANDA-IV filer, not yet having been successfully defended, was considered insufficient to block approval of a subsequent ANDA-IV filer. (During the heyday of the successful defense interpretation, however, doubts about its validity might have affected decisionmaking to some degree, in anticipation of its invalidity once tested.)

Moreover, the bottleneck does not apply in the same way to drugs for which the first ANDA was filed after December 2003. For those drugs, the MMA replaced the bottleneck with a complex forfeiture scheme discussed in Chapter 3. In addition, some settlements do not take advantage of the bottleneck—for example, because the generic firm alters its filing in a way that removes the block.

The approval bottleneck is thus sufficient, but not necessary, to demonstrate the feasibility of pay-for-delay settlement or the presence of allocative harm. There is also a downside to overreliance on the bottleneck as the primary means to demonstrate the feasibility of a settlement that produces an allocative harm. The absence of an approval bottleneck can give the erroneous impression that there is no activity of competitive concern. Some courts, such as Tamoxifen 2006 and Cipro 2009, have been distracted in just this manner. Attention to limits on exclusivity eligibility, not just FDA approval,
better identifies the full extent of the allocative harm.

C. The Guaranteed Bounty as a Source of Compensation

The specific form of the bounty’s implementation expands the potential for allocative harm in a second way. To see this effect, consider an ordinary patent validity suit with some probability of a judgment of invalidity. To be concrete, suppose that the probability of a judgment of invalidity is fifty percent. If the parties see the litigation to conclusion, then consumers have a fifty percent chance of receiving the incremental benefits of competition, rather than facing a monopolist for the remainder of the patent term. (Assume for now that the generic firm is unwilling to launch a competing product without a favorable district court judgment. Circumstances where such a launch is feasible are considered below.)

Two different kinds of settlement are just as good as litigation from a consumer’s point of view. One settlement solution is simply to agree to decide by some random means, such as a coin flip, whether entry occurs. Another of equal effect is for the parties to divide up the remaining term in accordance with the probability of success. If the chance of success is fifty percent, then the patentee might agree to permit competition halfway into the remaining term. Consumers receive the full benefit of competition, but for one half of the period; that is equivalent to a fifty percent chance of enjoying the benefits of competition for the entire period, ignoring litigation costs and changes in market conditions. In this setting, each outcome—a lawsuit with a probabilistic outcome, a randomized settlement, and a settlement that splits up entry in accordance with the probabilities—has the same effect upon expected patentee profits,
entrant profits, and consumer welfare.

As discussed in Part I, an agreement that divides up the remaining term into monopoly and competition periods is generally accepted. The underlying model on which these conclusions are premised, however, fits pharmaceutical regulation poorly. In suits involving an ANDA-IV filer, a division-of-term settlement and a probabilistic lawsuit are not equivalent. Providing a generic firm with fifty percent of the remaining patent term is not the same thing as a fifty percent chance of winning the suit—not for the generic firm, innovator, or consumers. The key source of profits for a generic firm is the exclusivity period. Rather than monopoly followed by general entry, there is an intermediate stage of duopoly between the two. This feature is not reflected in the standard model.

Key to the difference is an important feature of the Hatch-Waxman regulatory arrangement: If the parties agree to a negotiated entry date, the generic firm enjoys the exclusivity period when it finally enters the market. This result follows directly from the approval bottleneck discussed in Part II.B. That section demonstrated how a first-filing generic firm could retain its exclusivity eligibility, despite settlement. As noted, so long as the settling generic firm stays out of the market, later filers are denied FDA approval. In addition, once the generic firm does enter, it makes good on that eligibility, and enjoys the 180 days of exclusivity.

This effect of the statute holds true in the same set of important though limited situations in which the approval bottleneck can delay FDA approval of later ANDA-IV filers. By that, I mean those settlements reached after the demise of the successful defense requirement, where the relevant ANDA was filed prior to the MMA’s rule
change in December 2003. For settlements reached during the successful defense period, moreover, this feature might still be potentially relevant, if the anticipated demise of the successful defense requirement affected the terms of settlement.

By making the bounty a certainty rather than a probability, the innovator confers value upon the generic firm. That opportunity to confer value disrupts the equivalence between litigation and a term-dividing settlement. (Bulow 2004 hints at this point but does not develop it.) The disruption is most easily seen by considering two distinct aspects of the settlement negotiation.

First, it is costly to the innovator to allow the generic firm to enjoy the bounty with certainty rather than merely a probability. The innovator will accept a settlement only if the entry date is set late enough to compensate the innovator for the value thereby transferred to the generic firm. On average, that date leaves consumers with less benefit than they would receive through litigation.

To see this, it is helpful to consider a stylized model of the dynamics of negotiation. Consider a market served by an innovator, who is equipped with a single patent granting ten years of exclusivity, and by generic firms, exactly one of which initiates a challenge to the patent. The innovator and the generic firm litigate or negotiate to determine the division of profits for the remainder of the patent term. If the parties litigate, there is a trial, and the patent is found valid and infringed with some probability—say, to continue with our maintained assumption, fifty percent. If the patent is found valid and infringed, the generic firm is barred from entry, and the monopolist enjoys monopoly profits for the remainder of the term. Otherwise the generic firm enters immediately, leading to two stages of competition: an exclusivity
period set by statute, during which the innovator and generic firm each earn duopoly profits; and a residual period during which other firms can enter as well, and the two firms earn much lower profits.

The parties can choose to settle rather than litigate by agreeing upon the date of entry by the generic firm. Entry after negotiation resembles entry after litigation: There is a duopoly period followed by a residual period of competition. Entry after negotiation is certain, rather than probabilistic. Moreover, if the negotiated entry date is late enough, there is no final competition period, but instead monopoly followed by a truncated duopoly period. Suppose further that the parties decide whether to litigate or settle at the beginning of the ten-year period, and any agreement or trial is concluded instantaneously.

A few numerical assumptions ease the exposition. Suppose that under monopoly, the innovator receives 1000 each year, the generic firm and consumers nothing; that under duopoly, the innovator and generic firm each receive 500 per year, and consumers again nothing; and that under competition, consumers receive 1000 per year, and the innovator and generic firm each receive nothing. Think of each unit as a million dollars—$1 billion per year for the innovator under monopoly, and so forth—and the example roughly matches the magnitudes for a blockbuster drug.

These assumptions are unrealistic in two respects. First, the model assumes that total duopoly profits equal monopoly profits. By contrast, under most models of competition, producer surplus drops under duopoly compared to monopoly, and consumer surplus rises. This is a variation on the point already made, that duopoly profits are lower than monopoly profits. Pharmaceutical duopoly does tend to
approximate monopoly profits, but the more important point is that the polar assumption serves to elucidate the effect presented in the text. Second, the model assumes that firms earn no profits once full entry commences. But as acknowledged above, firms often enjoy some profits once the duopoly period has ended. These profits, if large enough, undercut the effect discussed above.

Under litigation, the innovator has a fifty percent chance of receiving 10,000 in monopoly profits and a fifty percent chance of receiving 250 in duopoly profits, an expected value of 5125. The generic firm has a fifty percent chance of receiving 250 and a fifty percent chance of receiving nothing, an expected value of 125. Consumers have a fifty percent chance of receiving 9500 (1000 per year for nine-and-a-half years; the first half-year is the duopoly period) and a fifty percent chance of receiving nothing, an expected value of 4750. This is depicted graphically in Figure 4. The length of the rectangle is ten years, and its height shows the division of expected benefits within a period.

Now consider settlement. Under settlement, the generic firm receives 250 with certainty, because the bounty is now guaranteed. The additional 125 to the generic firm, compared to litigation, must come from somewhere. The innovator also receives 250 during the duopoly period. To be indifferent between settlement and litigation, the innovator must earn at least 4875 during the monopoly period. That level of profit can be earned provided that entry begins 4.875 years into the remaining patent term or later. This result is depicted in Figure 4. Consumers, in order to equal their benefit from litigation of 4750, require that the entry date be no later than 4.75 years; assuming that entry date, consumers begin to receive 1000 per year six months after entry, or
beginning at year 5.25. If the entry date is 4.875 years, the level insisted upon by the innovator, consumers are worse off by 125 under settlement compared to litigation.

Moreover, the actual negotiated date of entry is likely to be substantially later than the threshold date that leaves the innovator indifferent between litigation and settlement. The innovator will bargain with the generic firm over the gains conferred by making the bounty a certainty. Securing a later entry date is very important to the innovator. For the generic firm, an earlier entry date is better, given the higher present value of earlier payment, but only modestly so. Enjoying the exclusivity period with certainty is more important to a generic firm than its timing. In fact, if future market demand is anticipated to increase, a generic firm might prefer the later entry date, so long as the increase in projected profits exceeds the discount from the delay in their receipt.

The innovator is likely to bargain not for a settlement that perfectly matches its profits under litigation, but for a more profitable settlement—that is, one with a later entry date. The generic firm is likely to agree, so long as it secures the duopoly period with certainty rather than having to take its chances in litigation. Suppose, for example, that the innovator and generic firm agree to an entry date nine years into the remaining patent term—that is, a year before expiration. Now the innovator earns with certainty nine years of monopoly profits (9000) plus 250 from the duopoly period; the generic firm earns 250 with certainty; and consumers see competition only in the last six months, for a total benefit of 500. This result is depicted graphically in Figure 4. Indeed, this is not even the latest entry date to which the parties might agree.

The assumptions of the stylized model are unrealistic, particularly with respect
to the generic firm, which normally earns some profit during the competition period, and hence has some reason to prefer earlier rather than later entry dates. Yet the simple depiction here is sufficient to show the problem for consumers from no-payment settlements—an innovator will be unwilling to accept any entry date that would leave consumers at least as well-off, and the date the innovator actually chooses is even worse for consumers. Delayed entry can thereby align the incentives of the innovator and generic firm, a point generally overlooked, and contrary to the view of Hovenkamp et al. 2003, which argues that delayed entry “does not align the incentives of pioneer and generic litigants: Generics will want the delay to be as short as possible, and patentees to make the delay as long as possible.”

D. The Complication of Litigation Expense

Considering litigation expense does not eliminate these allocative harms, and may, in fact, exacerbate them. To see why, it is useful to consider two respects in which saved litigation expense is thought to count in favor of settlement.

First, and as noted in Part I, saved litigation expense is thought to offset the allocative harm from the settlement. But although litigation expense is large in absolute terms, perhaps tens of millions of dollars, its size is dwarfed by the hundreds of millions or billions of dollars reallocated when parties enter a pay-for-delay settlement. The savings are insignificant except in the least important cases. Aside from its small role in any realistic assessment of the welfare effects of a settlement, saved expense is also an unlikely explanation of the parties’ motivation for entering the settlement.

Second, even those who favor antitrust liability for pay-for-delay settlements,
such as Hovenkamp et al. 2006 and Shapiro 2003b, make an exception for settlements with payments keyed to the size of litigation expense. Similarly, as a matter of practice, in decisions such as Schering 2003, the FTC effectively has granted safe harbor to settlements in which the innovator makes a payment equal to or less than saved litigation expense.

By differentiating pay-for-delay settlements that include large cash payments from those with payments that are equal to or less than saved litigation expense, the safe harbor usefully distinguishes those settlements likely to inflict the largest allocative harm. But the policy nevertheless permits some settlements that inflict allocative harm. That is true for two reasons. The first reason is an extension of the zero-payment settlement analysis of the previous section. Suppose, for example, that the innovator saves no litigation expense by settling. In that case an entry-splitting settlement that includes no cash payment is identical to the settlement discussed in the previous section. It fits within the safe harbor, yet entails an allocative harm.

Now suppose that the innovator saves some litigation expense by settling, but that the generic firm’s bargaining power is such that it is able to extract all of the benefit from the innovator’s saved expense. In that case, nothing has changed; a settlement that includes a payment equal to that saved expense is equivalent to the zero-payment settlement where there are no litigation savings.

If the innovator has some bargaining power, however, the safe harbor permits additional allocative harm. In that case, the innovator will be able not only to retain part of the gain from saved litigation expense, but also to bargain for part of the generic firm’s litigation savings. If the innovator has at least equal bargaining power, it should
need to pay no more than half of the difference between the parties’ saved litigation costs in order to secure a settlement. Allowing a larger payment, as the safe harbor does, permits the innovator to confer additional value upon the generic firm in exchange for additional delay, leading to additional allocative loss. Indeed, if the innovator has most of the bargaining power and the generic firm’s saved expense is large enough (it need not be as large as the innovator’s savings), the litigation savings component of the deal, considered alone, requires a net conferral of value from the generic firm to the innovator. In that case, the generic firm will not pay the innovator; instead, the parties will simply agree to a later entry date, thereby imposing a greater allocative harm. The problem is compounded by the potential for manipulation, as the innovator could inflate its cost estimate in order to permit a larger payment insulated from antitrust scrutiny.

III. Regulatory Design and Congressional Judgment

Part II demonstrated how the operation of the 180-day bounty expands the opportunity for allocative harm from settlement. The foregoing analysis establishes that the allocative harm of settlement extends to a wider range of settlements than generally recognized. Problematic settlements are feasible even where there is no formal bottleneck to FDA approval, because buying off the single firm with bounty eligibility carries a strong prospect of allocative harm. Settlements with small cash payments, moreover, can nevertheless entail payment for delay. Even where there is no cash payment, a term-dividing settlement provides the opportunity for an innovator to provide noncash compensation—the guarantee of the bounty itself—in exchange for delay.
This Part considers the objection, described in Part I, that the expected allocative losses from a pay-for-delay settlement ought to be tolerated. After all, these agreements settle litigation—and normally settlements are thought desirable, because they conserve litigation expense and benefit parties who are in the best position to arrange their own affairs. Moreover, the litigation settled is patent litigation, and patent policy favors innovation over consumer access; the interaction of patent policy with antitrust might be thought to permit allocatively harmful practices ordinarily condemned under antitrust law alone.

Here we come to a second mode of economic analysis of industry-specific regulation: unpacking a congressional judgment about the proper balance between innovation and competition. This judgment, like the judgment about innovation policy reflected in the Patent Act, may influence the scope and vigor of antitrust enforcement. For example, patent policy may contain a norm favoring innovation and favoring settlement that alters the antitrust treatment of practices involving patented goods. Even if patent policy generally contains such a norm, an industry-specific regulatory arrangement supplants that norm within its domain. To understand the alteration, it is necessary to understand in some detail how the regulatory regime differs in its effects from the usual effects of patent law.

This Part examines those differences and their relevance for antitrust enforcement. Part III.A motivates that examination, by first presenting why we might recognize, thanks to patent law, certain exceptions to the ordinary operation of antitrust law. Parts III.B and III.C show how the Hatch-Waxman Act alters the incentives set up by patent law, in essence setting up a tax on some drug development projects and a
subsidy for others. Part III.D explains how these alterations, understood as Congress’s industry-specific judgment about the proper balance between innovation and competition, provides a basis for special-purpose antitrust law.

A. The Uneasy Case for Patent Exceptionalism

If patent policy depends upon above-cost pricing, and antitrust policy is suspicious of firm practices that defend and extend above-cost pricing, then there is a case to be made for a reconciliation of means in which antitrust gives way, and the patentee is allowed to employ certain practices that would otherwise be prohibited. To make headway, it is useful to consider first whether antitrust law of its own accord provides a special accommodation to the makers of innovative goods, and then to assess whether the Patent Act alters the baseline of enforcement for patented goods.

Consider, first, innovation as an internal norm of antitrust. Such a norm may at first seem foreign to antitrust law. After all, low prices are an important goal of antitrust enforcement—even, some have claimed, the primary goal (Edlin 2002). There are important areas of antitrust doctrine in which low consumer prices trump other efficiency-promoting values. For example, under current U.S. doctrine, cost savings achieved through a merger are generally not cognizable unless they are “sufficient to reverse the merger’s potential to harm consumers in the relevant market, e.g., by preventing price increases in that market” (Department of Justice and FTC 1997). In addition, the Supreme Court has repeatedly invoked “consumer welfare” as the touchstone of antitrust analysis (Brooke Group 1993; NCAA 1984).

However, allocative efficiency does not exhaust the concerns of antitrust
analysis. Promoting innovation matters, too. Some innovation-promoting antitrust rules may have only a minimal conflict with allocative efficiency—for example, when an antitrust enforcement agency insists upon the maintenance of rivalrous research and development efforts as a condition of merger (for examples, see Katz and Shelanski 2005). A greater conflict is posed by a policy that advocates market concentration as an inducement or (more controversially) a platform for innovation (Schumpeter 1950).

Basic structures of antitrust doctrine reflect the need to provide a reward for innovation, even at some expense of allocation. As a general matter, monopolies are subject neither to dissolution by government decree nor to a duty to provide access to rivals at a discounted rate (Trinko 2004). Nor are product design decisions normally subject to disclosure to rivals, though disclosure would improve the rivals’ ability to compete in the provision of complementary goods (Berkey Photo 1979). A contrary policy would lower prices in the short run but reduce the prospective incentive to invest in new and improved products and processes, an important engine of economic growth. This dynamic benefit of policies that preserve monopoly profits offsets their static allocative cost. As an ordinary element of antitrust law, consumer access is balanced against the incentive to create.

The difficult question is how far to push the argument for dynamic efficiency. The higher the producer profits allowed, the larger the dynamic benefits. An agreement with a rival to divide markets normally attracts condemnation under section 1 of the Sherman Act. But an innovator might argue that the additional profits induce enough incremental innovation to make the practice beneficial overall. The argument is fundamentally similar for patented and unpatented (though costly-to-create) goods. An
innovator who builds a telecommunications network and one who designs a new drug are similarly positioned to argue that a certain profit-improving practice should be permitted, despite its adverse allocative consequences, in light of its salutary effect upon the incentive to innovate. The tradeoff inherent in providing incentives for creation while tolerating allocative distortion affects intellectual property and other assets alike (Elhauge 2003).

An argument favoring exemptions for innovative goods, however, likely fails as a matter of general antitrust law. It is difficult to establish convincingly that an exemption carries large benefits for future innovation. Nor is a generalist court equipped to make the necessary fine-grained determinations of industrial policy, relaxing antitrust here and tightening it there, in accordance with its views about desirable innovation and acceptable deadweight loss. Certainly such case-by-case determinations of incremental innovation and incremental, deadweight loss are projects ill-suited to the capacities of a generalist court. There is, therefore, often good reason to limit attention to allocative efficiency in practice, even if one is committed to a full range of efficiency arguments—including dynamic efficiency—in theory (Posner 2001).

Even without such a norm, the Patent Act provides a statutory foothold, external to antitrust law, for a patentee to insist upon a more innovation-protective antitrust policy than that available to innovators generally. There will not, of course, always be a conflict between antitrust law and patent policy. To the extent that the Sherman Act already reflects an acceptance of dynamic arguments, there may be no conflict in means. But often there will be a conflict, and in those cases the Patent Act provides a
basis for seeking an exception to the ordinary operation of antitrust.

The high-water mark in judicial recognition of patent exceptionalism is the Supreme Court’s holding in *General Electric* (1926) that a patentee may agree to a price-restricted license with its competitor. The extent of or rationale for exceptionalism is often left undeveloped. This is a problem in *General Electric* and other old cases, but the modern pay-for-delay cases fare little better. They are sprinkled with statements that, for example, antitrust liability should be withheld for “a rather simple reason: one of the parties owned a patent,” and that “[b]y their nature, patents create an environment of exclusion, and consequently, cripple competition” (*Schering* 2005). Such ipse dixit, if taken seriously, might justify a kind of naïve exceptionalism in which a court simply notes the conflict between antitrust and patent law and concludes against antitrust liability without further analysis.

A more sophisticated version of exceptionalism ties the contemplated exception to a specific provision of the Patent Act or to a policy closely related to its provisions. Such statute-oriented specificity emerges from the Supreme Court’s instruction in *Simpson v. Union Oil Co.*, explaining the rule of *General Electric*, that “[t]he patent laws which give a . . . monopoly on ‘making, using, or selling the invention’ are in pari materia with the antitrust laws and modify them pro tanto” (*Simpson* 1964). This version of *in pari materia* emphasizes that when two statutes govern the same activity, they must be reconciled by some means. In making that reconciliation, the Patent Act has a claim to primacy, as Congress’s more specific take upon how best to balance innovation and consumer access with respect to patented goods.

*Simpson* refers to the specific rights provided by the Patent Act—the exclusion
with respect to making, using, and selling, and a related right to license—not a general policy favoring patentee profit-taking. The necessity of specific statutory support also is indicated by the Court’s insistence elsewhere that exceptions created by the Patent Act must be “strictly construed” (Masonite 1942). Such constraints have prompted the recognition, for example, that a patentee enjoys no exception for restrictive practices that cover products not within the scope of the patent or that extend beyond its duration (Valley Drug 2003).

Without an explicit statutory provision to rely upon, a patentee claiming an exception may instead seek refuge in the innovation-protective policy of the Act. Yet every profit-enhancing practice of a monopolist, however damaging to allocation because of its effect on prices, might be defended on the ground that it increases innovation. As a way to cabin such an argument, it is helpful to consider what we might call the innovation efficiency of the practice, the ratio of incremental innovation to incremental deadweight loss produced by the practice, presented in Kaplow 1984 and used by others (Scotchmer 2004; Fisher 1988; Klemperer 1990; Gilbert and Shapiro 1990). The ratio gives shape to the Supreme Court’s declaration, made in the context of considering blanket licenses in copyright, that “we would not expect that any market arrangements reasonably necessary to effectuate the rights that are granted would be deemed a per se violation of the Sherman Act” (BMI 1979). Where a practice produces a large deadweight loss without much benefit for innovation, it will be more difficult to understand the arrangement as reasonably necessary to effectuate the Patent Act’s innovation policy, and the practice will be more vulnerable to antitrust condemnation.

The innovation interest is not limited to the patentee. An alleged infringer may
be an entrant also engaged in innovative activity. Identifying and negotiating with every patentee that holds rights that are possibly relevant to the entrant’s product is costly for the entrant, particularly in industries where innovation is cumulative. Identifying relevant patents is discouraged in practice, moreover, by the specter of enhanced damages for willful infringement, an outcome thought to be made more likely by prior awareness of relevant patents (Lemley and Tangri 2003). The likely outcome is that an entrant will frequently stumble into patent infringement suits in which it finds itself a defendant.

Seeing the litigation to conclusion is unlikely to be an attractive option for the defendant. Often, winning the litigation will be unrewarding for the entrant, due in part to a free-riding problem discussed in the next section. Yet a rule that prohibits all settlements that work an allocative harm will render some settlements unavailable. If all of the resulting confrontations must lead to a full adjudication of the patent, the result might be to reduce the supply of innovative entrants (Rosenberg and Shavell 2004). There is reason, therefore, to accept a certain amount of settlement, even settlement that works an allocative harm, in order to maintain incentives for a potential infringer’s innovative entry. (Even when the resolution of the suit forces the alleged infringer to exit the market, moreover, the limited period prior to exit is a source of some consumer benefit.)

Patent exceptionalism has sharp critics. The concept runs contrary to the enforcement agencies’ expressed view that “for the purpose of antitrust analysis, the Agencies regard intellectual property as being essentially comparable to any other form of property” (DOJ and FTC 1995) and to the government’s longstanding opposition to
A forceful argument can be made, too, that patent law at most confers rights of exclusion and enjoyment that match but do not exceed those enjoyed by owners of tangible property, and if so, exceptionalism is unwarranted (Melamed and Stoeppelwerth 2002). The present purpose is not to argue patent exceptionalism’s merits, but merely to note its possible basis in statute and precedent. Provided that paying for delay is within the scope of a Patent Act policy, patent exceptionalism provides a potential, and to some courts a persuasive, basis for insulating the practice from antitrust attack.

B. The Bounty as a Tax on Innovation

The previous section identifies some statutory basis for treating patentees differently under antitrust law. But patent law and antitrust law are not the only means by which innovative monopolists are regulated. Antitrust is *in pari materia* not only with patent law, but with industry-specific regulation as well. A reconsideration of the applicability of patent exceptionalism to pay-for-delay settlements in the pharmaceutical industry begins with an examination of the innovation and competition policy embodied in the Hatch-Waxman Act, compared to the treatment of patented goods generally.

That examination requires an investigation of the economic effects of the Act’s principal components. That investigation receives no assistance from legislative history, which is too scant to provide even arguable use here. The main source of such history is a House report accompanying an early version of the Act, but the key 180-day exclusivity period became law without informative discussion in that report and without
debate (House of Representatives 1984). (There is no comparable Senate report.) Moreover, it was apparently not contemplated at the time of passage that the regulatory scheme would facilitate collusion to the extent discussed above.

An important component of the innovation and competition policy of the Hatch-Waxman Act is the bounty provided by the 180-day exclusivity period. Without a bounty, the incentive to challenge patents is often much reduced. Normally, defensive nonmutual issue preclusion permits firms other than the original challenger to take advantage of a favorable legal judgment without repeating the time and expense of a suit (*Blonder-Tongue* 1971). If a favorable judgment is the only impediment to entry, then potential challengers will face a serious free-rider problem. Not only will a firm fail to internalize the full benefits of its challenge, since others can use the judgment as well, but in addition the gains will tend to be rapidly dissipated, as other firms enter and compete away the benefits of the favorable judgment. This result has led commentators to conclude that patent challenges are underprovided, both in the decision to bring a challenge and in the incentive to pursue it vigorously (Miller 2004; Thomas 2001). The bounty provides a substantial boost to the incentive to challenge.

Is a bounty really necessary? After all, it is possible that the beneficiaries of a public good could agree in advance to contribute to its provision. To make that work here, however, there would need to be some means to limit post-provision rivalry. As Lemley and Shapiro 2005 note, however, such coordination will be difficult. An agreement on post-judgment prices is likely to raise antitrust concerns. In addition, the agreement will be less effective to the extent that the incumbent whose patent is being challenged remains in the market, but subject to the quantity or price strictures of the
The bounty’s importance as an inducement to challenge varies with the type of challenge. Issue preclusion has an important effect where the absence of a favorable judgment is all that stands in the way of entry. This is particularly true of an invalidity challenge. It is true also of noninfringement challenges that establish a route of production available to many firms. For example, a district court might arrive at a narrow construction of patent claims, resulting in a clear, noninfringing, widely available route to offering a bioequivalent drug. In other cases, however, the noninfringement route pursued by the generic firm is not readily available to other firms, because it is difficult to accomplish or separately patentable. In that event, the bounty, though still valuable to the generic firm, may be less necessary as an inducement to trigger suit.

Consider, for example, K-Dur, the drug at issue in an antitrust challenge brought by the FTC, and discussed in the Appendix. K-Dur is no blockbuster; its sales at the time of settlement were $190 million. Its active ingredient is an unpatented potassium salt used to replace an electrolyte lost from the body as a side effect of certain anti-hypertension drugs. K-Dur’s advantage is a special patented coating that permits controlled release of the active ingredient. K-Dur is backed by a patent that, like any patent, is “probabilistic” and imperfect (Ayres and Klemperer 1999; Lemley and Shapiro 2005). But the source of patent weakness is different. For K-Dur, there is a significant opportunity to argue noninfringement, rather than invalidity—assuming, that is, that the filer can in fact come up with an alternative, noninfringing means of achieving bioequivalence.
This is exactly what happened with K-Dur. A generic rival concluded that it could manufacture a bioequivalent controlled-release product without infringing the patent. In particular, it contended that its product had a composition and viscosity different from that specified in the innovator’s patent. The likelihood that some generic drug company will be able to do this may be fairly high; if it does so, it is that expertise, which may itself be protected by a patent, that forms part of the generic firm’s ability to compete. This approach is less vulnerable to free-riding, less subject to a flood of profit-dissipating competitors, and less needful of the 180-day exclusivity to protect its bid for entry.

C. Other Hatch-Waxman Act Provisions as a Subsidy to Innovation

While the Hatch-Waxman regime promotes pre-expiration competition by means of litigation, a second set of provisions provides innovators with protection from pre-expiration competition. While a generic rival could in theory evade these regulatory delays by filing a full-blown NDA instead, including the safety and efficacy studies, typically this will not be worth the time and expense.

First, under § 355(j)(5)(F)(ii), if the innovator’s drug contains a novel active ingredient—more precisely, if it is a “new chemical entity,” a drug containing no “active moiety” already approved in another NDA—the FDA must not accept an ANDA-IV in the first four years after NDA approval. The delay is five years for ANDAs with Paragraph I, II, or III certifications. This delay, sometimes referred to as data exclusivity, is quite valuable, unless the drug holds so little future promise, as evaluated during the first few years of marketing, that a generic firm would not
otherwise have sought to initiate a challenge earlier than the four-year point. For other new drugs, there is an analogous delay of approval (not ANDA submission) of three years.

Second, ANDA submission triggers an initial, ministerial review by the FDA governed by regulation (21 C.F.R. § 314.101(b)(1)), normally completed within sixty days. The review is to confirm that the ANDA is sufficiently complete to permit substantive review. Upon completion of the review, the FDA notifies the ANDA-IV filer that its application has been received. This is brief, but hardly trivial, since a single month’s respite from competition may allocate hundreds of millions of dollars. Upon the completion of initial review, the generic firm sends notice of its filing to the innovator.

If the innovator initiates a patent suit, further delays ensue. One source of delay not unique to pharmaceuticals is the duration of the patent suit, which normally takes several years but can take longer, particularly in the hands of an innovator committed to drawing out the proceedings. The pharmaceutical innovator, compared to a patentee in another industry, receives additional protection during the pendency of the suit: an automatic stay of FDA approval.

The stay takes effect provided that the brand-name firm files suit within 45 days of receiving notice of the Paragraph IV certification. The stay, provided for in § 355(j)(5)(B)(iii), lasts for at least the first thirty months after the innovator’s receipt of notice. Under certain circumstances (see § 355(j)(5)(F)(ii)), the stay can last for more than three years. The district court can also lengthen or shorten the stay in response to uncooperative behavior by either party. If the suit drags on too long, the stay will
expire. The stay superficially resembles the preliminary injunction ordinarily available to patentees, but the pharmaceutical innovator need not show irrevocable harm or likelihood of success on the merits, nor post a bond from which the alleged infringer’s damages are paid if the patentee subsequently loses. As a result, not only is the stay automatic, but its expected cost is much lower than that of an injunction.²

The several years’ delay caused by the stay is an important source of profits where a generic firm would otherwise enter prior to the district court’s judgment. A generic firm would sometimes prefer not to “launch at risk,” even if permitted to do so; if a court eventually concluded that the innovator’s patent was valid and infringed, the generic firm would be responsible for lost profits. But if the generic firm’s likelihood of winning is sufficiently high and the discount at which it must sell to compete sufficiently slight, launching at risk will be an attractive strategy.

For example, suppose that the patent is valid and infringed with probability \( p \), and that entry takes the simple, unrealistic form of stealing share from the incumbent by selling at a discount. The incumbent earns a margin \( m \) on each unit; the entrant earns \( m' \). Entry implies a gain of \( m' \) on each unit but damages of \( m \), payable with probability \( p \). Entry is profitable provided \( p < m'/m \).

This analysis does not factor in the bounty, which may incline a generic firm toward caution, since it can wait for the district court to rule, then enjoy the bounty with less risk of paying damages. (Eliminating the risk entirely requires waiting until the conclusion of the appeal.) Factors favoring earlier entry include the time value of money, the risk of a declining future market for the drug (particularly if a competing

² This timing is usually not affected by the FDA’s evaluation of the ANDA to confirm compliance with its requirements, a process that runs in parallel with the timeline described here. This process frequently takes two years or more (Czaban 2006).
therapy is likely to become available), and the benefit of surprise in dealing with a threat from authorized generics (see the conclusion of this chapter for further discussion). Finally, a later ANDA filer may force the first filer’s hand, for a later filer’s victory triggers the first filer’s exclusivity period.

As matters stand, launches at risk do occur when the litigation has dragged on for so long that the stay expires, and the benefits of launch seem worth the expected costs if the generic firm later loses. Such launches were rare in the early years of the Hatch-Waxman Act (Dickinson 1999), and have been described as the exceptional case (Shapiro 2003a), but in fact such launches have brought early competition to major drugs such as Allegra, Neurontin, Paxil, Plavix, and Wellbutrin SR. More launches at risk would occur absent the stay.

Moreover, these are also the cases where an innovator would be least likely to secure a preliminary injunction, or would be responsible for the largest damages if it did secure an injunction and then lost the subsequent patent suit. A patentee’s decision to secure a preliminary injunction (if it can) resembles an entrant’s decision to launch at risk, in that each faces an expected penalty based upon the likelihood of losing the suit and the size of the other’s damages that must be reimbursed in the case of a loss. The two are dissimilar, however, in the key respect that seeking a preliminary injunction is here always profitable. The innovator’s profits saved are larger than the generic firm’s profits foregone, so that even if the patentee thought its loss certain, a preliminary injunction would still be desirable from the patentee’s standpoint. Ascertaining the proper level of damages, however, is a difficult question.

Taken together, the delays set up by the Hatch-Waxman Act provide an
important means for innovative drug makers to preserve the returns upon a new drug. For a new chemical entity backed by a patent, the delays provide about seven years of protection after the product is approved. Even if the drug were protected by *no* patent but had a new active ingredient, the delays would still secure about six years of protection: Without a patent to challenge, the generic firm cannot file an ANDA-IV, and therefore must wait five years before its ANDA is accepted, and likely another year or more for FDA approval.

A drug without a new active ingredient, like K-Dur, enjoys several years of protection, even if a challenge is immediate. Moreover, these figures understate the effect of delay enjoyed by an innovator. A drug must cross a certain threshold of profitability before a generic firm will find it worthwhile to prepare and file an ANDA-IV and then defend the ensuing patent suit. If a drug takes time to build demand, the generic firm will wait to file its challenge, and a substantial part of the delay is effectively held in reserve until that challenge occurs.

The combined effect of the tax and subsidy reflects contrary forces. Consumer access is promoted by the unique incentive to challenge patents. Innovation is supported by the term extensions, initial delay based upon data exclusivity, and automatic stay. But the two forces cannot readily be summed in an across-the-board manner that applies uniformly to all drugs. The combined effect is not functionally equivalent to a decrease or increase in the patent term. Increased competition is the more important factor for some drugs, increased innovation the more important factor for others. The overall result is a pivot in the reward structure—a relative increase in the returns on some drugs and decrease on others.
The factors determining the balance for a particular drug are its market importance, the likelihood that an innovator’s patent would be found invalid or not infringed if challenged, and the extent to which other challengers could take advantage of the judgment absent the exclusivity period. A typical blockbuster and K-Dur illustrate the alternatives. For some drugs, it is the increased threat from competition that predominates. This is likely the case for most blockbusters. For a popular drug with a patent covering a novel active ingredient, an invalidity challenge is economically feasible due to the large bounty prospect, despite the free-rider problem and the low likelihood of success. The delays dampen the effect to a substantial extent, but the overall effect is a reduction in reward. For example, a drug that earns the innovator $1 billion per year without competition and nothing otherwise, for which at least seven years of patent term are remaining upon its approval, and which has a fifty percent likelihood of losing its patent suit against a generic rival, has expected profits that are $3.5 billion ($1 billion per year $\times$ 7 years $\times$ 50 percent) higher than would be the case under immediate entry.

For other drugs, it is the increased protection from competition that predominates. For a drug faced with an infringement challenge not readily replicated by other generic firms, the bounty is less necessary to induce a challenge. If the challenge would have occurred in any event, the major effect of the regime is to protect the innovation for several rewarding years before subjecting it to potential competition.

The variation across different drugs may achieve, in a rough manner, an efficient balance between innovation and access across a range of drug development projects. With respect to a drug like K-Dur, increased protection may be a necessary inducement
to invest, since such a drug is highly vulnerable to the noninfringing results of reverse engineering, which may be initiated once the drug’s commercial success is established. The initial exclusivity period, slow adjudication, and the automatic stay protect the profits on such a drug for a limited period. The stay is particularly important, given the likely attraction of launching at risk. This protection helps justify the drug’s development and approval expense.

With respect to blockbusters, patent-busting might be unusually beneficial to consumers, relative to patent-busting on other drugs. This would be true if blockbusters have an unusually large amount of demand at lower price levels, relative to other drugs. Such demand might result if popularity spawns widespread market awareness, or because treatments that manage chronic conditions—as most blockbusters do—have a large number of consumers with low willingness to pay. The argument assumes (as seems reasonable) that the firm cannot easily price discriminate among consumers.

In that event, the consumer benefit from subjecting these drugs to early competition is unusually high, and the decentralization of the challenge scheme is an attractive feature; entrusting the early-competition decision to the government would create a risk of capture by interested parties. The size and scope of the reduction in the incentive to innovate, moreover, depends upon the degree to which the innovator knows in advance whether the project, if successful, is likely to be a big success that would attract a challenge. If a drug maker never has any advance warning, then the dampening effect on innovative incentives will be spread thinly across all drug development projects. But to the extent the innovator can anticipate success, the tax on innovation will be borne primarily by the projects that are prospective blockbusters. To the extent
that such projects have not only a high value conditional on success but also a high expected value, the tax will have less of a deterrent effect upon innovation.

Can the innovator anticipate success? Some evidence of awareness of future promise is provided by the prevalence of multiple drug development projects, running in parallel, which exploit the same chemical pathway. This is true, for example, of cholesterol-lowering statins such as Lipitor, Zocor, and Pravachol, and antidepressant selective serotonin reuptake inhibitors such as Paxil, Prozac, and Zoloft (DiMasi and Paquette 2004). This will tend to be the case when government or university research reveals the same promising pathway to multiple firms more or less simultaneously.

D. The Industry-Specific Case Against Pay-for-Delay Settlements

The particular shape of congressional intervention in the balance between innovation and access, together with important industry-specific features of the pay-for-delay problem in pharmaceuticals, serve to undercut the Patent Act–based case for an exception to the ordinary operation of antitrust law. The argument applies in different ways to the innovator-focused and infringer-focused arguments for an exception.

With respect to innovators, the practice in question is a poor fit with Patent Act policy, because permitting pay-for-delay settlements is a highly innovation-inefficient means of increasing the incentive to innovate. To see this, consider as a benchmark a competitive practice that had the effect of increasing the length of the patent term at no incremental expense to the patentee. Arranging a longer term might be expected to increase producer profits and consumer allocative losses in equal measure (assuming, among other things, that the producer faces the same demand curve in each period). If
the social benefits of innovation increase proportionately with profits, then the ratio between innovation and deadweight loss is unchanged with respect to term length.

If instead, as is frequently presumed, additional profits have a **declining** impact upon the social benefits of incremental innovation, then a longer term entails a lower ratio—that is, less innovation “bang” for the additional deadweight loss “buck.” Such a practice is difficult to justify by reference to Patent Act policy, for the reason introduced in Part III.A. Congress’s selection of a particular patent term length implements a choice about the balance between innovation and acceptable deadweight loss. If Congress had chosen a longer term, it would have implemented a more innovation-protective policy with respect to patentees; but Congress did not do that. A “reasonable effectuation” of the Patent Act’s innovation protectiveness does not require permitting a practice that is less innovation-efficient than, but otherwise identical to, a major innovation-protective term of the Patent Act. Therefore, to the extent that a privately-arranged term lengthening is less innovation-efficient than the current period of exclusivity, it cannot be insulated from antitrust attack by reference to the policies of the Patent Act.

This approach differs from Kaplow 1984 in several respects. The ratio test is used not to directly calculate a ratio, and assign antitrust liability based upon the size of the ratio, but instead to compare the ratio to a ratio set by the Patent Act. Another difference is that here, we do not need to calculate the ratios. We can simply compare them along the dimension on which they differ. A decisive comparison is unavailable, by contrast, where the practice has a **higher** ratio than that implied by the patent term, or is not readily comparable to an element of patent policy.
Pay-for-delay settlements resemble an increase in effective term length, but in an important respect they are even less innovation-efficient. In exchange for receiving a reprieve from competition, the patentee must make a sizable payment. This payment reduces its profits and hence the incremental innovation incentive gained by arranging for the extension. The point is general: Gains from a practice that must be shared among, say, cartel members, dampen the dynamic benefit of increased profits. This deficit in innovation efficiency makes the agreements more difficult to justify as a reasonable effectuation of the Patent Act. In short, the Patent Act’s general policy of innovation protectiveness has, at best, a weak claim to insulating pay-for-delay settlements from antitrust attack.

Moving from the general case of patents to the specific case of pharmaceuticals further weakens the argument for insulation. As already noted, antitrust is in pari materia not only with patent law, but with industry-specific regulation as well. Compared to the Patent Act, the Hatch-Waxman Act provides within its domain a more specific and hence more relevant account of the congressionally implemented balance between innovation and competition.

The balance set by the Hatch-Waxman Act is a deliberate effort to promote consumer access through litigated challenges. For most drugs, the Hatch-Waxman Act is less innovation-protective than the Patent Act; as noted previously, the tax on blockbusters is a concession to consumer access at the expense of innovation. For a few drugs, it is actually more innovation-protective, thanks to the innovation subsidy provided by the industry-specific delays. In either case, the ordinary operation of the Act sets a particular balance between innovation and competition. The balance set for a
particular drug is disrupted by a settlement favoring somewhat more innovation at the further expense of consumer access.

The disruption to the congressional balance caused by settlement, moreover, is difficult to reconcile with the Hatch-Waxman scheme. With the Patent Act, a general norm in favor of innovation might at least be relied upon; by contrast, the Hatch-Waxman Act provides a calibrated outcome for different types of drugs. The Patent Act is silent about the role of litigation and the extent to which litigation can be avoided in the interest of preserving profits. In the Hatch-Waxman Act, by contrast, the promotion and delay of litigation are central preoccupations of the regulatory regime. An open-ended permission for innovators to set innovation policy by self-help is less plausible, as Congress has taken explicit steps to fill those gaps. Since litigation is the instrument by which the regulatory arrangement accomplishes its ends, it is difficult to argue that an end-run on the instrument is consistent with the scheme. Given that the regime explicitly provides for innovation protection in certain cases—an effective lengthening of the patent term for certain drugs, but a limited one—it is implausible to attribute to that regime a tolerance for an additional, highly innovation-inefficient means to accrue additional profits.

The infringer’s argument against antitrust liability is also weaker in the pharmaceutical context, compared to the general case. First, the generic firm lacks an innovator’s interest. The generic firms simply make use of the Hatch-Waxman scheme to offer a bioequivalent drug. Even if a Patent Act policy favoring innovation helps some infringers, it cannot be thought to apply here.

Limiting the generic firm’s ability to extract a benefit from unpromising
litigation has some effect on an infringer’s incentives, though not on its innovation incentives. To be clear, a limitation on settlement does not force the generic firm to see the litigation to completion—it can simply walk away from the suit. However, such a limitation on consumer-disregarding settlements does lower the value of the generic firm’s abandonment option (Grundfest and Huang 2006), an option that matters most when a party develops new information about its prospects during the course of litigation. The difference in reward implies that some marginal challenges will not be brought. There is little reason, however, to think that preserving the full value of this option is necessary to effectuate a Hatch-Waxman Act policy of promoting challenges, not least because the incentive to challenge is already so large.

As noted above, the generic firm can simply walk away from the suit. It is possible to imagine a more aggressive rule, however, in which the generic firm is prohibited from abandoning a challenge once initiated; compared to the assumption in the text, this would increase the fraction of challenges that result in early competition, but at the expense of some challenges not being brought.

Second, and again unlike many infringers outside the pharmaceutical context, the generic firm has deliberately stepped, not stumbled, into the infringement controversy. It does not move in uncertain terrain filled with hidden patent dangers; the patents protecting pharmaceutical innovations are open and notorious, compiled in the Orange Book, as described in Chapter 1. The generic firm volunteers for and seeks out the challenge by filing the Paragraph IV certification, which invites a lawsuit by the innovator (Hovenkamp et al. 2004). Here, and unusually, Congress has recruited and offered to compensate generic firms to bring patent challenges. Far from being
unwilling private attorneys general, generic firms have been deputized, in effect, to act on the public’s behalf. The explicit use of litigation to achieve the balance undercuts the preference for settlement sometimes discerned in ordinary patent policy.

In summary, the analysis in this Part reinforces the conclusion that pay-for-delay settlements are properly accorded a presumption of illegality as unreasonable restraints of trade. It also undermines, in a domain-specific way, the patent policy arguments sometimes thought to justify a patent-based exception to antitrust as a general matter. Finally, the analysis offers industry-specific support for the proposition that pharmaceutical consumers have an entitlement to the average level of competition implied by litigation, a proposition more difficult to sustain as a general matter.

Conclusion

Examining pay-for-delay settlements from the perspective of regulatory design yields two main results. First, the industry-specific bounty renders feasible an allocatively harmful settlement in a surprisingly wide array of circumstances. Because only the first-filing generic firm has potential access to the exclusivity period, an innovator has an especially strong incentive to pay to neutralize that source of potential competition. Because a guaranteed bounty is a valuable source of compensation to a first-filing generic firm, settlements that divide the remaining patent term confer a noncash payment for delay. Allowing an innovator to make multimillion dollar payments up to the amount of saved litigation expense exacerbates the allocative harm.

Second, the Hatch-Waxman Act produces a specific pattern of encouragement to and limitations upon innovative activity. That industry-specific pattern, rather than the
arguably innovation-protective policy of the Patent Act, provides the basis for an *in pari materia* analysis with antitrust law. The Hatch-Waxman Act’s calibration between innovation and competition is disrupted if firms are free to engage in self-help. The resulting disruption is difficult to square with the policies that animate the Hatch-Waxman Act, particularly in light of the inefficiency of pay-for-delay settlements as a means to provide additional reward to innovators.

Beyond the analysis of pay-for-delay settlements and other competitive practices in the pharmaceutical industry, a careful engagement with regulatory facts and economic theory within a specific industry is a promising method of antitrust analysis. The approach advanced here requires a close look at the economic effects of the regulation and the legislative instrument by which it achieves those effects. The project entails two distinct uses of regulatory economics: to understand the incentives of the parties as shaped by their legal constraints, and to assess what a law really means in light of the incentives it produces.

Such an economically aware and institutionally informed examination is particularly important in industries that are in the midst of deregulation. Such industries are an area of renewed interest in antitrust. Deregulation enlarges the domain of antitrust, as Hovenkamp 2005 has explained; it does so in part by altering the contours of liability. In some industries, the process of deregulation has occurred in an incomplete fashion, and partial deregulation may give rise to heightened antitrust concern.

Under partial deregulation, the regulatory regime balances innovation and competition by decentralized mechanisms, rather than by the central command of price
regulation. Under full regulation, there may be little role for antitrust, given its redundancy upon a regulator actively managing the antitrust function. Under partial deregulation, however, redundancy is less likely. The use of a decentralized mechanism by Congress risks nullification by unilateral or concerted action by self-interested firms, with allocatively harmful effects. Where the mechanism is not well preserved by the industry-specific regulatory agency, there may be a heightened role for antitrust intervention.

One virtue of an industry-focused approach is the presence of built-in limiting principles. An antitrust decisionmaker can resolve one set of cases without having to reconsider an entire category of conduct. For example, a court can resolve pay-for-delay settlements in the pharmaceutical industry—a set of cases of great theoretical significance and practical importance—without reconsidering the relationship of antitrust and patent generally. Another consequence, of course, is to put off answering broader questions—here, whether consumer-disregarding settlements of patent litigation in other industries are actionable as antitrust violations. But in an area of legal and economic inquiry so complex, and in which we lack even basic information about the facts on the ground in other industries, including the prevalence and structure of such settlements, this limitation is a virtue rather than a vice.

The difficulty of making sense of an enactment’s effects heightens the importance of deep industry expertise. The FTC’s role in pharmaceutical enforcement is illustrative. As far back as 2001, about a quarter of the FTC’s competition investigations were devoted to pharmaceuticals (Muris 2002). The FTC has produced detailed reports about industry competition (FTC 2002, 2009) and, more generally,
about the intersection of patent and antitrust (FTC 2003). It has brought enforcement actions challenging a variety of industry practices, aside from pay-for-delay settlements, including sham litigation, abusive Orange Book filings, and agreements among generic manufacturers (FTC 2006). We will return to the question of expertise in Chapter 3.

Moreover, the FTC sees the full range of cases due to its national enforcement scope and augments its stock of knowledge by combining the analyses of staff economists with information gleaned from civil investigatory demands of market players. In addition, the 2003 amendments to the statutory scheme (MMA 2003) require that industry settlements be filed with the FTC on an ongoing basis, which has provided continuing intelligence about industry practices.

Such expertise is particularly important in dealing with the panoply of strategies employed by pharmaceutical firms. Apart from the settlement cases, most such strategies entail beating competitors rather than joining them. Drug makers have displayed a great deal of ingenuity in preserving the profits from an innovative drug. For example:

*New-but-related drugs.* A separately patentable alteration to an existing drug is profitable provided that doctors and patients can be convinced to switch over as protection on the old drug ends (due to expiration or successful challenge). The most famous transition is from the anti-heartburn drug Prilosec to Nexium, an enantiomer of Prilosec’s active ingredient, omeprazole (Gladwell 2004).

*New patents on the same drug.* For example, a firm may assert patents on metabolites (the compound a drug is converted to within the body), intermediates that appear during the production process, or alternative crystalline forms. An important
aspect of this strategy has involved an interaction with the regulatory system. As noted in Chapter 1, an ANDA-IV must address every patent that is listed by the drug manufacturer in the Orange Book. Under an early FDA interpretation, if a brand-name firm added additional patents after the commencement of an ANDA-IV challenge, the generic drug maker had to amend its certification, which triggered further infringement challenges, which, in turn, was understood to trigger additional and later 30-month stays. Eventually, this loophole was closed. The filing of multiple stays by Bristol-Myers with respect to BuSpar was part of the basis for a consent decree with the FTC. For another drug, Paxil, indirect and direct purchaser class action suits resulted in settlements of $65 million (Nichols 2005) and $100 million (Stop & Shop 2005), respectively.

Some of the strategies are very difficult to justify by reference to a plausible consumer benefit. That is not to say that such techniques are all illegal or even troubling—new drugs and price-lowering distribution strategies, for example, potentially provide considerable consumer benefit. But the proliferation of such strategies does give rise to a bewildering array of choices for antitrust enforcers—what then-Commissioner Jon Leibowitz likened to a game of “Whack-a-Mole” (Leibowitz 2005).

An important test of that expertise comes in the ongoing debate over “authorized generics” (Reiffen and Ward 2005; Kong and Seldon 2004; FTC 2009). The basic idea is that an innovator, faced with competition from a first-filing generic firm, recruits an additional generic firm to sell an unbranded version of the drug under the innovator’s own license. The presence of an additional generic competitor, selling during and after
the bounty period, lowers prices in the generic segment of the market. This effect on the
generic segment of the market is typically a fifty percent discount on the innovator’s
price, compared to the thirty percent discount with just one generic firm (Morgan
Stanley 2004). Consumers benefit in the short run from lower prices, and the innovator
enjoys incremental profits from the additional revenue stream; only the independent
generic firm loses out.

Over the last several years, an authorized generic product has become a familiar
accompaniment to a pre-expiration launch by a generic firm (Abboud 2004). But
fighting-brand pharmaceuticals were used in a limited way before that. In the 1990s,
innovator firms engaged in a certain amount of own-brand generic sales. Then, too, the
activity raised antitrust concern (Kamien and Zang 1994; Yang 1994). In the late 1990s
the innovators for the most part exited the generics business, as they discovered that
selling generic drugs was not their forte, and as they improved in their ability to shift
customers from one product to its successor (Freudenheim 1997). The resurgence of
authorized generics may be attributable to three features: the patent expiration of a large
number of blockbuster drugs, which creates an unusually large opportunity for generic
competition; an increase in the number of exclusivity periods granted, particularly as
evergreening strategies involving later-added, weak patents are successfully challenged
by generic firms; and the increased penetration of generic entry, which creates a sizable
profit opportunity for the innovator, provided that the additional entry does not affect
pricing and volume too much in the branded segment of the market.

Generic drug makers complain that the use of authorized generics, in reducing
the benefits of the 180-day exclusivity period, is contrary to the purpose of, and hence
violates, the Hatch-Waxman Act. This argument has failed on a textual reading of the Act, which merely excludes subsequent ANDA filers (Teva 2005). Generic firms have also argued that the use of authorized generics violates antitrust law by reducing generic profits to such an extent that a challenge is not worth pursuing, thus deterring generic entry.

The underlying antitrust concern is that the practice, though beneficial in its short-run allocative effect, will discourage future entry, ultimately leading to higher prices. Acting to deter a rival’s procompetitive actions is a general strategy analogous to, for example, the price-matching policies of large retail stores (Edlin 1997). The structure of at least some authorized generic licenses (for example, Asahi Glass 2003) provides for withdrawal should independent generic entry cease. The authorized generic mechanism also has a unique feature that potentially enhances its deterrence. If the innovator licenses an outside firm, its contract is an observable commitment to entry, which may provide a source of credibility. Such an ability to precommit might make seeing through the threat unnecessary in practice—though the direct profitability of the additional distribution mechanism may, aside from lessening the antitrust concern, make precommitment unnecessary.

Unless authorized generics actually deter entry in practice, or—an important complication—slow the filing of ANDA-IVs or lessen the vigor of their pursuit, there is no basis for antitrust concern. Anecdotal evidence suggests that authorized generics might have little negative effect on generic entry, at least for the most important drugs. For example, Apotex earned a large profit in its challenge to Paxil despite competition from an authorized generic. According to Apotex’s own figures, its profits were reduced
from the $530-to-$575 million range to the $150-to-$200 million range because of the authorized generic entry (Apotex 2004). Substantial empirical work remains to be done to resolve the issue decisively. The FTC is uniquely positioned, due to its expertise and power, to collect and assess the relevant information, and it has indeed begun to do so, issuing an interim report (FTC 2009) that assesses the price effect of an authorized generic, though without assessing what the effect on challenges has been.

The underlying impulse to tailor innovation policy by industry resembles the parallel project by patent scholars to understand patent law in an industry-specific fashion (Burk and Lemley 2003). In both contexts, the perspective implies that a holding reached within a particular industry’s factual setting is unlikely to have ready appliability to other industries. One important difference between the projects, however, is that the industry-specific approach in patent law operates primarily through judicial interpretation; it must necessarily do so, given the single statutory scheme that governs patent doctrine across most industries.

The approach here, by contrast, places more emphasis upon Congress and expert agencies. Congressional enactments govern the balance between innovation and competition, modulating the vigor of antitrust enforcement in an industry-specific fashion. The effect is to place the overall thrust of innovation policy more firmly in the hands of the legislative branch, perhaps quieting congressional complaints of judicial circumvention in other areas of competition policy. The competition regulator, meanwhile, plays an important role in decoding the meaning of a legislative enactment as it bears upon industry economics and antitrust law. That role is particularly important where, as in pharmaceuticals and other industries, courts need help in recognizing and
tailoring antitrust analysis to the “distinctive economic and legal setting” (Trinko 2004) of a regulated industry.
Chapter 3: A Public Survey of Pay-for-Delay Agreements

Introduction

Antitrust policymaking in the United States has a tension at its core. Antitrust law “maintain[s] certain basic rules of competition” as a way to preserve low prices, efficient production, and robust innovation (Whinston 2006). In regulating a particular type of behavior, a decisionmaker may choose a rule that minimizes costly errors—false condemnations and false exonerations—even at the expense of accuracy in a particular case. Courts, as the actors charged with setting substantive antitrust policy, routinely make such choices. Unfortunately, courts lack the information needed to select optimal rules.

Consider, for example, predatory pricing. Antitrust law permits price-cutting to exclude a rival, provided that the price does not fall below cost, on the view that a more aggressive rule yields too many false condemnations (Brooke Group 1993). That lenient rule increases false exonerations, but the Supreme Court has concluded that these are unlikely, as predation is “rarely tried, and even more rarely successful.” But how does a court come to know this? And is a court the right institution to uncover the answer?

This chapter identifies and examines an “aggregation deficit” in antitrust analysis: the troubling lack of information about the frequency and costliness of anticompetitive activity. Aggregation matters for both the substance and institutional structure of antitrust policy. In setting substantive antitrust rules, courts make rough guesses, informed by economic theory and the facts of a specific case, about the distribution of real world economic conduct. What a decisionmaker actually needs is aggregate information on which to base a cost-minimizing substantive antitrust rule. In
selecting an antitrust decisionmaker, moreover, we ought to favor the institution that has superior access to aggregate information, all else being equal.

As a vehicle for considering the substantive and institutional dimensions of an aggregate approach, this chapter, like Chapter 2, focuses on patent settlements between a brand-name drug maker and its generic rival. As explained there, settlements result from a generic drug maker’s effort to market a competing version of a brand-name product. The brand-name firm responds with a patent infringement suit that claims its product is protected by one or more patents, and the generic firm counters that the patent is invalid or not infringed by the proposed generic product. The brand-name firm, rather than take a chance that the generic firm might win that argument in court, thereby ending its monopoly on the product, settles the litigation by paying the generic firm to abandon the challenge and delay entry. Does this agreement violate antitrust law? As I argue in Chapter 2, this question is the most important unresolved issue in U.S. antitrust policy, measured by economic importance and repeated high-level judicial attention.

Identifying the proper scope of liability, however, is not a simple task. Some settlements do not raise pay-for-delay concerns. For other settlements, it is difficult to tell whether a payment was made. Before an optimal antitrust rule can be developed, policymakers need accurate information regarding the scope and nature of the problem. As an initial step toward erasing this deficit, this chapter assesses the problem of entry-delaying settlements by aggregating publicly available data about these settlements and considering the overall picture that emerges. This approach draws upon a new dataset of drug patent settlements, developed from a wide range of public sources. The resulting dataset provides, for the first time, a vivid picture of the frequency and distribution of
settlement activity. Viewing the settlements collectively permits new insights about enforcement priorities, the optimal substantive rule, and the choice of decisionmaker.

The analysis reveals an evolution in the terms of settlement. Whereas early settlements simply traded cash for delay, modern settlements show sophistication in the means by which payment and delay are provided. One example is the use of side deals, consummated at the same time as settlement of the patent litigation, in which the generic firm contributes unrelated value, such as a separate patent license, ostensibly in exchange for payment. That tactic undermines reliable case-by-case characterization of settlements as collusive or not: In a particular instance, it is difficult to tell whether the brand-name firm’s payment is consideration for delay, for the unrelated value, or both.

An aggregate approach permits us to address the question in a different way. It reveals that these sorts of deals are a frequent component of settlements, but rare outside of settlement. Thus, the overall pattern suggests that they provide a disguised means to confer payment. This supports the adoption of a presumption that a brand-name firm’s payment to a generic firm, when contemporaneous with a generic firm’s agreement to delay entry, is consideration for delay, not for the goods or services acquired in the side deal.

As an institutional matter, the aggregate approach undermines the case for courts as primary antitrust policymakers. A court is largely limited to the facts of a particular case. It lacks the capacity to collect information about the distribution of activity in the economy. To be sure, parties can supply the court with aggregate analyses based upon public information, but public disclosures contain important gaps. Moreover, courts are likely to have trouble processing this information. Agencies have a decisive advantage
in collecting and synthesizing aggregate information given their expertise, access to confidential information about regulated firms, and freedom to examine issues over a long period of time, outside the litigation context. Thus, the analysis suggests that the FTC should do more to exploit its informational advantage as a plaintiff, amicus, and rulemaker.

Finally, the aggregate perspective provides a basis for predicting the success or failure of antitrust enforcement over time. As applied to settlements, the prediction is pessimistic. Settlement has continued to evolve—even beyond side deals—in response to the enforcement emphases of particular litigants and courts. Settling parties have been able to achieve the same entry-delaying effect of the earliest settlements, while devising new disguises for payment or even the very existence of agreement. As litigants respond dynamically to judicial scrutiny with new and complex settlement structures, existing antitrust institutions have trouble keeping up.

This chapter proceeds in six parts. Part I assesses the state of the literature, identifying a gap in the empirical understanding of settlement. Part II describes the data collection effort. Part III introduces a typology of settlement and provides an initial assessment of the size of the pay-for-delay problem. Part IV describes the changing structure of entry-delaying settlements and spells out how these features recommend making the settlement issue an enforcement priority. Part V identifies an error cost-minimizing rule for side deals, explaining why they should be presumed to convey payment when accompanied by an agreement to delay entry. Finally, Part VI addresses the question of institutional choice. It first shows why courts make poor aggregators, and proceeds to consider how agencies can help fill the gap by aggregating data and
promulgating rules.

I. The Aggregation Gap

This Part assesses the current understanding of the pay-for-delay settlement problem. Although settlements have received a great deal of attention, almost all of it has focused upon the theoretical issues raised in individual cases, at the expense of important factual questions that also arise. I describe this neglect and its connection to the larger problem of an aggregation deficit in antitrust.

Before doing so, it is useful to review a few points from Chapter 2. As explained there, pay-for-delay settlements restrict a particular kind of competition between brand-name and generic firms. The two drug makers have a powerful incentive to settle. For a blockbuster drug with billions of dollars in annual sales, a brand-name firm has billions to lose from generic competition. Moreover, entry hurts the brand-name firm more than it helps the generic firm. Entry lowers total producer profits by introducing price competition, particularly once other generic firms are free to enter after the 180-day period ends. There is therefore a large gain from trade for the two firms. A settlement in which the brand-name firm pays the generic firm, and the generic firm agrees to delay entry, is profitable for both firms. Because later filers generally have much less incentive to challenge a brand-name drug patent, including no eligibility for the 180-day period, buying off the first filer is an effective means to remove the most potent entry threat.

Such settlements, if they include payment, reduce expected static consumer welfare. Early competition benefits consumers by lowering drug prices sooner. The
consumer benefit is probabilistic, since it is not certain that entry would occur; the brand-name firm might win the suit. Settlements without payment reflect the perceived strength of the patent. For example, suppose a generic firm has a fifty percent chance of success, and ignore the complication of the 180-day period emphasized in Chapter 2. Then the parties should settle upon an entry date halfway between immediate entry and patent expiration. That result is equal to the average result of litigation, in which the consumer has a fifty percent chance of enjoying the full benefit of immediate competition and a fifty percent chance of receiving no benefit.

By contrast, bargains that reflect not only perceived patent strength but also payments from brand-name to generic manufacturers will induce the generic firm to accept a later entry date, which decreases consumer welfare. Thus, a pay-for-delay settlement transfers wealth from consumers to drug makers, in the form of continued high pharmaceutical prices, with brand-name firms sharing a portion of that transfer with the generic firm. The higher price also alters the purchase decisions of consumers and insurance providers, introducing an additional welfare loss.

As I explain in Chapter 2, the consumer-disregarding effect of pay-for-delay settlements requires their condemnation as a violation of antitrust law. Allocating markets in this fashion is a restraint on trade in violation of section 1 of the Sherman Act, and may also be condemned as illegal monopolization. It is therefore no surprise that the FTC and private parties have brought numerous cases, often together with state attorneys general, arguing that certain pay-for-delay settlements violate antitrust law.

Chapter 2 considers a range of arguments offered by settling parties. For example, permitting settlement increases the brand-name firm’s profit, and hence its
expected reward for developing innovative drugs, the marketing of which provides great benefits to consumers. Put another way, the static harm of settlement from high prices today must be weighed against the dynamic benefit of more and better drugs in the future. As Chapter 2 explains, this is a doubtful argument, particularly where, as here, the increase in innovative incentive from delaying competition is partially offset by the necessary payments to the generic firm.

Similarly, some parties have noted that settlements in other industries are similarly consumer-disregarding, raising the specter of a widespread expansion of liability if these settlements are prohibited. It is true that market division through patent settlement is a real possibility in other industries, and to that extent, antitrust liability may be warranted there too. In addition, as explained in Chapter 2, the Hatch-Waxman Act reflects a specific effort to promote consumer access through litigated challenges, a feature that makes the case for prohibition particularly strong in this industry.

Thus far, courts have taken a relatively sympathetic, albeit highly uneven, stance toward pay-for-delay settlements. The courts denying liability have accepted, as a doctrinal matter, a maximalist view of the patent right. They have adopted the view that any settlement is permissible, provided it restricts no more entry than the nominal scope of the patent if valid and infringed. Courts permitting settlement add the caveat that the brand-name firm must not have engaged in fraud upon the patent office or sham litigation. As a result, brand-name firms are effectively permitted to buy private term extensions to their patents. The maximalist view thus produces the absurd result that an ironclad patent and a trivial patent have the same exclusionary force. Each can support a settlement that restricts generic entry until the nominal expiration date of the patent.
The maximalist perspective has a second pernicious effect. It ignores the fact that the nominal scope of the patents at issue, particularly the expiration date of the last-expiring patent, is highly malleable. A sophisticated brand-name drug maker can produce a steady stream of patents, with successively later expiration dates, which in turn support a settlement date that is even later than the expiration of effective protection.

A settlement involving the blockbuster drug Lipitor, Pfizer’s most important product, provides an example. Pfizer sued Ranbaxy, the first-filing generic firm, over Pfizer’s two strongest patents, expiring in March 2010 and June 2011, winning as to the first patent but losing as to the second (Pfizer 2006). Analysts therefore expected entry in March 2010 (Risinger 2008), or at the very latest in June 2011 when the second patent expired. However, when the parties eventually settled, generic entry was set for November 2011, later than the expiration of either patent (Pfizer 2008a). The parties explained this result by noting that, shortly before settlement, Pfizer had also sued Ranbaxy on two minor patents that expire in 2016. The two patents at issue were not listed in the Orange Book (Drug Industry Daily 2008), which contains listings of those patents, filed by the brand-name firm, that “count” for Hatch-Waxman purposes. The main effect of the inclusion of these patents was to permit the parties to choose an entry date later than the expiration of the two main patents at issue.

Beyond this theoretical question—do pay-for-delay settlements violate antitrust law?—there is a set of factual questions that must be answered. For example, how frequently do pay-for-delay settlements occur? Knowing the answer is necessary to decide whether to make the settlement issue an enforcement priority. A second factual
question arises in many modern settlements. If settlement and delay occur as part of a larger set of transactions between the two firms, how do we know that the payment was made in exchange for delay, rather than for some other valuable consideration? Often, this is a difficult question. In the only case involving a side deal that has been fully litigated so far, attempts to determine whether the particular settlement was anticompetitive produced divergent results at each level of review. These factual questions have been neglected by scholars so far.

This gap in our understanding of modern settlement practice exemplifies a general problem in antitrust enforcement. Given a theoretical model of anticompetitive behavior, true under specific factual circumstances, how do we establish with confidence that those circumstances are present in a particular case? If that determination is imperfect, how do we identify a cost-minimizing rule—for instance, that alleged predation is reviewed leniently because predation is “rarely tried, and even more rarely successful,” as discussed above, or that resale price maintenance ought to be accorded rule-of-reason treatment because its procompetitive uses are not merely “infrequent or hypothetical” (Leegin 2007)?

Because a court lacks the capacity to independently collect the information necessary to develop an optimal rule, it relies upon others, including academics and other governmental institutions. In considering predation, for example, the Supreme Court has explicitly relied upon a “consensus among commentators” that the practice is rarely tried or successful (Matsushita 1986). If the external consensus changes, the Court suggests, so too may the substantive rule. Agencies and Congress play a similar role. For example, Justice Breyer, dissenting from the Court’s recent decision to end a
longstanding per se ban on resale price maintenance, thought any change should await solid information about “how often are harms or benefits [from the practice] likely to occur” (Leegin 2007). He also questioned how readily the two can be distinguished; in other words, “[h]ow easy is it to separate the beneficial sheep from the antitrust goats?” Such information must be supplied by others, if it is to be collected at all, since courts, unlike Congress and the FTC, are not “well-equipped to gather empirical evidence outside the context of a single case.”

Real world evidence about the frequency and distribution of anticompetitive activity helps to build the requisite consensus among commentators. Such work has furthered our understanding of predation (Bolton et al. 2000), vertical contracting (Cooper et al. 2005), and other competitive practices. Industry-specific analyses have been important too in industries as diverse as motion picture exhibition (Davis 2005) and telecommunications (Shelanski 2000); for a review, see Bresnahan 1989. In addition to measuring the aggregate costs of a type of antitrust violation, this study adds a distinctive dimension: the effort to understand the evolution of a practice over time. Understanding this evolution provides evidence about how well existing antitrust instruments can be expected to cope. Frequent or rapid mutations in the practices of regulated firms raise doubts about whether common law processes can effectively regulate those practices.

Whether by legislative reform or judicial decisions, “economic policy must be contrived with a view to the typical rather than the exceptional,” in the apt phrase of Stigler 1952. Both legislators and judges would benefit from a clear idea of how often and in what form settlements occur, and how effective we can expect judicial
management to be. This is a fitting moment to examine real world evidence of settlements, before the Supreme Court or Congress establishes a new rule. The Supreme Court has not weighed in on the settlement question, but if and when it does, its rule will be difficult to undo, thanks to the infrequency of antitrust review, the operation of stare decisis, and a fear of upsetting reliance interests. The next Part begins the examination necessary to formulate an optimal rule for pay-for-delay settlements.

An agency such as the FTC is well-positioned to fill these informational gaps. The agency has a statutory mandate to collect, study, and publish information about particular industries. It has general authority under the FTC Act to require firms to divulge confidential information relevant to antitrust policymaking. In the particular context of settlement, the FTC’s position is even stronger: As mentioned in Chapter 2, it has unique access to the details of every brand-generic settlement since December 2003, due to drug makers’ special statutory obligation to file all such settlements with the agency. This aggregate information complements other sources of FTC expertise developed and used in litigation, congressional testimony, and public hearings.

The FTC sometimes uses this advantage to good effect. In 2002, the agency published an important survey of brand-generic drug competition, drawing upon information supplied by drug makers under FTC compulsion as well as information collected independently by the FDA (FTC 2002). That study indicated the importance of the pay-for-delay settlement problem and made a variety of policy recommendations. However, there has been no follow-up to the 2002 study. More generally, as Scherer 1990 notes, industry studies—once a staple product of the FTC—have become less frequent.
The FTC’s conclusions, based on its aggregate information, can be deployed in a variety of policymaking settings. In the case of FTC 2002, the conclusions were used in amicus briefs, legislative advocacy, and litigation brought by the agency. But in each of these settings, the agency is essentially supplying its information to an external decisionmaker. I say “essentially,” because in the case of adjudication, the FTC makes an initial determination, which is then reviewed by an appeals court. The agency has available to it a more aggressive option, however, which emphasizes the FTC’s role as a decisionmaker in its own right: antitrust rulemaking. In Part IV, I argue that the FTC’s aggregation advantage is a reason to favor antitrust rulemaking, and that pay-for-delay settlement is an attractive candidate for a rule. But first, I will lay out what an aggregate approach can tell us about drug patent settlements.

II. Data Collection

To examine the frequency and evolution of brand-generic settlements since 1984, I collected a novel dataset. The object was to identify and synthesize all public information about the frequency and terms of settlement. The effort drew upon press releases, trade publications, financial analyst reports, transcripts of analyst calls with management, court filings of patent and antitrust litigation, SEC filings, FDA dockets, and FTC reports. The broadest search was a review of all articles in the Factiva database mentioning “settlement” and a “new drug application.” The database includes newspapers, magazines, trade journals, press releases, company presentations at analyst conferences, and transcripts of calls between company executives and equity analysts. The search included linguistic variants of “settlement” and the abbreviations “NDA”
and "ANDA." The Factiva search found a number of settlements that were not evident in other sources, such as analyst reports. In many cases, articles in Factiva filled in important settlement details.

For settlements involving twelve drugs, the actual settlement agreement was available. These, and their sources, are listed in Table 2. In addition to the terms of settlement, I recorded the annual sales figures at the time of settlement and noted whether the generic firm was eligible for the exclusivity period. To determine eligibility, I assessed whether the drug was subject to the exclusivity period, whether the settling generic firm was a first filer, whether any exclusivity eligibility had already been triggered at the time of settlement, and whether the settlement itself included a forfeiture of retained exclusivity. The second determination is the most difficult, both because the FDA considers the identity of the first filer to be confidential information, and because there are often multiple first filers. I based the determination on FDA letters granting ANDA approval with exclusivity (which are not confidential), generic-firm press releases reporting presumed first-filer status, and a comparison of complaints in patent suits with FDA reports of the date of a first ANDA filing, which is not confidential on a prospective basis starting in 2003.

I also determined whether a particular provision of the MMA 2003, discussed below, applied to the settlement. To be included in the set, the agreement must pertain to patent litigation resulting from an ANDA filing by a generic firm. This criteria rules out, for example, an agreement over the contraceptive Ovcon 35, which was not a patent dispute but did feature an agreement between the brand-name and generic firm that was challenged as anticompetitive by the FTC. It also omits a few drugs where a
settlement as to another drug discouraged the filing of an ANDA in the first place. The search period extended from 1984, when the Hatch-Waxman Act was passed, through August 2008, and therefore ignores subsequent settlement activity.

This work yielded information for 143 settlements involving 101 brand-name drugs. For 28 drugs, the brand-name drug maker settled with multiple generic firms. Multiple settlements can be the result of settlements with multiple first filers sharing the exclusivity entitlement (as mentioned in Chapter 2), or settlements with later filers who lack eligibility for the exclusivity period. Although the focus of the subsequent analysis is settlements with first filers, in some cases settlements with later filers can raise pay-for-delay issues as well.

Several checks confirm that the dataset contains nearly all significant settlements that delay entry. For example, FTC 2002 catalogued 14 troubling settlements in 2002, but did not name names: 8 cash or side deal settlements, 2 “supply agreements,” and 4 retained exclusivity settlements. Of these, I can match 7, 2, and zero settlements, respectively, to my dataset. Of the 11 settlements in the FTC’s update for fiscal 2005, I can account for 8, as well as 26 of 28 in the 2006 update and 20 of 33 in the 2007 update (FTC 2006, 2007, 2008). Barr has stated that it reached settlements as to 14 drugs (Downey 2007). My data likewise contain 14 settlements as of the time of Barr’s statement. Similarly, the data contain 10 Teva settlements by early 2007, which is identical to Teva’s own statement (Whitehouse 2007).

The dataset oversamples settlements that restrict entry for important drugs. Important drugs receive more extensive coverage in public disclosures, and settlements that restrict entry tend to receive more attention. Thus, omitted settlements are likely to
be for minor drugs, or settlements that had no effect on entry. For those settlements in
the dataset, publicly available information contains significant gaps. In particular, price
terms are normally omitted, and detailed settlement terms are sometimes missing. Even
with these limitations, the new dataset is a useful tool for examining the extent and
evolution of settlement; indeed, it may be the most comprehensive examination of
brand-generic settlements until and unless the FTC uses its power of compulsion to
produce a complete dataset.

III. A Typology of Settlement

Of the 143 settlements in the dataset, 60 settlements include both delayed
generic entry and possible contemporaneous provision of value by the brand-name firm.
The 60 settlements involve 51 out of the 101 drugs in the dataset. For an additional two
drugs, Prefest and Mircette, the Hatch-Waxman dispute was resolved through
acquisition: The generic firm bought out the brand-name firm’s rights to the drug, thus
ending the possibility of competition between the two (Krauskopf and McKay 2004;
Barr 2005). Neither deal was challenged by the FTC, however, suggesting that the firms
lacked market power in the first place.

As to the remaining 48 drugs, my data collection effort identified no pay-for-
delay issue, and so far as I know, the settlements raise none. Of the 81 settlements in
this category, 67 pertain to 48 drugs whose settlements appear to raise no pay-for-delay
issue. The remaining 14 settlements pertain to drugs in which the brand-name firm
reached at least one other settlement that does raise a pay-for-delay issue. To avoid
double-counting, these latter settlements are not included in the number of drugs in this
category.

Some such settlements include an agreement on the generic entry date, without any payment. These negotiated outcomes likely reflect the perceived strength of the relevant patents. Their existence demonstrates that settlement without payment is feasible (Leibowitz 2008). (Even if no-delay settlements were infeasible, however, the main reasons to condemn pay-for-delay settlements would still hold.) Table 3 summarizes the three categories of agreement.

For the 51 drugs raising pay-for-delay issues, payment and delay take a variety of forms. For 21 of the 51, the compensation was wholly or partly monetary. This category includes “underpayment” settlements discussed below. Sometimes the payment was an open conferral of cash. For other drugs, the possible payment was embedded within a more complicated transaction. The caveat “possible” is used because in some cases public information leaves it unclear whether the settlement included compensation, an issue explored in more detail in Part III. These 21 drugs are listed in Table 4, together with details about the various forms of payment, which are explained later in this chapter. On average, they had annual U.S. sales, measured in the year of settlement and adjusted for inflation, of $1.3 billion.

The 21 drugs include blockbusters such as Lipitor (more than $7 billion in annual sales) and Nexium (more than $3 billion). Five drugs with annual sales exceeding $2 billion account for more than two-thirds of the total, measured by annual sales. More than half are new versions of existing therapeutic agents, whose patents are generally thought to be weaker because they tend to be obvious (and hence invalid) and are easily worked around. This group consists of Sinemet CR, K-Dur, Naprelan,
Niaspan, Effexor XR, Propecia, Adderall XR, AndroGel, Wellbutrin XL, Nexium, and Aggrenox. Even for those drugs that are not new versions, some of the relevant patents are noticeably weak. For example, Altace was protected by a patent not on the basic compound, but an enantiomer, and was subsequently invalidated (Aventis 2007). Provigil is protected not by a compound patent, which expired, but by a particle-size patent (Cephalon 2010). Some of the settlements in Table 4 have lapsed, and generic entry has occurred, while others continued to block entry as of March 2009. Ten drugs in the latter category account for about $17 billion in annual sales.

The effect of delayed entry can be enormous. For the questionable settlements in Table 4, a one-year delay in generic entry represents, under conservative assumptions, a transfer from consumers to producers of about $14 billion. Suppose generic entry achieves 75% penetration and that the generic product is priced at a two-thirds discount, relative to the brand-name drug. These figures are a simplification, because in reality, penetration and the discount (particularly during the 180-day period) are smaller at first, but quickly increase. Under these assumptions, the avoided transfer is one-half of annual sales, or $661 million per drug. Across 21 drugs, the total is about $14 billion. This figure does not include welfare losses caused by pricing distortions. One of the settlements, Plavix, never took full effect, as explained in detail in the Appendix. With Plavix removed, the transfer from a one-year delay is $12 billion.

Whether the one-year benchmark is an overestimate or an underestimate is often difficult to assess in a particular case using public information. Part of the total delay caused by settlement is attributable to the strength of the patent itself, rather than payment. Since the pre-expiration period covered by settlement is several years—the
average period, weighted by sales (and excluding Plavix), is 4.1 years—the benchmark is likely conservative.

A more nuanced figure might be developed by offering a specific prediction about what would have happened in each case absent the settlement. The particular circumstances of a settlement can provide important indications of the likely alternative outcome. A weak patent, and likely early entry, might be identified by an analysis of the patent’s validity and scope, or inferentially by a large payment. Another basis for inference is preparations by a generic firm to launch “at risk,” before a court has ruled on invalidity or noninfringement. Launches at risk suggest that the patent protection is weak, because the generic firm does not fear the prospect of damages (which would exceed the generic firm’s profits if imposed) or a preliminary injunction (which would spoil the expensive preparations for a generic launch).

For some drugs, public statements by management or the expectations of financial analysts help to provide a specific measure of delay. In the case of Provigil, for example, the drug maker’s CEO said that due to settlements, “We were able to get six more years of patent protection. That’s $4 billion in sales that no one expected” (George 2006). The CEO’s statement reflects the firm’s pre-settlement expectation of entry in 2006 (Cephalon 2005), and settlements delaying entry until 2012 (Cephalon 2010). In the case of Lipitor, the settlement delayed anticipated entry by nearly two years (Risinger 2008). Overall, the $12 billion benchmark estimate is likely to be conservative.

For settlements involving 25 drugs, the brand-name firm compensated the generic firm as part of an entry-delaying agreement, but the compensation was not
monetary. Instead, compensation took the form of retained exclusivity. As explained in Chapter 1, the 180-day period is valuable to the generic firm. One hundred eighty days of duopoly is worth hundreds of millions of dollars in the case of a blockbuster.

Perhaps surprisingly, the entitlement can also be sold to another generic firm. A generic firm can either selectively waive its entitlement in favor of a particular later filer, or relinquish it entirely. This is a profitable strategy where the firm with the entitlement has been unable to secure FDA approval—for example, due to difficulties in formulating or manufacturing the product—and a later filer is ready to go to market, but for the fact that it is “bottled up” behind the first filer in the sense discussed in Chapter 2. Selective waiver has been permitted for numerous drugs, including Zantac, Zoloft, and Wellbutrin XL. The FDA has insisted that selective waiver, as opposed to relinquishment, can occur only once the exclusivity has been triggered through a favorable court ruling or commercial marketing (FDA 2004).

The value of this opportunity, however, is discounted by the uncertainty that the generic firm might lose the litigation, and thus never enjoy the exclusivity period. Other risks include the possibility that a later-filing generic firm wins a patent suit, thereby triggering the first filer’s exclusivity period before the generic firm secures FDA approval, or that the patent expires before the generic firm wins the suit. A brand-name firm’s agreement to drop the patent fight is valuable to the generic firm because it raises the probability of enjoying the exclusivity. Moreover, as discussed below, settlement does not remove entitlement to the exclusivity period.

These 25 drugs are listed in Table 5. The list omits settlements where there was no delayed entry or where the available data was ambiguous about continued
entitlement to the exclusivity period. Entry as to one drug on the list, Exelon, was not disclosed, but was assumed to be at least 180 days prior to patent expiration. This list is likely underinclusive.

The ability to settle with retained exclusivity disrupts the alignment of interests between the generic firm and consumers. Ordinarily, late entry dates are bad for consumers, but also bad for the alleged infringer, whose profits are a function of the amount of time on the market, and who therefore can be expected to fight for an earlier entry date. Here, by contrast, the generic firm cares more about protecting its 180-day duopoly entitlement, and less about when exactly that entry occurs. It is therefore willing to trade a later entry date for a greater chance of enjoying the 180 days, as explained in Chapter 2. Meanwhile, consumers and taxpayers finance the continued sale of drugs at the higher, brand-name price.

This argument has an important limit. If the generic firm’s pre-expiration entry lasts for less than 180 days, then its profits are, roughly speaking, linearly increasing as it pushes for an earlier entry date. In that case, the alignment between the generic firm and consumers is more nearly maintained. Of the 25 drugs listed in Table 5, 7 have entry dates so late that they have less than 180 days of exclusive sales. Frequently, this occurs when a brand-name firm secures a six-month pediatric extension under § 355a that is tacked onto the end of the patent term. For the remaining 18 drugs, the misalignment critique applies.

The 25 drugs have average annual sales of $580 million. Of these, the 18 drugs with “full” exclusivity have average sales of $442 million. If guaranteed exclusivity induces a delay of one year for each of these drugs, the transfer, using the same calculus
described above, would be about $4 billion.

The preservation of exclusivity can take a second form. In some cases, a generic firm wins a patent challenge, but is blocked from approval by a second patent that the generic firm either did not challenge at all, or challenged unsuccessfully. In such a case, the generic firm “wastes” the exclusivity resulting from that partial victory, which is triggered and expires while the generic firm is blocked from entering by the second patent (Beers 2008). Once the second patent expires, the generic firm enters, but without exclusivity.

A generic firm can avoid wasting its exclusivity by abandoning its challenge, and agreeing to enter with exclusivity upon the expiration of the second patent. This benefits the brand-name firm, and harms consumers, for the same reason: Prices are higher during the (preserved) duopoly exclusivity period than with full competition from other generic firms.

The Zoloft settlement between Pfizer and Zenith is an apt example. Pfizer had two patents on Zoloft: a strong patent expiring in 2006, and a weak patent expiring in 2013. In 1999, Zenith challenged the 2013 patent but not the 2006 patent. Winning as to the 2013 patent would have wasted the exclusivity, unless that happened after the expiration of the 2006 patent. (Patent suits are slow, but not that slow.) Instead, Zenith agreed to enter with exclusivity upon the expiration of the basic patent.

The Lipitor settlement appears to contain another variant. When Ranbaxy won its challenge to one patent in the Federal Circuit, this triggered exclusivity, but prematurely, since the other valid and infringed patent prevented FDA approval (Pfizer 2006a). The combined result would have been to permit entry without exclusivity in
March 2010. (The patents expiring in 2016 were never listed in the Orange Book, and did not affect that result.) However, Pfizer had three more Orange Book-listed patents in reserve, on which Ranbaxy was likely the first ANDA filer, but Pfizer did not sue. Under pre-MMA law, each patent provided a fresh opportunity for exclusivity (Apotex 2006). By declining to sue Ranbaxy on these patents, Pfizer preserved Ranbaxy’s exclusivity despite the initial trigger, a preferable result for both parties. The MMA 2003 replaced this “patent-by-patent” approach to exclusivity with § 355(j)(5)(B)(iv)(II)(bb), which provides a single opportunity for exclusivity for each product (Thomas 2005).

In addition to the drugs for which the only form of compensation is retained exclusivity, most of the drugs in Table 2, after the first five, have secured an assured 180 days of generic sales. The first five settlements included no pre-expiration entry, for reasons discussed below. Other settlements explicitly trigger exclusivity, or involve generic firms that are ineligible for exclusivity in the first place. This is the case when the generic firm is not a first filer, or when the brand-name drug does not give rise to exclusivity eligibility.

Five pay-for-delay settlements fit neither of these categories. Three are “interim” agreements, which restrict entry while the patent infringement suit is pending but do not resolve the suit. After such agreements were targeted for antitrust enforcement in the late 1990s, parties turned to the monetary and retained exclusivity settlements discussed above. The remaining two settlements are supply agreements involving two drugs, in which the generic firm did not retain exclusivity eligibility. In the case of the first drug, Procardia XL, the generic firm received an immediate license
not only on the 30-milligram strength for which it was the first filer, but on two other strengths as well (Great Lakes Health Plan 2001). In the case of the second drug, Wellbutrin SR, the generic firm relinquished any eligibility for the exclusivity period, and received a license to sell not only the 100-milligram strength for which it was first filer, but also a second strength (Andrx 2003).

A summary of the four categories of pay-for-delay settlements appears in Table 6. Again, the number of settlements is larger than the number of drugs, because—for a few drugs—the brand-name firm entered multiple settlements. For example, for the drug Provigil, the brand-name firm entered monetary settlements with four first filers (Cephalon 2010). In addition, the drugs K-Dur, Adderall XR, AndroGel, and Hytrin had monetary settlements with generic firms that filed ANDAs with Paragraph IV certifications, but were not first filers.

The firms that have entered settlements with both payment and delay are quite diverse: 28 brand-name firms and 25 generic firms in all. Accounting for mergers, the set of generic firms falls to 20. Teva and Barr are considered separately, though they merged in December 2008. The most frequent brand-name settler is Glaxo, with 2 settlements in Table 4 (Zantac and Lamictal) and 8 in Table 5 (Lamictal CD, Imitrex (both tablets and injection), Valtrex, Avandia, Avandamet, Avandaryl, and Paxil CR). Teva and Barr are the most frequent generic settlers, with 11 and 9 settlements, respectively. Barr has been more aggressive than Teva: 6 of its settlements (Nolvadex, Cipro, Niaspan, Provigil, Adderall XR, and Aggrenox), compared to 4 of Teva’s (Lamictal, Effexor XR, Provigil, and Wellbutrin XL), appear in Table 2. One generic firm, Ranbaxy, has played a role disproportionate to its settlement count, reaching
settlements involving the blockbusters Lipitor and Nexium in the span of a few months in 2008.

Although many individual drug makers enter into multiple brand-generic settlements, repeat negotiations between brand-generic pairs are rare. Of the settlements in Tables 2 and 3, Glaxo negotiated with Teva over Lamictal, then later over Avandia, Avandaryl, and Avandamet. Glaxo settled with Genpharm over Zantac, then settled with Genpharm’s successor, Mylan, over Paxil CR. Otherwise, these settlements are the result of one-time negotiations between brand-name and generic firms.

IV. The Evolution in Settlement

Three factors have shaped a continuing evolution in the structure and content of brand-generic settlements: 1) the waxing and waning of antitrust enforcement, 2) a change in judicial interpretation of the Hatch-Waxman Act, and 3) major statutory amendments to the Act in 2003. This evolution poses challenges when choosing an optimal substantive antitrust rule and antitrust decisionmaker, topics taken up in Parts III and IV, respectively.

1. Antitrust Challenges. — The form of settlement varies significantly with the level of perceived antitrust risk, particularly as to monetary settlements. Table 4 depicts this pattern. Monetary settlements occurred at a rate of about one per year from 1993 through 1999. In 2000, the FTC initiated antitrust actions against several settlements, and monetary settlements subsided. In 2005, the government and private purchaser plaintiffs lost antitrust suits in the Eleventh and Second Circuits, respectively (Tamoxifen 2005; Schering-Plough 2005). That year saw monetary settlements as to
three drugs, and in 2006, six more. Moreover, some settlements may be timed to correspond to a depletion in FTC enforcement capacity. In 2008, shortly after the FTC challenged one monetary settlement, there was a renewed flurry of monetary settlements, including Lipitor and Nexium.

The intensity of antitrust enforcement affects not only the fact, but also the form, of monetary settlements. The first monetary settlements—including the first five listed in Table 2—blocked entry until patent expiration, and the brand-name firm paid cash. Starting in 1997, and with increasing frequency after 2000, settling firms changed the standard form of settlements in two ways, both likely responses to increased pressure from antitrust enforcers. (The first such settlement, K-Dur, was negotiated in 1997, and predated increased antitrust pressure (Schering 2005).) First, settlements began to include some pre-expiration entry. That shift provides drug makers with the rhetorical opportunity to argue that the settlement guarantees some competition. Some entry looks better than no entry. From this perspective, the law has shifted in the drug makers’ favor even further than they may have anticipated, given the prevailing view of appellate courts that it is fine to pay for settlements with no pre-expiration entry.

Second, starting in 1997, settlements frequently included not only payment and delay, but also additional contractual terms that tend to obscure whether payment has occurred. The forms of these disguises, and their importance for case-by-case litigation, are discussed in Part III.

2. Judicial Interpretation. — The shift toward settlements with pre-expiration entry has a second cause. Prior to 1998, the FDA had insisted that, in order to enjoy the 180-day exclusivity period, a generic firm must successfully defend its pre-expiration
challenge. In 1998, that view was defeated in the courts, on the ground that it was contrary to the text of the Hatch-Waxman Act, as explained in Chapter 2. After that, a first-filing generic firm could expect to enjoy exclusivity provided it did not lose the patent suit, even if it settled. That made it possible to compensate using retained exclusivity, provided that entry occurred before patent expiration.

The end of the successful defense requirement also created a new form of delay with respect to nonsettling firms. As explained in Chapter 2, this is due to a statutory quirk in the 180-day exclusivity provision: A later-filed ANDA may not be approved until 180 days after either the first filer’s initiation of commercial marketing or a court determination of invalidity or noninfringement. A settlement with the first filer eliminates the possibility of commercial marketing or a court ruling. The 180 days is never triggered, and the later ANDA filer is stuck, for the FDA lacks authority to approve the application, blocking subsequent entry.

Of the 21 monetary settlements described in Table 4, at least 11 appear to create a bottleneck. As for the others, the first 5 settlements predated the demise of the successful defense requirement, and so their effect, at least as of the date of settlement, is debatable. Four recent settlements—Wellbutrin XL, Nexium, Caduet, and Aggrenox—are governed by the new rules, considered below. In the remaining settlement, AndroGel, the first filer abandoned any claim to the bottleneck. Of the 25 drugs described in Table 5, 9 appear to create a bottleneck under the old rules. The remaining 16 are subject to the new rules.

This resulting “bottleneck,” however, is defeasible. If a second generic firm files an ANDA, is sued by the brand-name firm, and wins the patent suit, that decision
triggers the first filer’s exclusivity period. The second ANDA filer can enter 180 days later. (In several early settlements, as discussed in the Appendix, the generic firm disavowed exclusivity eligibility by changing its certification from Paragraph IV to Paragraph III.) To avoid that outcome, the brand-name firm may decline to sue the second generic firm, in which case the generic firm must bring a declaratory judgment suit challenging the patents, win that suit, and then wait 180 days. For some later filers, this route was blocked by the Federal Circuit’s view that the generic firm lacked standing to bring suit, a roadblock that was later cleared by judicial interpretation (MedImmune 2007).

3. Statutory Change. — Statutory change represents a third possible source of evolution. The actual change wrought by the MMA, however, has been unexpectedly small. The changes were designed, in part, to curb anticompetitive settlements. The most important change was a new forfeiture procedure (§ 355(j)(5)(D)), which causes a generic firm to lose its entitlement to the exclusivity period under certain circumstances described below. The MMA’s passage led commentators such as Bernard 2008 to conclude that the settlement problem had been resolved.

As of 2008, however, there was little evidence that settlements featuring both payment and delayed entry have become less popular. As noted above in Table 4, monetary settlements have been a common occurrence after 2003; if anything, they appear to have increased in frequency. The incidence of monetary settlements for blockbuster drugs has increased. The most important settlements, preserving brand-name profits on blockbusters such as Lipitor, Nexium, and Plavix, occurred after the statutory change. The only blockbuster settlement that predates the MMA is Zantac.
That 1995 settlement also preceded significant antitrust enforcement efforts and avoided antitrust scrutiny.

One reason for the limited effect is that the new forfeiture regime only applies prospectively. It is limited to drugs for which the first ANDA was filed after December 8, 2003 (MMA 2003). (There is an immaterial exception. One basis for forfeiture, an unappealed or unappealable determination that the agreement violates antitrust law, applies also to “old” ANDAs.) Most drugs, therefore, are governed by the old regime. Patent litigation frequently takes four or five years to reach settlement. In the Lipitor litigation, for example, a generic firm first filed an ANDA in 2003, but the firms did not settle until 2008. All but 4 of the monetary settlements depicted in Table 4, and 9 of the 25 retained exclusivity settlements in Table 5, were reached under the pre-MMA rules. In short, even if the pre-MMA regime is only transitional, it remains important.

Moreover, even when fully applicable, the new forfeiture rules do little to curb pay-for-delay settlement. Like the old rules, they permit a brand-name firm to neutralize the first filer’s challenge through settlement. That first filer still has the largest incentive to challenge the patent because only it is eligible to receive the 180-day reward. Further, the new rules appear to retain a version of the bottleneck (FDA 2008).

The rules on forfeiture are famously complex. Under § 355(j)(5)(D)(i)(I), forfeiture for failure to market applies only upon the satisfaction of two statutory conditions. The first condition is relatively easy to satisfy—75 days after the first filer’s effective date of FDA approval or 30 months after the application was filed, whichever comes first. The second condition is triggered only if an appeals court rules that the relevant patents are invalid or not infringed, if a settlement reaches a similar result, or if
the brand-name firm withdraws the relevant patents from the Orange Book. Without a trigger, neither is there a forfeiture. (This is not the only possible interpretation, since a court might conclude instead that the certification asserting patent invalidity or noninfringement was not “lawfully maintained,” and hence no longer support the generic firm’s status as a “first applicant” under § 355(j)(5)(B)(iv)(II)(bb).)

The new bottleneck, like the old one, is defeasible; a later-filing generic firm can break the logjam by winning its challenge and waiting 180 days. The post-MMA rules make the relevant condition for defeasement an appeals court win, rather than a district court win—a condition now applicable to both post-MMA and pre-MMA drugs. The complex details are set out in a footnote. This change delays further the moment of generic entry.

The foregoing survey has several implications for antitrust enforcement. First, it demonstrates that the settlement issue is a first-order enforcement question. The one-year benchmark measures discussed in Part III, $12 billion for monetary settlements and $4 billion for retained exclusivity settlements, imply a total $16 billion transfer from buyers to sellers. Thus, the size of the buyer overcharge from pay-for-delay settlements likely exceeds $16 billion. Again, that figure leaves out any effect from increased utilization due to competitive prices. The large implications for consumer welfare justify vigorous FTC and private enforcement efforts and continued scholarly investigation of the evolution and effect of settlements.

---

Prior to the MMA, a generic firm’s district court win triggered the running of the exclusivity period (FDA 2000). The FDA had previously taken the view that the generic firm could wait until an appeals court ruling without triggering exclusivity, in order to avoid the choice between launching at risk and losing exclusivity. The MMA restores the appeals court trigger for pre-MMA ANDAs (MMA 2003). For new ANDAs, the rule is analogous. Forfeiture (rather than triggering) of exclusivity occurs 75 days after a generic firm’s appeals court win, § 355(j)(5)(D)(i)(I)(bb)(AA), and provided that the first condition discussed in the text is also satisfied.
The survey also underscores the importance of prompt Supreme Court review. In terms of their practical importance, the impact of drug patent settlements is at least comparable to other antitrust issues on which the Supreme Court has granted certiorari. By way of comparison, resale price maintenance, the subject of a recent major Supreme Court case, has long been avoidable for most well-counseled firms (Leegin 2007).

Moreover, settlement has become a patent issue, not only an antitrust issue. Although framed as an antitrust case by plaintiffs, the Federal Circuit has embraced the view that settlement is essentially a patent issue, governed by patent law—indeed, arguably governed by Federal Circuit law (Cipro 2008)—and that patent law trumps antitrust doctrine within the nominal scope of the patent. The settlement issue fits well with other patent cases on which the Court has taken certiorari in recent years, and with the Court’s effort to combat perceived hypertrophy in the claimed extent of patent protection (KSR 2007; eBay 2006).

The MMA provisions targeting anticompetitive settlements provide no basis for postponing review. The “transitional” pre-MMA rules have had a significant ongoing impact. One of the first pay-for-delay settlements concerned an ANDA filed in 1985; the certiorari petition in the resulting antitrust suit was filed 21 years later (Joblove 2007). Antitrust challenges regarding ANDAs filed in 2003 or earlier are likely to remain pending for quite some time. Because post-MMA ANDAs are governed by similar rules, a Court decision about a pre-MMA case largely controls the analysis for post-MMA cases as well.

This aggregate survey reveals a final advantage of prompt review. Antitrust challenges to early settlements are still making their way to the Court. An example is
Cipro, currently the subject of a request for en banc review in the Second Circuit. These contain payment and delay, but not much else. Later settlements, however, add contractual complexity. They add difficult factual layers—Was there payment? Was there delay?—atop the legal question of whether payment in exchange for delay violates antitrust law. For a Court that dislikes wading into factual complexity, the early cases provide a more attractive vehicle for setting a clear rule.

V. Identifying an Error Cost-Minimizing Rule

This Part examines how an aggregate approach affects the choice of a substantive antitrust rule. Part V.A highlights one particularly troubling element of the evolution in settlements: the rise of side deals that disguise the fact of payment in a pay-for-delay settlement. Part V.B demonstrates that the exchanges seen in these side deals, though common in settlements, are uncommon otherwise. Part V.C argues that the absence of similar deals outside the settlement context provides a basis for presuming that side deals are disguised payments for delay, not for value.

A. The Rise of Side Deals

As explained in Part II.B, the earliest settlements were straightforward affairs. The brand-name firm paid cash in exchange for the generic firm’s delayed entry. The largest “naked” cash payment was nearly $400 million, which Bayer agreed to pay Barr in settling litigation over Cipro, a major antibiotic (Cipro 2010).

In the wake of increased antitrust scrutiny, naked payments have given way to more complex arrangements. Today, side deals take two complementary forms:
overpayment by the brand-name firm for value contributed by the generic firm, and underpayment by the generic firm for value provided by the brand-name firm.

1. **Overpayment by the Brand-Name Firm.** — In the most common type of side deal, the generic firm contributes—in addition to delayed entry—some further value, such as an unrelated product license. The additional term provides an opportunity to overstate the value contributed by the generic firm and claim that the cash is consideration for the contributed value, rather than for delayed entry. In reviewing K-Dur, the earliest settlement with this type of side deal, the Eleventh Circuit accepted such a factual assertion, which provided a basis for rejecting antitrust liability (*Schering* 2005).

Side deals are now a regular feature of entry-delaying settlements. The contributed value can include a wide range of product development, manufacturing, and promotional services. In some deals, the generic firm offers a product or patent license, or agrees to develop a new product. In one variant, the generic firm develops a new formulation of the brand-name drug. In other deals, it agrees to furnish manufacturing services to the brand-name producer, or to provide inventory, or even to provide “backup” manufacturing services. In some cases, the generic firm provides promotional services as to the product at issue, related drugs, or unrelated products. For some drugs, the brand-name firm reaches entry-delaying settlements with multiple generic firms, each with side deals. These deals, and the drugs to which they correspond, are summarized in Table 4.

Some of these arrangements are suspect on their face. It may seem clear that the brand-name firm does not need a patent license that does not clearly cover its product,
new drug development that is unrelated to its current core business, a new source of raw material supply, backup manufacturing, or additional promotion. For example, in the case of a settlement involving the wakefulness drug Provigil, the brand-name firm, Cephalon, apparently was aware of one generic firm’s intellectual property for three years before showing any interest in seeking a license (FTC 2009). However, not all such settlements are facially absurd. In some cases, the generic firm has plausible expertise in the subject of the side deal. For example, a brand-generic agreement involving Adderall XR described Barr investments in drug delivery technology, to be exploited in new product development by the brand-name firm as part of settlement (see Appendix for details). It is very difficult to be certain that a deal is collusive without a deep and complex inquiry into the business judgment of the two drug makers.

2. **Underpayment by the Generic Firm.** — The brand-name firm, rather than paying too much, can charge too little. One mechanism involves “authorized generic” sales. These are sales made by a generic firm under the brand-name firm’s FDA approval. The brand-name firm supplies the product to the generic firm at a discount, which the generic firm then resells under its own label at a profitable price. The compensation is buried in the discounted price offered by the brand-name firm.

In several early settlements, the authorized generic product was launched at the time of settlement (*Great Lakes Health Plan* 2001; *Tamoxifen* 2006). This practice fell out of favor after a court concluded that the authorized generic sales triggered the 180-day period (*Mylan* 2001). Some modern settlements avoid the trigger problem by providing for authorized generic sales only after another generic firm enters, or of a drug other than the subject of the generic firm’s ANDA filing. For example, in settling
its Nexium litigation, AstraZeneca made Ranbaxy an authorized generic distributor of Prilosec and Plendil. The Effexor XR settlement granted the generic firm an early license to sell an immediate-release version of Effexor. The Niaspan settlement provided a license as to Advicor. The Propecia settlement appears to be a fourth example. There, Merck made Dr. Reddy’s an authorized generic distributor of Proscar and Zocor around the same time that the parties settled litigation over Propecia. Another option is to make the generic firm an authorized generic in another country.

In a related form of discounted sale, which avoids the trigger issue, the brand-name firm sells an entire product line to the generic firm. One settlement involving an extended-release version of a drug (Adderall XR), for example, transferred (for a possibly discounted price) the immediate-release version to the generic firm. In a more complicated set of deals, a brand-name firm may have sold a generic firm rights to one product, and the generic firm delayed entry in two other products.4 (A further variant of this strategy, simultaneous settlement of multiple drugs with uneven entry terms, is considered in Part VI.C.) Once again, it is very difficult as a practical matter for a decisionmaker to know whether the transfer price provides compensation from the brand-name firm to the generic firm, and if so, how much.

B. Infrequency Outside of Settlement

Outside of settlement, brand-name firms seldom contract with generic firms for help with the activities that form the basis of side deals. Indeed, as a general matter, 

---

4 Galen sold Barr rights to Loestrin, and Barr delayed entry as to two other products, Estrostep and Femhrt (Lavery 2003). This transaction is not included in Table 4 because the Loestrin sale was completed first. That ordering limited the degree to which a Loestrin sale could confer compensation upon Barr, in exchange for delayed entry on Estrostep and Femhrt, because Barr could simply walk away with its Loestrin “quid” without providing an Estrostep/Femhrt “quo.”
brand-name and generic firms seldom execute major deals outside the settlement context, with the exception of authorized generic arrangements, which necessarily are reached between a brand-name firm and a generic firm.

A review of the annual securities filings of settling drug makers supports this proposition. To examine the extent of business dealings outside of settlement, five major brand-name firms (Abbott, AstraZeneca, Bristol-Myers Squibb, GlaxoSmithKline, and Pfizer) and five major generic firms (Barr, Mylan, Ranbaxy, Teva, and Watson) were chosen based upon their frequency of settlement activity and economic importance. For each brand-name firm, annual filings between 2000 and 2007 were searched for the names of the five generic firms. Each resulting “hit” led to further examination, to see whether the discussion indicated a business relationship between the two firms, as opposed to, say, a description of litigation or competition. The business transactions were examined further using articles in the trade press and other materials. The same exercise was performed for each of the generic firms, as to each of the five brand-name firms.

The resulting inquiry into twenty-five total brand-generic business dealings—each of five brand-name firms, with each of five generic firms—produced just two responsive business arrangements, both of them involving Ranbaxy. Ranbaxy has an arrangement with Glaxo to take “hit” molecules that show initial promise, and help develop and winnow them into “candidates” for further development by Glaxo, and potentially to develop them further to the point of a new investigational drug application in India. Ranbaxy also bought the rights to a set of minor dermatology drugs from Bristol-Myers Squibb. There were also several other business arrangements that do not
match the terms of the side deals discussed above, such as an agreement by Bristol-
Myers Squibb to commercialize a drug initially developed by generic firms. The
absence of significant evidence of business arrangements is not decisive; such non-
settlement deals could exist, yet be too insignificant to report in an annual filing. If so,
however, they are apparently not of first-rank importance to the operations of the firm.

Further evidence about the firms’ limited business dealings, outside of
settlement, is revealed by one specific type of side deal known as co-promotion. Brand-
name firms frequently enter co-promotion arrangements to augment their promotion
efforts—for example, to reach physicians that their own detailing team does not visit. In
a second search, the same annual filings were reviewed for mentions of promotion, and
those mentions which pertained to product promotion were examined further. That
search produced many examples in which a brand-name firm recruited other brand-
name firms to help promote a drug, but no significant examples, outside the settlement
context, in which the brand-name firm recruited a generic firm to promote a brand-name
drug. A minor exception is Ranbaxy’s promotion of a Sanofi vaccine in India. On the
other hand, generic firms do occasionally have significant branded drugs, and the search
did reveal instances when they have hired brand-name firms to help market the drug,
most prominently Teva’s multiple sclerosis drug Copaxone, for which Teva recruited a
predecessor of Sanofi-Aventis.

This result is not surprising, considering the business of generic firms. Generally, they do not have substantial promotion teams, for they seldom have major
branded drugs to promote. The absence of generic provision of other services, outside
the settlement context, is equally unsurprising. The research and development capacity
of generic firms is generally limited, although some firms have made efforts to develop
a brand-name drug business—most notably Teva, whose efforts have produced two
products, Copaxone and Azilect. This is not the core business of generic firms.

Nor do they have powerful manufacturing capabilities such that they would be
the obvious and efficient alternative supplier for a brand-name firm. As an example,
Cephalon agreed to buy Provigil’s active ingredient from a third generic firm, even
though the firm had not manufactured the product and Cephalon already had an
adequate source of supply (Cephalon 2009). The contrast is less severe in side deals
featuring transferred assets. It is quite common for a brand-name firm to set up an
authorized generic arrangement with some generic firm. Transfers of product lines to
other drug makers are common as well.

C. Adopting a Presumption of Payment

Viewed in isolation, it is difficult to tell whether a side deal represents payment
for value or disguised payment for delayed generic entry. A broader comparison of side
deals in conjunction with settlements, versus brand-generic deals outside this context,
tells a different story. At least with respect to overpayment side deals, the absence of
brand-generic deals outside of settlement is a strong reason to suspect that the deals are
used to pay for delay.

In such cases, it is appropriate to impose a presumption that the side deal
provides disguised payment to the generic firm. Under this pay-for-delay presumption,
drug makers would be free to come forward with evidence that their unusual deal was
for value and therefore raises no anticompetitive issues. That burden is most
appropriately placed upon them, as the least-cost providers of the necessary information. An alternative approach, also supportable by the evidence from aggregation, would make this presumption conclusive.

That conclusion is not, by itself, enough to impose liability. It resolves the “factual” question of whether a settlement containing a side deal constitutes payment for delay, but not the “theoretical” question, emphasized in Chapter 2, of whether pay-for-delay settlements violate antitrust law.

This proposal, like any aggressive antitrust rule, is potentially overinclusive. It raises the probability of false condemnation. But here, the rarity of such arrangements outside of settlement lowers the likelihood of false positives. The error cost analysis has a further component: How costly are false positives when they occur? Not very costly, as it turns out, because the generic firm is seldom a distinctive source of the particular value in question.

The rule comports with the comparative rigor with which we treat collusive activity generally. Antitrust’s lenient approach to exclusionary conduct reflects an error cost calculation focused upon false positives. As noted in the introduction to this chapter, decisionmakers think that true positives are rare and difficult to distinguish, and also that false positives are particularly costly, because they amount to condemnation of the “very conduct” (competitive price cuts) that antitrust is supposed to protect. As Gavil 2004, Salop 2006, and Leslie 2006 have noted, false negatives are an important countervailing problem. For collusion, by contrast, avoiding false negatives is the important goal, particularly where false positives are rare and low-cost, and where no significant equilibrating factors tend to restore competition. That relatively aggressive
approach is shared even by “Chicago School” analysts, who support an aggressive enforcement emphasis upon collusion (Posner 1969).

What about underpayment side deals? The likelihood of false positives is higher, compared to overpayment deals, because authorized generic arrangements and product transfers frequently occur outside the context of settlement. The cost of false positives remains low, however, due to the absence of distinctive value arising from dealing with the particular generic firm that happens to be locked in a patent suit with the brand-name firm.

The high cost of false negatives and low cost of false positives support a presumption in the underpayment context, just as in the overpayment context. A more conservative alternative would be to make the presumption applicable only to future settlements. That way, parties have ample notice that they must not reach underpayment deals with parties with which they are settling. Given the absence of distinctive value offered by the settling firm, that route places at most a minimal burden upon parties that wish to reach authorized generic or asset transfer arrangements.

This policy suggestion could be implemented by several routes. For example, it could be adopted by a court considering a particular case, using the federal courts’ common lawmaking authority under the Sherman Act. Alternatively, it could be instituted through new congressional legislation, or promulgated as an agency rule by the FTC. The next part considers the strengths and weaknesses of these alternative routes.
VI. Institutional Choice: What Should the FTC’s Role Be?

This Part turns to the institutional question of who should employ this aggregate approach to antitrust questions. Part VI.A explains why an agency—here, the FTC—is better positioned to collect and synthesize aggregate information, relative to courts. Part VI.B argues that this advantage in wielding aggregate information favors a shift in substantive policymaking authority from courts to agencies.

A. Information Gathering and Synthesis

A court establishing antitrust policy faces the fundamental problem that it has little capacity to collect aggregate data. The disadvantage of a court as a fact-finder is a familiar idea from the literature on institutional choice, well-canvassed in Lemos 2006. The problem is particularly acute here. At best, a single court needs many years to develop a sense of the overall distribution of cases, as antitrust cases appear only rarely on its generalist docket. The Supreme Court is in a slightly better position, since it is exposed to appeals from all over the country. However, many instances of anticompetitive behavior are never litigated, and courts have particularly limited ability to observe nonpublic data about settlements outside the case at bar.

Private parties cannot entirely fill the gap. These plaintiffs struggle to learn the content of settlements, with some early agreements, such as the Zantac settlement discussed in Chapter 2, escaping notice entirely. Later settlements have been shielded from scrutiny due to the difficulty of discerning, from public information, the extent of pay-for-delay deals. This information gap partially explains why so few of the most recent settlements have been challenged.
This chapter helps fill the gap, but it is not a complete solution. My data does not include nonpublic details that would help build confidence about whether a side deal conveys payment. For example, how much did the brand-name firm agree to pay for a co-promotion agreement? How much did a generic firm pay for a product transfer? Is payment conditioned on successful performance by the other party? Was a particular product development deal a long-felt need of the firm, which shopped for alternative sources? How was the service provided valued internally by the payor? Public data for most settlements lack these details.

Outside the context of side deals, two other issues are important. First, do the parties expect the generic firm to retain exclusivity when it enters the market? In some cases, one or both parties divulge their view publicly, but in other cases they do not. Second, how often does the brand-name firm contract with this counterparty and other generic firms outside the context of settlement? Reciprocally, what is each generic firm’s experience with brand-name firms outside the context of settlement? Public information of the type collected in Part II paints only an incomplete picture of the frequency of particular arrangements outside the settlement context. With details such as these, an inference of payment for each case could be strengthened, and, more importantly, the inference of payment across cases could be strengthened as well.

The FTC already has in place all the tools it needs to perform this task. As noted in Part I.B, it receives information about each settlement and has statutory authority to require firms to produce additional information of the types discussed above. That authority ought to be used to collect two types of information. First, the FTC should seek full details about each settlement—at least enough information to answer the
questions listed above. Some of these questions may be answerable by examination of the agreement itself. To the extent they are not, the gaps could be filled using voluntary questionnaires or, if necessary, compulsory process. Second, the FTC should collect from each brand-name firm a detailed catalogue of its dealings with generic firms, and vice versa for generic firms.

This information would be the key input in a comprehensive study of side deals. It would provide a firm basis for the FTC to endorse or reject the conclusion offered in Part II, based upon public information, that contemporaneous side deals should possess a presumption of payment. If the information is sufficiently lopsided, error cost minimization might suggest the more aggressive rule should be instituted, making the presumption of illegality conclusive and effectively banning contemporaneous side deals.

In this respect, the analysis in Part II provides a rough draft for a more comprehensive, future agency report. The public data presents a prima facie case that something is amiss regarding the increasing utilization of side deals. For skeptical readers of this chapter, who may think that the survey results reported in Part II are too weak to justify a presumption of payment through side deals, the case for deploying the agency as an aggregator should be even stronger; agency action is necessary to fill these informational gaps and better explain whether and when compensation is conferred for delay.

The FTC has not fully exploited its information gathering advantage. Of the drugs with monetary settlements in Table 2, two-thirds occurred after the end of the FTC’s last major study in 2002. Moreover, all of the retained exclusivity settlements in
Table 3 post-date the study. To be sure, the FTC evaluates each individual agreement to determine whether further investigation is appropriate, and no doubt it asks some of the questions detailed above in considering its response. But it does not synthesize the resulting information, aside from very general annual summaries of settlement activity. As this chapter reveals, only such an aggregate approach can generate a useful picture of—and rule for—brand-generic settlement.

B. Antitrust Rulemaking

The previous section advocates a focused increase in the FTC’s “competition policy research and development” (Muris 2003). If the FTC accepted the suggestion, it would eventually reach a firm, empirically grounded conclusion about the optimal policy for side deals, and thus either confirm or reject the conclusion reached in Part II. That conclusion could be deployed in a variety of policymaking settings, including litigation brought by the FTC, amicus practice, and advocacy for congressional legislation. This section considers a further possibility, that a comprehensive aggregate study of settlement practice could form the basis for substantive policymaking by the FTC in the form of rulemaking.

There is of course an enormous literature on the choice of courts versus agencies, adjudication versus rulemaking, and rules versus standards, and this chapter does not engage the full complexity of those debates. My goal here is simply to suggest how the virtues of an aggregate perspective on settlement practice shift the balance in a way that favors agency rulemaking. In other words, the settlement issue highlights certain advantages of moving away from a court-centered model of antitrust law.
Why bother with rulemaking? Even if the expert agency is better than a court at arriving at a correct policy conclusion, thanks to its superior capacity for aggregation, it does not necessarily follow that the agency ought to set policy. It could instead simply furnish the information to a court or Congress, which might then implement the same conclusion, with some of the same benefits—for example, efficiency compared to case-by-case adjudication, and certainty for businesses about the range of acceptable practices.

Put another way, why would we care whether the agency itself makes policy in the first instance, rather than acting as an input to a court? The question suggests a bureaucratic version of the Coase Theorem. If there is no friction in communicating an expert policy conclusion from the agency to the court, then it does not matter which of the two has policymaking authority. If, on the other hand, the agency’s message arrives garbled or is ignored by the court, that provides a reason to prefer that the agency reach a substantive policy judgment of its own, rather than merely furnishing advice to the court.

One reason to expect the court to do a less effective job is that courts have trouble correctly identifying anticompetitive strategic behavior (Hovenkamp 2005), particularly in a setting as complex as the Hatch-Waxman Act. That view is borne out by a recent appeals court opinion about settlement. The court relied, as a reason to deny antitrust liability, upon the mistaken idea that a settlement with one generic firm would spur other generic firms to action, and that these firms would have the large incentive provided by the exclusivity period (Tamoxifen 2006). In fact, later filers are ineligible for the exclusivity period. This error was unforced; the point does not appear to have
been argued below. (As noted in the previous chapter, in 2010, the Second Circuit acknowledged this error (*Arkansas Carpenters* 2010).) The same court took comfort in the view that often there is more than one generic challenger, and the court concluded that multiple challengers are difficult to buy off. In fact, however, multiple settlements do happen.

Courts have also had trouble evaluating the facts of particular cases. For example, in the case discussed above, the plaintiff had argued that the brand-name firm compensated the generic firm not only with cash, but also through authorized generic sales. The court ignored this idea entirely. In a second case focused on side deals, the appeals court essentially ignored the extensive evidence that the payment was for delay, rather than the separate value offered by the generic firm (*Schering* 2005). This pattern is likely to continue, given the evidence of complexity discussed in Part III.A.

An expert agency, essentially by definition, is less likely to make mistakes identifying the strategic behavior of parties. To be sure, this information could be communicated to a court. But as a practical matter, courts have not welcomed the information about settlements supplied by the FTC. In a key case brought by the FTC, the appeals court largely ignored the analysis employed by the agency, granted essentially no deference to its findings of fact, and indeed berated the agency for failing to follow the appeals court’s earlier rule (*Schering* 2005; *Crane* 2008). For the most part, courts have also ignored the results of the FTC’s extensive 2002 study and its subsequent annual summary updates, as well as its amicus recommendations based on this data.

A second reason to expect courts to be less effective than the FTC is that
antitrust courts are obliged to impose treble damages when they condemn behavior as a violation of the Sherman Act. The large measure of damages may strike a court as excessive, particularly where the conduct seems ambiguous or complicated, such that the parties might not be expected to know that their behavior violated antitrust law. That impression may be reinforced where the conduct is out in the open, rather than hidden, so that a usual justification for a damages multiple—the difficulty of detection—is missing. As suggested in Lemos 2008, the combined effect is to make a court gun-shy, and to cause it to select a deliberately underinclusive antitrust rule, a trend exemplified by cases such as Twombly 2008 and Matsushita 1986, which interpreted the appropriate standard for pleading and summary judgment in antitrust suits. Indeed, courts rejecting antitrust liability for settlements have repeatedly adverted to treble damages in their analysis (Tamoxifen 2005; Valley Drug 2003).

The FTC is less constrained. Its substantive conclusions would be made under the Federal Trade Commission Act’s prohibition of “unfair methods of competition,” rather than under the Sherman Act. The Supreme Court has stated repeatedly that the FTC Act’s prohibitions are broader than those of the Sherman Act (Indiana Federation of Dentists 1986; Brown Shoe 1966). Thus, behavior that constitutes unfair competition does not necessarily also violate the Sherman Act’s prohibitions of unreasonable restraints of trade or monopolization.

This conclusion is resisted by some observers, such as Posner 2005, who think it is “no longer tenable” to treat the FTC Act as broader than the Sherman Act. However, the Supreme Court’s rejection of strict equivalence can be justified on an eminently pragmatic ground. The argument for equivalence rests upon the proposition that, as
Posner 2005 puts it, “the Sherman and Clayton Acts have been interpreted so broadly that they no longer contain gaps that a broad interpretation of Section 5 of the FTC Act might be needed to fill.” But to the extent that the Sherman Act, as actually interpreted by courts, contains important gaps, as exemplified by the lack of liability for pay-for-delay settlements, the quoted statement does not hold. Where, as here, courts reach incorrect conclusions about liability, distinguishing the two statutes is useful because it allows the FTC to enjoin settlements without being automatically reversed by a court equipped with the (erroneous) view that antitrust law does not extend so far.

Moreover, nonequivalence is particularly useful where, as here, treble damages lead courts to constrict the scope of liability. Even if it is appropriate for courts to constrict liability to compensate for the heightened false-positive risk created by treble damages, it does not follow that the FTC must adhere to the same path. The FTC seeks injunctive relief, not treble damages. That difference reduces concerns about false positives and overdeterrence. Put another way, the FTC’s optimal scope of liability may well be broader than the courts’. Nonequivalence allows the FTC to take advantage of that difference, compared to the Sherman Act, which, as Arthur 2002 notes, applies a harsher penalty to a narrower class of activity.5

A third advantage of the FTC is that it is less subject to the constraint of stare decisis. Lower courts are bound by their own or Supreme Court precedent. The Supreme Court, for its part, is not quick to revisit antitrust doctrine—it took close to a century for the Court to reverse its per se rule against resale price maintenance (Dr.

5 The nonequivalence also provides an answer to the understandable concern, raised in Crane 2005, that the shared authority of the Antitrust Division and FTC over antitrust matters might undermine the FTC’s claim to Chevron deference. Crane points out that the FTC has characterized its powers as being co-extensive “for the most part” with the Sherman Act, and suggests that it might be necessary to combine the two enforcers before granting Chevron deference. But the quoted language also underscores the non-identity of the two statutes, consistent with the Supreme Court’s long-held view.
Miles 1911; Leegin 2007)—and frequently feels constrained to follow its own previous views. The FTC is freer to change course, provided that the new interpretation is a reasonable understanding of the FTC Act.

One way for the FTC to exploit these advantages is to promulgate a legislative rule—that is, a rule having the force of law and entitled to deference by a court.6 A Supreme Court case, Chevron 1984, imposes upon courts a “duty to defer to reasonable agency interpretations . . . [of an ambiguous] statute that an agency is charged with administering” (Merrill and Hickman 2001). FTC rulemaking has been suggested periodically by commentators (Elman 1964; Baker 1993; Kovacic 1997; Balto 2005) as a way to shift decisionmaking authority to the FTC and fill gaps in the coverage of other antitrust statutes. The rulemaking route, though not without controversy, is an attractive and feasible means to take full advantage of the aggregate approach to settlement.

The FTC possesses the power to promulgate rules with the force of law that are subject to Chevron deference. The FTC has promulgated one such antitrust rule in 1968 (FTC 1968), pursuant to sections 2(d) and 2(e) of the Clayton Act. The rule was eventually rescinded, apparently without having been used by the agency in law enforcement (FTC 1994).

The FTC’s rules and operating procedures do not deny the agency’s possession of this authority, to be administered under ordinary notice-and-comment procedures, but neither do they fully spell it out. According to the FTC’s Operating Manual (FTC 2010), “the Commission has statutory authority under FTCA § 6(g) to promulgate rules respecting unfair methods of competition,” and a particular case (Petroleum Refiners, 580, 1968).

---

6 Some commentators use the term “substantive rules” instead of “legislative rules” for the rules I have in mind, but the choice of terminology is unimportant for my purposes. Gersen 2007 provides an account of the overlap.
discussed below) gives the FTC “authority to promulgate rules with substantive effect.” The agency’s Rules of Procedure accommodate antitrust rulemaking too (FTC 2008).

The FTC’s rulemaking power arises from 15 U.S.C. § 46(g), its general rulemaking authority, which the D.C. Circuit interpreted in *National Petroleum Refiners Ass’n* (1973) as a grant to make rules with the force of law. The FTC has two distinct missions, consumer protection and antitrust, and *Petroleum Refiners* specifically dealt with a consumer protection rule, not an antitrust rule. But the relevant statutory language covers both consumer protection and antitrust rules, and the applicability of the court’s ruling to both types of rules is fairly implied in its opinion.

The *Petroleum Refiners* result is doubtful as an original matter. As Merrill and Watts 2002 argue, when Congress passed the FTC Act and other statutes, it omitted any sanction for rule violation, thereby evidencing an intent to deny legislative rulemaking authority. Nevertheless, today it is relatively settled that an ambiguous statute, such as the FTC Act, suffices to confer that authority. After all, as Merrill and Watts also discuss, similar rulings have been made for other statutes, and overruling *Petroleum Refiners* today would jeopardize rulemaking in these other contexts, including the grant analyzed in *Chevron* itself. These prudential considerations tend to confirm the viability of rulemaking authority. Although as noted above the FTC reportedly sought candidates for antitrust rulemaking after *Petroleum Refiners*, it has not yet found any. A rulemaking focused on settlements is an attractive candidate if this procedural route is pursued again.

Rulemaking is not the only way to shift substantive policymaking authority from courts to the FTC. The FTC can bring individual cases through agency adjudication.
using its authority under § 5(b) of the FTC Act, or directly in an action in district court. The FTC has taken both routes in attacking settlements. The agency adjudication route resulted in an appeals court loss; two cases in district court are pending.

Rulemaking has significant, familiar advantages over the adjudicatory route. Rulemaking permits affected parties to test aggregate data in an open way, with ample opportunity for rebuttal. As legal scholars such as Pierce 1988 have noted, the opportunity for input and testing improves the quality of resulting policy. The resulting rule thus has a superior claim to judicial deference, compared to judicial review of a single case: The rule has been thoroughly vetted under notice and comment, after a broad, deep review of the full terrain of behavior by regulated parties. It is this superior breadth and greater vetting, rather than the doctrinal force of Chevron itself, that presents the strongest reason to think that a rule might succeed where adjudication has failed.

Rulemaking helps in another way. The FTC Act is broader than the Sherman Act, as noted above, but the degree of its additional breadth has been a subject of controversy. Some lower courts have regarded with skepticism the FTC’s efforts to regulate behavior not already governed by the Sherman Act (Ethyl 1984; Boise Cascade 1980). A powerful way for the FTC to overcome this skepticism would be to support its claim to authority with aggregation, buttressed by notice-and-comment rulemaking. In this manner, the FTC could combine, in a mutually reinforcing manner, the two ways in which its authority stands out relative to ordinary, judicial antitrust policymaking: having a statute with broader reach than the Sherman Act, and possessing the power to collect information beyond the reach of the judiciary.
At the same time, some of agencies’ distinctive disadvantages seem less pronounced here. A shift from courts to agencies raises concerns about an agency’s comparatively greater vulnerability to capture by regulated parties. As applied to the FTC, this concern finds some support in the early history of the FTC, where a protectionist attitude toward small businesses in certain industries can be plausibly attributed to capture (Posner 1969). Moreover, the settlement issue is currently of concentrated interest only to the pharmaceutical industry, making the capture concern particularly salient, although one could imagine insurers and other drug purchasers providing a counterweight.

On the other hand, the modern FTC is a much more effective organization today than the agency that received so much criticism several decades ago, and has erased the taint of the earlier capture critique (Posner 2005). Its newfound success can be attributed in part to a bipartisan consensus about the role of economic analysis in modern antitrust law. In any event, whatever the general merits of this characterization, it seems inapplicable to the settlement issue, where FTC commissioners across the political spectrum have been unanimous in their view that settlements raise serious competitive concerns.

C. Responding to Novel Forms of Regulatory Avoidance

Settlement practice continues to evolve to exploit regulatory complexity. The usual assumptions about settlement are that it entails an agreement, by which cash or its equivalent is exchanged for entry, and where the entry date is constrained to be no later than patent expiration. In fact, the forms of payment and even the fact of agreement are
manipulable. The following examples from recent settlement practice bear this out.

1. **Multiple Settlement with Uneven Entry.** — In some instances, the brand-name and generic drug makers settle several disputes at the same time, affording the brand-name firm an opportunity to pay the generic firm for delayed entry on one drug by granting early generic entry on a second drug. Consider, for example, Lamictal, a blockbuster epilepsy treatment that is offered in both chewable and nonchewable forms. A generic firm launched a pre-expiration challenge to each form. Both challenges centered upon the same patent. In the joint settlement of both disputes, the generic firm received a license to the chewable version that permitted entry three years before entry on the nonchewable version (Teva 2005).

Uneven entry does not automatically raise pay-for-delay concerns. For example, a one-year delay as to one drug might exactly offset a one-year acceleration of entry on a second drug of equal importance. More generally, if a generic firm’s interests are aligned with consumer interests, there is little to worry about, because a generic firm will insist upon early enough entry (and increased consumer welfare) on one drug to compensate for the reduced generic entry (and consumer welfare) on the other drug. Of course, in such a situation, it is difficult to see why the parties would bother with uneven entry. The explanation is that the drug with early entry is the one on which the parties expect comparatively little incremental entry from other generic firms. In the case of Lamictal, the nonchewable version is far more important than the chewable version; in fact, the chewable version had low enough sales (about $50 million annually) as to be unlikely to attract additional generic challengers. A second example is Barr’s settlement of Provigil litigation, wherein Cephalon granted a slightly earlier
entry date as to another drug, Actiq, on which Barr already had a license providing for entry under certain circumstances.

2. Probabilistic Payment. — A special case of the uneven entry strategy arises when entry as to one of the drugs has already occurred, and there are accrued damages—probabilistic, as the patent suit has not yet been resolved—that the brand-name firm can forgive as part of the settlement. Lipitor is again exemplary. Pfizer and Ranbaxy had done battle on a second significant drug, Accupril. Ranbaxy had launched a generic version of Accupril without waiting for a district court to rule whether Pfizer’s patent was valid and infringed (Pfizer 2005). Pfizer secured a preliminary injunction, which was affirmed by the Federal Circuit. The rulings indicated a high likelihood that Pfizer would win on the merits. Accupril has annual sales of about $400 million, so the expected damages were likely at least tens of millions of dollars. Thus, at this point, Pfizer’s damages claim against Ranbaxy, although probabilistic, was large in expected value. The Lipitor settlement also “resolved” the Accupril dispute, likely by forgiving the accumulated expected damages. The residual uncertainty about the terms of this settlement helps illustrate why the FTC’s role is so important.

The forgiveness strategy can be applied not only across several drugs, but also across several strengths of a single drug. For example, in Wellbutrin XL, the generic firm had challenged the patent applicable to two different strengths of the drug. It launched at risk as to only one strength. The subsequent settlement forgave accumulated damages on the first strength, and delayed entry on the second (Biovail 2007). The logical next step is a launch followed by waiver for a single drug. Rockoff 2009 describes an example involving the acne medication Solodyn. There, the generic firm
launched at risk, then agreed to cease sales until 2011, in exchange for a damages waiver.

3. “No Authorized Generic” Provisions. — As previously explained, retained exclusivity is a source of compensation to a generic drug maker. That compensation is reduced, however, if a brand-name drug maker launches an authorized generic product to compete with the generic entrant, in addition to the brand-name firm’s existing branded product. The brand-name firm can increase the generic entrant’s profits from exclusivity by agreeing not to launch an authorized generic product. This decision is costly to the brand-name firm, which foregoes extra profits from its own authorized generic competitor. Numerous recent agreements include a “no authorized generic” term. Drug settlements with such a term include Adderall XR, Plavix, Effexor XR, and Nexium.

This source of payment to induce delay has attracted some antitrust enforcement attention. The clause was reportedly one reason why antitrust enforcers rejected the Plavix agreement (Carrerou et al. 2006). Logically, if a “no authorized generic” provision raises an antitrust problem, then so does retained exclusivity itself, for the effect of the provision is to raise the value of retained exclusivity.

4. Avoiding Agreement. — Through careful design, settling parties can arrange for delayed entry without any formal agreement as to timing. The parties can condition periodic payment upon nonentry, and make payment a function of brand-name profits that depend upon nonentry. An example is the settlement in Naprelan, which took the form of a royalty paid on brand-name sales (Andrx 2003).

One settlement, involving the drug Altace, appears to have used a variant of this
strategy. There, the brand-name firm acquired a new tablet formulation of the drug and agreed to pay a royalty on its sales (King 2007). This gave the generic firm an incentive not to enter precipitously, as early entry might jeopardize the orderly transition to a new and more profitable formulation. In addition, periodic cash payments, purportedly in exchange for developing the new formulation, were made contingent on unspecified events. This may have been directly for nonentry or indirectly for a successful transition; it is impossible to tell based on the limited data available.

* * *

These examples demonstrate that drug makers are adept at achieving a particular substantive outcome—brand-name compensation of generic firms, combined with delayed generic entry—while altering the form of settlement to evade the most obvious risks of antitrust liability. The continuing shift in strategy here resembles the economics of tax shelters. As a regulatory prohibition becomes more stringent, the cost of noncompliance rises. Some regulated parties will give up and simply comply, resulting in a welfare gain. Others will continue to avoid regulation, and instead shift to new strategies. These strategies are more costly to the firm, for otherwise they would have been chosen in the first place; this shift represents a social loss (Kaplow 1990; Weisbach 2002). Thus, whether an increase in enforcement is warranted depends upon the amount of residual noncompliance, the increase in social costliness of the new behavior, the cost of administering the system itself, and costs resulting from overinclusion—for example, restricting some value-increasing side deals.

The review of settlement behavior in this chapter paints a mixed picture. To be

---

7 A royalty on sales for the acquired tablet product does not appear in the parties’ disclosure of the agreements, but it is mentioned in FTC 2007, which describes a “complex set of transactions” that fit the Altace transactions.
sure, this process of continuing evolution threatens the ability of existing antitrust institutions, particularly courts, to keep pace. Courts are increasingly unlikely to be an effective check on settlement. In part, this is because they are poor aggregators. In addition, courts must be fed cases by either a government agency or a private plaintiff. The FTC has limited case-by-case enforcement capacity; practically speaking, it can bring at most a few pharmaceutical antitrust cases at a time, and they are likely to last for five years or more. That capacity is small, compared to the frequency of pay-for-delay settlements—although a successful enforcement action or two would likely reduce the frequency. Private plaintiffs, meanwhile, are reluctant to bring cases. Having lost the simplest cash-for-delay agreements, why should they take a chance challenging more complex settlements?

Continuing evolution makes the crisis in case-by-case adjudication more acute for another reason. A single appellate or Supreme Court opinion imposing liability does not fully resolve liability for the newest settlements. A win on a simple case is a very helpful start, but only sets the stage in making sense of the more complicated cases. Thus, even if it were settled as a theoretical matter that paying for delayed entry is prohibited, and settled as a factual matter that side deals provide a disguised means to pay for delay, it does not necessarily follow that the newest settlements also violate antitrust law. Given the malleability of side deals, even an on-point judicial decision imposing liability would not preclude firms from arguing that their arrangements were conceptually or factually distinct. On the other hand, if a court is forced to start with one of the most complex cases, without the benefit of affirmative precedent on the simpler cases, correctly identifying liability seems less likely.
Here, too, FTC rulemaking can help. As to new forms of payment, for example, the FTC could set a rule stating that *any* conferral of value by a brand-name firm, if made contemporaneously with a generic firm’s agreement to delay entry, will be considered to exchange payment for delay. Probabilistic damages would clearly fit within that definition. Agreements to preserve exclusivity, whether simply by agreeing not to contest ANDA approval or by an affirmative agreement not to launch an authorized generic product, would also be included, as just another form of payment. In these examples, the aggregation approach helps to identify and respond to emergent settlement practices before they become more prevalent. For settlements without formal agreement, moreover, FTC rulemaking would be even more effective, because the applicability of the FTC Act, unlike section 1 of the Sherman Act, is not conditioned on the existence of an agreement.

Agency rulemaking is not the only possible route for implementing a broadly applicable rule. If a court can be persuaded to think broadly about the implications of settlement, it may adopt a similarly broad ruling, though the considerations above tend to make that less likely. Legislative action is also a possibility. The MMA closed some loopholes, though it preserved others, including the bottleneck and retained exclusivity. Its requirement that an appeals court trigger exclusivity also worsened the delays in an important respect, through a provision buried seven steps deep in the statutory structure: § 355(j)(5)(D)(i)(I)(bb)(AA).

Thus, although Congress could directly implement any of the presumptions or rules discussed above, its ability to do so quickly and correctly is open to significant doubt. These difficulties also suggest a reframing of the legislative project. Rather than
closing identified loopholes with another new layer of complexity, it would be better to remove existing complexity—in particular, by ending retained exclusivity. Simply put, if a generic firm ends its litigation against a brand-name firm, it should no longer be eligible for the exclusivity period.

In a sense, this provision would offer a partial return to the FDA’s original view, described in Chapter 2, that a generic firm must earn exclusivity by winning a patent suit. It would reduce both the amount of payment conferred in a settlement, and the extent to which a settlement delays entry. The provision should apply to both new and existing settlements, like other statutory provisions that have been given retrospective effect. Finally, the political economy of such a statutory change is attractive. If, as some observers might argue, retained exclusivity is not a valuable form of compensation for delay, then its omission from the settlement equation will not be missed, and there is little reason for drug makers to resist this statutory change.

A rule directed to all contemporaneous conferrals of value by a brand-name firm would appear to resolve the pay-for-delay issue, closing the avenue for escape to yet further forms of regulatory avoidance. This appears to be true even as to informal contemporaneous understandings reached between parties, as in the Altace example above. In tax planning, informal alternatives to contract greatly expand the opportunity for avoidance (Raskolnikov 2008).

That problem is less severe for drug patent settlements. For one thing, repeat interactions are much less frequent, as explained above, although continued consolidation in the industry provides one ground for caution. Further consolidation, as it reduces the number of industry players, could increase the frequency of repeat play,
and hence the opportunity for informal arrangements. It would also raise the cost of overinclusion, by reducing the number of available alternative counterparties with which to conduct business arrangements for non-settlement purposes.

In addition, the negative consequence of curtailing brand-generic interactions—of tolerating an overinclusive ban on the content of side deals—is small. Eliminating continued entitlement to the exclusivity period, despite settlement, would also simplify consideration of any arrangement reached by the brand-name and generic firms. Thus, it seems unlikely that effective avoidance would survive the promulgation of a strong rule. If that judgment is incorrect, the agency’s ability to respond flexibly, without being subject to stare decisis, may prove to be a significant advantage.

Conclusion

Examining in detail the terms and effects of drug patent settlements reveals several important points. Drug patent settlements that restrict generic competition are an increasingly important, unresolved problem in antitrust enforcement. The evolution in settlement structure makes it less likely that courts will correctly identify and condemn them. There is therefore much reason to fear a continuation and intensification of false negatives if the current policy persists. Case-by-case evaluation is a failure and is likely to remain so, at least absent intervention by the Supreme Court. One partial response is to impose a presumption of payment where side deals accompany delayed entry. This would force firms to explain their increasingly questionable side deals, and would potentially discourage such complex dealmaking in the future.

The analysis presented here supports several further measures. This study
reveals persistent gaps in public knowledge about settlements, both in their existence and their terms. These are gaps that the FTC is uniquely positioned to fill. The FTC should step in to collate the extensive information it already has, supplement it with additional factfinding, and disseminate authoritative information of the type offered here. In that respect, this study represents a prima facie case that additional information gathering is necessary, and can serve as a first draft for the FTC’s future work. So long as settlements and their terms remain hidden, it will be difficult to do integrative work of the kind suggested here, and difficult to develop the “consensus among commentators” that is a key step in developing appropriate antitrust policy. The additional insight will help academics and policymakers in revising, if necessary, the initial conclusion presented here that pay-for-delay settlements are frequently tried and frequently successful.
Chapter 4: Patent Portfolios and Generic Drug Challenges

(Joint with Bhaven Sampat)

Introduction

The pharmaceutical industry is the rare setting in which the patent system works as it is supposed to—or so the story goes. Innovative new drugs have supplied dramatic improvements in longevity and quality of life over the past century (Murphy and Topel 2000; Lichtenberg 2007, 2009; Lichtenberg and Virabhak 2007). Effective patent protection has provided a critical stimulus to that lifesaving research. Patents are widely thought to have a unique role in stimulating drug research and development, compared to other industries (Levin et al. 1987; Cohen, Nelson and Walsh 2000). It is easy to see why. A patent on a new drug compound is difficult to invent around, and thus is effective in preventing imitation. Without a patent, imitation is straightforward, without any need to duplicate the innovator’s enormous efforts to discover and test a new drug.

The profitability of new drugs is limited—some would say threatened—by the aggressive entry of so-called generic drug makers, which offer a close, lower priced copy of the brand-name drug. Once generic firms enter the market, prices fall, often to less than 10 percent of the price of the brand-name drug. Generic drug utilization has seen explosive growth over the past 25 years, from 20 percent of prescriptions then to 70 percent today (Frank 2007; Engelberg et al. 2009). Generic drugs have saved consumers an estimated $700 billion between 1999 and 2008 (IMS Health 2009). Some of that savings is lost innovator profit, raising concerns that aggressive generic competition, by reducing innovator incentives, might also contribute to the “current crisis in industry R&D pipelines” (Higgins and Graham 2009, p. 370).
From the standpoint of innovators, the matter is dire enough when generic drug entry occurs after patent protection expires. But generic drug makers often seek Food and Drug Administration (FDA) approval and market entry prior to patent expiration, taking the view that one or more brand-name patents are invalid or not infringed by the generic product. Such “challenges,” often called “Paragraph IV challenges,” occur pursuant to a special legislative regime that governs the approval process for generic drugs, commonly known as the Hatch-Waxman Act. Frequently, these challenges result in patent litigation. In many cases, the generic firm wins the challenge, resulting in entry, and lower drug prices, much earlier than would otherwise be the case.

As we describe in detail below, the prevalence of challenges have risen dramatically over the past 25 years, placing them at the center of a vigorous debate about drug innovation and access (Engelberg 1999; FTC 2002; Grabowski 2004; Hemphill 2006). The concern that challenges reduce brand-name patent protection, and hence research incentives, is strongest if drug patents fit the ideal type ascribed to them by Levin et al. 1987 and others, in which a novel molecule is covered by single patent that clearly covers the molecule. In that case, a rise in generic challenges might indicate a rise in “prospecting,” in which generic drug makers challenge basic compound patents—or basic compound patents on high-sales drugs—indiscriminately, even though most challenges of this type are long-shots, in the hope of winning a few. And

---


9 Voet (2005) is typical: “[T]he validity of virtually all major patented drugs is being challenged not necessarily because they are not meritorious patents, but only because that is the road to riches. Thus major generic companies have scores of such suits ongoing and generic companies rely on the law of averages—if you place enough bets, you are sure to win a few of them . . . .”
if so, more and earlier challenges might imply less protection and resulting reward for innovative drugs.

As we show, however, many drug patents do not cover the compound. Brand-name firms have sought increasing recourse to ancillary patents on chemical variants, alternative formulations, methods of use, and relatively minor aspects of the drug. These patents have the effect of extending the total duration of protection for a brand-name drug, compared to the protection offered by the compound patent alone. The parallel rise of generic challenges and brand-name patenting suggests an alternative account, in which challenges provide a means for generic firms to restore the status quo, by challenging and thereby clearing away patents of questionable validity or relevance (Engelberg 1999; Bulow 2003). While the source of much debate, these conflicting accounts have been subject to little systematic empirical work.

In this paper, we examine the growth in generic challenges and contemporaneous increase in brand-name patenting, and the relationship between patent portfolio building and the likelihood of a challenge. We bring novel data to bear by examining a new dataset of brand-name drugs approved by the FDA. Among other advances, we construct detailed measures of the patent protection covering each drug and the timing of a challenge (if any).

We find evidence against the null hypothesis that generic drug makers are indiscriminate in their attacks, or that market size is the primary determinant of challenges. Market size (as proxied by sales) does matter, unsurprisingly, as does the technical difficulty of making the drug. Patents, however, matter too. Non-compound patents, particularly non-compound patents that extend the duration of protection
beyond the expiration of the patent, make challenges more likely. Our results provide support for the proposition that generic drug makers opportunistically challenge weaker patents that block entry.

This paper proceeds in six parts. Part I reviews previous studies of generic drug entry and patent policy that inform our inquiry. Part II describes the regime set up by the Hatch-Waxman Act and the incentives it creates, including a special reward available to certain generic drug makers that challenge brand-name patents. Part III offers new descriptive results that trace the growth of brand-name patent portfolios and the contemporaneous increase in generic challenges. Part IV describes our data and explains our empirical approach, which relates the likelihood of a challenge to a drug’s patent portfolio, sales, and other characteristics. Part V reports the results. Part VI assesses the implications of our results for several debates about drug and patent policy.

I. Previous Studies

Though there is a significant theoretical and empirical work on generic entry (Scott Morton 1996, Reiffen and Ward 2005), there is little previous work on Paragraph IV challenges.

A few recent papers have discussed patent challenges as part of broader work on the changing length of market exclusivity periods in pharmaceuticals (Grabowski and Kyle 2007; Grabowski 2004). There has also been some work on how sales affect challenges: Berndt et al. (2007b) considers the effect of sales on the intensity of challenge, conditional on a challenge occurring, for a proprietary sample of drugs.
Filson and Oweis (2010) use two events in the history of Paragraph IV challenges to evaluate the effect on decision making by pharmaceutical startups.

Our analysis is a large-sample study of the determinants of challenges, and the first to assess the relationship between brand-name patents and generic challenges or entry. By focusing on the role of patents, our project draws on a substantial literature, most of it theoretical, evaluating the partial protection provided by a patent. As economists understand this regulatory entitlement, a patent is “probabilistic,” in the sense that it provides not an absolute privilege to exclude alleged infringers, but only a right to try to exclude, through litigation whose outcome is uncertain ex ante (Lemley and Shapiro 2005).

This uncertainty is due partly to ambiguity in the breadth of patent claims, creating doubt as to whether they cover the product of an alleged infringer (Bessen and Meurer 2008). The uncertainty is also a natural consequence of the light scrutiny that patents receive during the application process, due in part to sharp resource constraints facing patent examiners (Jaffe and Lerner 2004) and to differences in strictness across patent examiners (Lemley and Sampat 2010). As a consequence, at the time of issuance it is uncertain whether a patent in fact reflects a nonobvious advance over the prior art, as is necessary for a patent to validly issue.

The question whether light ex ante review is good policy is a specific instance of the more general, longstanding inquiry by legal scholars and economists about the virtues of ex post review relative to its alternatives. The lightness of review ex ante

---

10 The choice between ex ante and ex post resolution of uncertainty about the validity and breadth of a patent is also, at its base, an inquiry about the merits of private litigation compared to alternative modalities of regulation as a means to determine the existence and scope of private rights.
might be a rational response given the substantial cost entailed in reviewing each patent. Such “rational ignorance” is cost-effective provided that most patents have little economic importance, and the set of important patents cannot be identified early (Lemley 2001), but can be identified later. One necessary condition for effective ex post review is that the reviewer must have adequate incentives to review. That condition might not be satisfied for patents generally. The reason is that the alleged infringer, the economic actor in the best position to challenge the validity and scope of a patent, faces a free-rider problem, because other alleged infringers can quickly take advantage of a favorable judgment. This result has led commentators to conclude that patent challenges are underprovided, both in the decision to bring a challenge and in the incentive to pursue it vigorously (Farrell and Merges 2004; Miller 2004; Thomas 2001). That free-rider problem is overcome in the pharmaceutical industry, due to a special incentive to challenge patents discussed below. A second condition is that the likelihood of intensive review should increase with the deadweight loss associated with exclusion.

Prior to that review, even weak patents can have important effects on competition. They can slow down rivals by obliging them to search for, evaluate, and litigate patents that are unlikely to be found valid and infringed. Moreover, patents do not always exist in isolation as single entities. In some industries, single firms collect extensive portfolios that they assert, or threaten to assert, against other firms (Hall and Ziedonis 2001). In general, portfolios can be expected to discourage entry, permitting the incumbent to exclude substitutes or extract revenue from producers of complements. Portfolio building has not generally been associated with the pharmaceutical industry,

Posner (2009) reviews the values at stake, emphasizing the relative strength of ex post litigation and litigation-like regulatory processes in making the most use of situation-specific facts.
which is typically understood as a “discrete product” industry in which a single patent covers a single product (Levin et al. 1987; Cohen et al. 2000). As we report below, however, brand-name drug makers are building patent portfolios, raising the question of what effect this might have on generic competition.

As discussed in the introduction, some commentators view challenges as a mode of ex post review that can identify and eliminate low-quality patents (Engelberg 1999). This is particularly important given concerns that the U.S. Patent and Trademark Office (PTO) and FDA are unable to do so (Eisenberg 2007), allegations about “evergreening” of patent portfolios by brand-name drug makers, and the consumer harm from such patents (Thomas 2005). On other hand, Grabowski and Kyle (2007), Grabowski (2004), and Higgins and Graham (2009) take the view these challenges are mostly driven by sales, and that they create uncertainty and reduce returns to R&D for brand-name firms. The null hypothesis in this strand of the literature—at least implicitly—is that patent characteristics don’t matter.

II. The Regulatory Regime for Pre-Expiration Challenges

Generic patent challenges target brand-name drugs that are already on the market. Under federal law, a brand-name firm must demonstrate that a new drug is safe and effective before the FDA will approve it for marketing. Making that demonstration as part of a New Drug Application (NDA) is a lengthy, expensive process, consuming years and many millions of dollars to conduct the necessary clinical trials (DiMasi et al. 2003).
Once the brand-name firm places a drug on the market, a generic firm may seek to market a competing, “therapeutically equivalent” version of the same drug by filing an Abbreviated New Drug Application, or ANDA, with the FDA. The generic firm must demonstrate that the rate and extent of absorption of the active ingredient (AI) are the same in both drugs.\footnote{21 U.S.C. § 355(j)(8)(B). The applicant must also demonstrate that the generic drug contains the same conditions of use, route of administration, dosage form, strength, and labeling. § 355(j)(2)(A).} The generic drug maker must also be able to actually make the drug, a task that is easier for some dosage forms of drugs (such as tablets or capsules) than others (such as, say, a transdermal patch). New clinical trials are not required. An ANDA costs about $1 million to prepare (FDA 2003a).

Most new drugs are protected by one or more patents. Those patents are listed by the brand-name firm in an FDA compendium commonly known as the Orange Book. Additional details about the Orange Book, and other regulatory matters, are contained in the Note on Data Collection. The generic firm, faced with this array of patents, may choose not to challenge any patents, in which case the FDA delays ANDA approval until expiration of the last listed patent.

In many cases, however, the generic firm attempts to enter prior to patent expiration. In that case, the generic firm files an ANDA containing a “Paragraph IV” certification, asserting that one or more listed patents are invalid or not infringed by the proposed generic product. The filing of such an ANDA is an act of patent infringement. In response to the ANDA, the brand-name firm may file a patent infringement suit to establish validity and infringement. This pattern—launch, challenge, sue—is frequent for major drugs, and it has become the norm for the best-selling drugs (Hemphill 2006,
Litigation raises the expense of a Paragraph IV challenge to $10 million or more (Goodman 2004).

A generic drug maker has a special incentive to challenge a patent, particularly if the patent is probably invalid or not infringed. That is due to a special feature of the Hatch-Waxman regime. Under certain circumstances, the first generic firm to file an ANDA is entitled, upon FDA approval, to a 180-day exclusive right to market its product in competition with the brand-name firm before other generic firms may enter. This exclusivity period provides a bounty to generic firms that incur the costs of Paragraph IV challenges.

There are two types of brand-name drug that are subject to Paragraph IV challenges. Some drugs, called “new chemical entities” (NCEs), contain a novel AI. NCEs are often thought to be the most innovative drugs (NIHCM 2002). NCEs receive special regulatory protection from challenges, in that the FDA may not accept an ANDA for filing during the first four years after approval of the brand-name drug maker’s NDA. Other drugs are essentially improved versions or variants that contain a previously approved AI. For example, the drug may be offered in a liquid dosage form rather than the original tablet, reformulated so that the drug can be taken just once a day, or combined with another existing drug. Reformulation as a strategy for extending drug life is a particular focus of brand-name drug makers (Perett 2008). For non-NCE drugs, an ANDA may be filed immediately after NDA approval.
III. The Rise in Brand-Name Patent Portfolios and Generic Patent Challenges

To assess changes in brand-name patenting and generic challenges over time, we constructed a new dataset that combines detailed information about brand-name drugs, including patent protection, with detailed data about Paragraph IV challenges for each drug. We start with the set of 1481 new non-injection brand-name drugs approved between 1985 and 2008, collected from an FDA database (FDA 2009). A drug is one or more AIs and a dosage form (e.g., extended-release tablet). We aggregate multiple strengths (e.g., 10 milligrams) of the same drug. For each drug, FDA data discloses the applicant name and approval date. We then match this information to data about patent protection and Paragraph IV challenges.

A. Patent Portfolios

For each drug, we collected information about applicable patent protection from current and archival editions of the Orange Book (FDA 1985-2009). The Orange Book is a comprehensive account of a drug’s patent protection, including any patent containing at least one claim that covers the drug’s AI, its formulation, or a “method of use” pertaining to an approved indication (e.g., inhibiting cholesterol biosynthesis). Removing those drugs that have no Orange Book patents, and hence are not subject to Paragraph IV challenges, yields a set of 1032 drugs.

Our first analyses examine the growth of portfolio building by brand-name firms. Our measure is the number of unique patents that are listed in (any edition of) the Orange Book. We collect the drugs into a series of six three-year approval cohorts starting in 1985, the first full year after the Hatch-Waxman Act was passed, and ending
in 2002. We stop with 2002, because later cohorts are censored: some patents are added to the Orange Book years after the drug is approved. 692 new drugs were approved during this 18-year period.

Figure 5 shows trends over time in the number of patents per drug. Drugs in the first cohort, approved between 1985 and 1987, have an average of 1.9 patents per drug. In the final (2000 to 2002) cohort, the mean slightly more than doubles to 3.9 patents per drug. The median increases from 1.5 to 2.5 patents per drug. Much of the growth in patenting is attributable to the right tail. The top quartile point grows steadily, from two patents per drug to five patents per drug. In other words, the top twenty-five percent of patent portfolios, among drug approvals in the first several years of the Act, had two or more patents per drug, while the top portfolios fifteen years later were more than double that size.

Using Orange Book data, we also determine whether the drug received protection as an NCE. As Figure 5 shows, the increase in patenting plays out differently for NCE and non-NCE drugs. For the first three cohorts, NCEs have more patents on average. Starting with the 1994-1996 cohort, non-NCE drugs take the lead. In the top quartile, non-NCE growth is particularly marked. Non-NCE and NCE drugs start at the same place in the first cohort, 2 patents per drug, but non-NCEs triple by the final cohort, to 6 patents per drug, while NCEs merely double.

Not all patents are created equal. Some patents are more likely to exclude generic entry than others. There is a rough hierarchy in patent strength—that is, in the likelihood that a brand-name firm will convince the court that its patent is valid and infringed by the drug proposed to be made in the generic firm’s ANDA. Patents that
claim the AI—those basic patents that cover the drug compound—are generally the strongest. They are infringed by making the generic drug product, almost by definition; otherwise bioequivalence is lacking. To make an invalidity argument, a generic drug maker is left contending that the drug was previously disclosed or that the patentee engaged in inequitable conduct during the application process. These are difficult arguments to win.

In comparison to AI patents, patents for particular formulations—for example, a chemical mechanism providing sustained release of the drug substance over time—are more open to attack. In that case, the generic drug maker can argue not only invalidity but also noninfringement. For example, the generic firm can argue, often with success, that it employs a different, noninfringing mechanism for accomplishing the sustained release of the drug. Other patents listed in the Orange Book—for particular salt forms, particle sizes, and methods of use—are also open to challenge.

To explore the importance of portfolio building, and the respective roles of AI and non-AI patents, we examined the claims of each Orange Book-listed patent, 3036 patents in all. One of the authors developed a coding guide and worked with a former PTO examiner of drug patent applications to code each patent according to whether it contained at least one AI claim. This coding is arguably overinclusive: as elaborated in the Note, certain types of claims are counted as AI claims even though they are weaker than the “classic” compound claim in significant respects. Thus, there are likely more non-AI patents than our counts suggest. In recent years, the use of non-AI patents has increased substantially, and may account for the increase in nominal patent term. Figure 6 traces out the trend. In the first cohort of NCEs, 61 percent have at least one non-AI
patent listed in the Orange Book. By the time of the 2000 to 2002 cohort, 84 percent have at least one such patent. Non-NCEs have seen similar growth. As discussed in Part V, these patents have a significant effect upon the likelihood of a Paragraph IV challenge.

Patents pertaining to a single drug differ along a second dimension, their expiration date. Those that expire later potentially provide a substantial temporal extension in a brand-name drug maker’s exclusivity. We call the lag between a drug’s approval date and the date of its last-expiring patent the “nominal” patent life for that drug. Brand-name firms and other market participants often use this date in their announcements and discussions of when a drug goes off-patent. Figure 7 shows the increase in nominal patent life over time, again grouped by approval cohort. NCE drugs have seen a more than three-year increase in nominal patent life, from 12.2 years to 15.5 years. Non-NCEs have seen a smaller increase, from 14.8 years to 16.6 years.

Non-AI patents that expire later than the compound patent on a drug seem particularly likely to attract Paragraph IV challenges. The generic firm has an additional incentive to selectively challenge the later-expiring patent, even if it concedes infringement and validity of the compound patent, in order to secure entry upon expiration of the compound patent. Moreover, the generic firm need not challenge the compound patent in order to enjoy the 180-day exclusivity period.

Despite the risk of a challenge, including a later-expiring patent within the patent portfolio has significant benefits for the brand-name firm. Most obviously, it potentially extends the total duration of protection, beyond the expiration date of the first patent. Moreover, if the patent attracts a selective challenge to the later-expiring
patent only, that is an unambiguously better outcome for the brand-name firm, compared to the patent never having issued. At worst, the brand-name firm loses the challenge, and entry occurs when it would have anyway. And even if the challenge is successful, the award of 180 days itself provides a measure of protection to the brand-name firm, since it restricts the extent of generic entry for half a year. At best, the brand-name firm might win the challenge or discourage entry altogether.12

B. Paragraph IV Challenges

During the same period that patent portfolios have increased, Paragraph IV challenges have grown as well. To explore this trend, we also collected detailed information about Paragraph IV challenges. We determined which drugs attracted challenges between 1984 and December 2009 using a list of such drugs, which we call the “Paragraph IV List,” maintained by the FDA (FDA 2010). Comparing the Paragraph IV List to the set of approved brand-name drugs reveals that 299 drugs (out of 692) have been subjected to Paragraph IV challenge by August 2009. Overall, the fraction of drugs challenged has increased from 22 percent of drugs approved in the first cohort to 55 percent in the last cohort. We report these trends in Figure 8. Overall, a slightly larger share of NCEs (44 percent) than non-NCEs (43 percent) have been subjected to Paragraph IV challenges. The most striking change, however, has been the rise in challenges against non-NCEs, from just 15 percent of non-NCEs approved in the first cohort, to more than half (58 percent) of those approved in the last cohort.

12 Including the late-expiring patent might create a risk for the brand-name firm. The generic firm might decide that as long as it is challenging the second patent, it might as well challenge the first patent, too, given that it is already incurring the cost of a challenge and litigation. This spillover effect could be a significant negative consequence of the accumulation of late-expiring patents.
To understand the role of Paragraph IV challenges in reducing effective brand-name patent life, we need to know when these challenges occur relative to FDA approval. The FDA’s Paragraph IV List reports the date of first challenges that occur after March 2004. The lack of systematic data before this period has hindered empirical analyses of challenges. To relax this data constraint, we extended this data back to 2000 by hand, using a variety of public sources. (See Note for details.)

One benefit of information on timing of challenges is that we can construct a measure of “early” challenges, within one year of challenge eligibility, for a range of drugs. For this analysis, the right truncation is less severe, and so we can use a wider range of approval years. For NCEs, we can construct this measure for drugs approved between 1996 and 2004. (An NCE approved in 1996 can be challenged starting in 2000, on account of the regulatory delay described above. An NCE approved in 2004 can be challenged starting in 2008.) For non-NCEs, which can be challenged immediately, we can construct this measure for drugs approved between 2000 and 2008. Similarly, we can construct a three-year challenge measure for NCEs approved between 1994 and 2002, and for non-NCEs approved between 1998 and 2006.

Figure 9 reports the results. NCEs have seen a dramatic increase in year-one challenges over the past decade, from seven percent of drugs approved in 1996 (and hence challengeable starting in 2000) to 69 percent of drugs approved in 2004 (and hence challengeable starting in 2008). Non-NCEs are seldom challenged in year one of eligibility. That is not surprising, because it takes time to design a viable challenge, and it may not be immediately clear that a drug is worth challenging. The three-year measure is more appropriate for non-NCEs. On this measure, non-NCEs have seen
significant growth from 19 percent of drugs approved in 1998 to 33 percent of drugs approved in 2006.

Patenting and challenges have thus both increased sharply in the quarter century since the passage of Hatch-Waxman. In the next section, we explore how the two are related in the cross-section, examining how patents and other drug characteristics affect the likelihood of challenge, for all drugs eligible for challenge over the past decade.

IV. Data and Empirical Approach

A. Data: Description and Summary Statistics

Our sample consists of prescription drugs that first became eligible for generic challenge between 2000 and 2008. That range reflects the fact that we know the timing of challenges that occur between 2000 and 2009, and allow for at least one full year of observation. For non-NCE drugs that can be challenged immediately upon approval, these are drugs approved between 2000 and 2008. For NCEs that become eligible after four years of exclusivity, these are drugs approved between 1996 and 2004. (In addition, a few drugs have an NCE exclusivity period of shorter duration, as explained in the Note.)

We obtained sales data from the National Sales Perspective database of IMS Health, the leading provider of sales data. Specifically, we collected monthly sales for January for each year between 2001 and 2009. For convenience, we convert all sales figures to annualized sales in billions of dollars, inflated to 2010 values using the CPI
We omit a small number of drugs for which IMS Health lacked sales. The resulting sample contains 479 drugs.

Half of the 479 drugs were challenged by the end of 2009. Among drugs eligible for challenge in 2006 or earlier, for which we at least three years of observations, 35 percent were challenged in the first three years.

What explains which drugs are challenged? Table 7 describes each of the variables we examine. Our main analyses are duration models, following a drug over time. However, for expository convenience, we begin in Table 8 by presenting static descriptive information, providing summary information for each of the 479 drugs. The drugs were approved between 1990 and 2008. Most of the approvals before 2000 (35 percent of the sample) are for drugs with NCE protection. In addition, the dataset includes a few previously approved antibiotic drugs (8 percent of the sample), which were made eligible for challenge in 2008, under a complex scheme discussed in the Note. As discussed above, dosage forms differ in their ease of manufacture and replication, and so may also be important to the likelihood of challenge. Two-thirds of the drugs in our sample are orally administered. We also calculated average sales for each drug (across each year we observe the drug). These range from zero to $5.9 billion, with a mean of $205 million. Nine drugs in our sample have average annual sales exceeding $2 billion (Zyprexa, Lipitor, Protonix, Celebrex, Advair Diskus, Nexium, Cymbalta, Abilify, and Vytorin).

We employ three types of measures of the quality of the patents on a drug. The first set is based on information on the type of patents on a drug—whether the drug has only AI patents, only non-AI patents, or both types—using the expert coding described

---

13 This data is available at http://www.bls.gov/data/inflation_calculator.htm.
above. (Recall that, by definition, each of the drugs in our sample has at least one patent. Otherwise it would not be eligible for a challenge.) Of the drugs in our sample, 12 percent have only AI patents, 47 percent only non-AI patents, and 41 percent have both. To examine the hypothesis that late expiring non-AI patents could have distinctive effects, we also determined whether the non-AI patent adds market life beyond the AI patent, or is redundant to it. For a drug with both AI and non-AI patents, about three-quarters of the non-AI patents expire after the AI patents. Thus about 8 percent of the drugs in our sample have redundant non-AI patents, and 33 percent have non-AI patents generating extra market life. On average (across all drugs) the extra life from non-AI patents is 1.9 years.

In addition to the measures based on our own coding, we also employ a suite of widely used measures of patent quality, including the number of times a patent is cited (by later-issued patents) and the number of countries in which patent protection is sought (Hegde and Sampat 2009; Sampat 2010). Forward citations to a patent are commonly used measures of patent value, on the theory that higher quality patents generate more citations. Using information from a private patent data vendor, Delphion, we determined the number of citations for all Orange Book listed patents issued until the end of 2008. For each patent associated with a drug in our sample, we determined whether it is in the bottom quartile of citations for all Orange Book listed patents with the same issue year. (Normalizing by issue year is necessary because more recently issued patents have a shorter window to be cited.) For each drug, we created a variable counting the number of these “low quality” patents. On average, drugs in our sample had 0.98 of such patents. We constructed a similar measure for family size, the number
of countries where a patent application was filed. This too is a commonly used measure of patent quality, on the theory that more important patents will be filed in more jurisdictions (Lanjouw et al. 1998; Lanjouw and Schankerman 2004). On average, the drugs in our sample have 0.79 patents in the bottom quartile of this measure of patent quality.

A last measure is the number of patents applied for before or during the calendar year of drug approval—what we call “early” patents—and the number of “late” patents applied for after approval. Late filed patents are generally viewed as being lower quality than others, central to the alleged “evergreening” strategies discussed earlier. Drugs in our sample have 3.3 early patents, on average, and just 0.36 late-filed patents.

**Empirical Framework**

Our estimation strategy relates the likelihood of a first challenge to a drug at a point in time to the extent and type of patent protection, drug sales, and other drug characteristics. For each of the drugs we have an observation for each year following its initial eligibility for challenge, until it either receives its first challenge or is censored when our data ends on the last day of 2009. We specify the probability of challenge for drug $i$ at time $t$, $P_{it}$, as

$$P_{it} = \alpha + \beta SALES_{it} + \gamma PATENTQUAL_{it} + \delta_i + \lambda(t) + \theta_i.$$ 

$SALES$ is a set of time-varying sales measures for each drug (e.g., sales in the past year). $PATENTQUAL$ includes the patent measures discussed above. $\delta$ is a set of measures of drug characteristics, including chemical type, therapeutic class dummies, and dosage form. The last two terms capture the timing of the challenge. $\lambda$ is a measure of duration.
dependence, years since eligibility, which is operationalized as a set of dummy variables. \( \theta \) is a set of challenge year dummies. We are thus estimating discrete-time duration models of the likelihood of challenge.\(^\text{14}\) These models will be estimated using logistic regressions, with standard errors clustered on drugs.

The dependent variable is whether the drug receives its first challenge in a given year, e.g. in the first year after eligibility, the second year, and so forth. Each of the 479 drugs in our sample has an observation for each year after eligibility, until it is either challenged or censored. We also drop a small number of observations (60 observations) in which no challenge has yet occurred, but all patents have expired. At that point, the drug is not at risk of challenge. The resulting dataset has 1807 drug-year observations.

Table 9 describes the data structure. In year one after eligibility, all 479 drugs are at risk of challenge. Of these, 77 drugs, or 16 percent, receive a challenge. Of the remaining 402 drugs, 40 are censored: these are drugs that were eligible in 2008, and 2009 is the last observed year of possible challenge. Five more drugs leave the dataset because their patents expire. That leaves 357 drugs at risk of challenge in year 2. Of these, 43 drugs (12 percent) were challenged and 22 (6 percent) were censored. And so on. The first column can thus be interpreted as a discrete time hazard rate, without conditioning on additional predictors.

V. Results

Table 10 shows results from the baseline models. In the most parsimonious specification, Model 1, we relate whether a drug is challenged to indicators for sales of sales.
the drug in that year, time elapsed since eligibility, indicators for each potential challenge year, and therapeutic class effects. Sales have a large, positive, and statistically significant impact on the hazard of challenge. A billion-dollar increase in annual brand-name sales is associated with an 11 percentage point increase in the (annual) probability of challenge. Neither the NCE indicator nor the control for antibiotics is significant. Consistent with our expectations, oral drugs have a much higher hazard of challenge than others, raising the probability of challenge by 9 percentage points.

Model 2 introduces the patent characteristics. Having only non-AI patents has a positive and statistically significant effect on the likelihood of challenge, relative to the left out category (AI patents only). Drugs with both types of patents also have a higher likelihood of challenge. Having at least one non-AI patent in addition to having at least one AI patent raises the probability of challenge by 9 percentage points, compared to having no non-AI patents.

Model 3 unpacks the “both” category into cases where non-AI patents are redundant to the AI patent (i.e., adds no extra time), and cases where they extend nominal patent life. Non-AI patents generating extra market life have a positive and significant effect on likelihood of challenge, as does a redundant non-AI patent.\(^{15}\) Model 4 shows that the amount of extra time generated by non-AI patents has a positive and significant effect. Relative to non-AI patents generating no extra patent life, those with the most extra life (in the top category, about 2.6 years) extra life have a 25 percentage point higher hazard of challenge. Model 5 employs non-parametric

\(^{15}\) The difference between the coefficients is not statistically significant (p=0.67).
indicators of the amount of extra time generated by non-AI patents, and shows that the likelihood of challenges increases monotonically with this measure.

These results suggest that sales have a strong and positive effect on the likelihood of challenge. But conditional on sales, challenges are also responsive to the presence of weak patents, particularly those generating considerable extra market life.\textsuperscript{16}

Though the tables suppress eligibility year indicators, these variables do provide insights on challenge dynamics across types of drugs. To illustrate this, we estimated model 2 separately for NCEs and non-NCEs. Figure 10 plots the predicted values from these models (at the means of all other variables). For NCEs, the hazard of challenge shows a sharp decrease after year one, and is generally decreasing over time. By contrast, for non-NCEs the hazard remains more or less constant. This may reflect differential learning dynamics. For NCEs, potential challenges have the exclusivity period (in most cases, four years) to observe the drug, during which time they can determine feasibility of reverse engineering and manufacturing, and about the market for the drug. Accordingly, for these drugs most of the most profitable challenges will occur early after eligibility. For non-NCEs, where there is no exclusivity period, this learning occurs throughout the eligibility period.

To illustrate the magnitude of the impact of non-AI patents, we also estimated a variant of model 2 where we collapsed the patent information to two categories, indicating drugs only an AI patent (the left out category) and drugs with any non-AI patent (alone or together with an AI patent). Figure 11 plots the predicted hazards from

\textsuperscript{16} These models constrain the effect of sales to be constant across NCEs and non-NCEs. However, these sales figures are taken at different vintages for NCEs, since they are not eligible for challenge until the NCE exclusivity period elapses. Results on the patent variables are very similar if we allow NCE and non-NCE sales to have different coefficients.
these models over time, separately for drugs with AI patents and for drugs with non-AI patents. Across all drugs, the hazard is generally decreasing over time. But (as would be expected from the regression results) the hazard is much larger for drugs with non-AI patents.

Table 11 shows results from the other patent quality measures, including citations, family size, and the number of early versus late patents. Models 6 and 7 show that drugs with more patents in the first quartile of citations, or the first quartile of the family size distribution, are associated with significantly higher rates of challenge. Model 8 shows that the number of early patents has a negative and statistically insignificant association with the likelihood of challenge, while the number of late patents has a strong, positive, and statistically significant effect. For example, drugs with one late applied patent have a 2.5 percentage point higher likelihood of challenge than drugs with none, evaluated at the means of all other variables.

In the specifications above, sales are introduced continuously. We relax this assumption in the models reported in Table 12, by including indicators for sales quartiles (with the first quartile left out). The sales quartiles are calculated relative to all other drugs sold at a point in time. Model 9 suggests sales do have a non-linear effect on challenges. Thus the predicted hazard of challenge is 1 percent for the bottom quartile of sales, 6 percent for the second, 16 percent for the third, and 28 percent for the top quartile. In addition, allowing sales to affect challenges via this more flexible functional form does not affect the patent results discussed above.
Additional Analyses

These results suggest that while sales matter for challenges, so do patents. In particular, challenges seem to be driven by low-quality patents. In this section, we discuss various threats to identification, and our attempts to rule them out.

One concern is that our results reflect omitted variable bias: the expected profitability of the drug, not captured by drug-specific sales or the time invariant therapeutic class indicators, affects both the extent of patenting and the likelihood of a challenge. To examine this, we also estimated our baseline models including the three year distributed lag of sales, in addition to contemporaneous sales. (Doing so required beginning the analyses in 2003, since sales data are not available before January 2000. This change results in a reduction of sample size of 183 drug-years, to 1624.) This approach aims to control for trends in sales, which are arguably more predictive of future profits than point-in-time sales figures.

Table 13 shows the results. In each of the models, the coefficients on sales and lagged sales are haphazard, reflecting collinearity between the measures. (Wald tests show that the sales measures are jointly significant all each of the models.) More importantly, the direction, magnitude, and significance of estimates on all of the patent variables are unchanged.

As another proxy for expectation of sales, we also re-estimated the baseline models, adding a measure of expected market size. Specifically, we use total sales for all other drugs in a therapy class, three years forward, to proxy for expected sales. This measure reflects a “rational expectations” assumption, that firms anticipate the future size of the market (and thus expected profits), and make decisions to challenge and file
patents in response. In constructing this measure for market size, we focus on all other drugs in the market, since the size of the market for a given drug would be directly affected by a successful challenge (for example, if generic entry means wider utilization), which would introduce a different source of reverse causality to our analysis.

Table 14 shows the results. Since we use sales data in a class three years after a given observation year, we had to drop the 636 observations from 2008 and 2009 from these analyses, leaving 1171 observations. In each of the models, future sales (conditional on current sales) has a positive but statistically insignificant impact on the likelihood of challenge. This could reflect that the therapeutic class indicators were adequate controls for future sales, that future sales don’t matter for challenges. It could also be that this proxy for future market size is not a good one. Subject to this caveat, our main patent results are robust to introduction of this control.17

VI. Discussion and Conclusion

In the quarter century since the passage of the Hatch-Waxman Act, the practice of listing questionable patents on the Orange Book has grown rapidly. There has been a concomitant increase in Paragraph IV challenges. The interplay of these two trends—and the patterns of litigation, settlement, and entry that result—determine the effective patent life for new drugs. In effect, the policies and rules governing these activities

17 The analyses above focused on omitted variable bias—specifically, that profit expectations drive both challenges and patenting decisions. A related threat to validity is that we have it backwards. Rather than patenting practices causing challenges, the anticipation of a challenge leads to increased patenting. As a check on this hypothesis, in unreported analyses we focused on orally administered drugs, which are easier to imitate, and thus more vulnerable to challenge. We found no evidence that these drugs acquired more non-AI patents or that they were more likely to late patents. These results are available on request.
determine how we balance incentives for dynamic efficiency (research and development incentives) and static efficiency (price competition) in pharmaceuticals.

We provide the first comprehensive evidence that allegations of “evergreening” are real, and that evergreening has grown over time, using data on all (non-injectable) drugs approved since the Hatch-Waxman Act was passed, and all Orange Book patents on those drugs. We find that nominal patent life has increased. How do we square this result with previous work, including the finding of Grabowski and Kyle (2007) that “effective patent life” has decreased over the past decade? One possible explanation is a difference in empirical approach. We examine drugs on a forward-looking basis, while previous work on effective life is conditional on generic entry, and hence backward looking. A second explanation is that thanks to Paragraph IV challenges, nominal patent life is cut back in effect to the baseline set by the expiration of the compound patent.

Intriguingly, effective patent life might be reduced even below that point, if a generic drug maker decided to challenge the compound patent too—for example, if a compound patent challenge were viewed as a low-cost add-on to a challenge on a later-expiring non-AI patent that would occur in any event. This line of thinking—that increases over the past decades in nominal patent life generated by evergreening have been reined in by Paragraph IV challenges—is the subject of ongoing work.

In our econometric analyses in this paper, we focus on all drugs first eligible for challenge between 2000 and 2008. Across these drugs, there is a strong association between patent portfolios and generic challenges. One striking result from our regressions is the importance of weak patents for the likelihood of a Paragraph IV challenge. Non-AI patents have a strong positive relationship with challenges,
especially those generating extra market life. Conditional on sales, drugs with more questionable patents, a larger number of late patents, and greater patent life generated by late patents are much more likely to draw challenges. While these results are based on our own characterization of patents, drugs with “low quality” patents using standard patent indicators—citations and family size—are also more likely to be challenged. These findings suggest that the characterization of challenges as frivolous attacks that reduce patent life (and perhaps, as a result, research and development incentives) is too simple. We strongly reject the null hypothesis that the composition and quality of a drug’s patent portfolio don’t matter.

We thus find qualified support for the proposition advanced in Bulow (2004) that the major contribution of generic drug makers is to make the patent system work more effectively. Innovations that support strong patents should be rewarded appropriately, while patents that do not cover a significant innovation should not. The likelier it is that weak patents will be weeded out, the lower the cost of granting a long patent reward in the first place. Given that pharmaceuticals is the rare industry in which patents are a direct and central impediment to competition, this weeding-out process is particularly important.

The role of large sales in encouraging Paragraph IV challenges also provides qualified support for the idea that ex-post review is an effective way to test probabilistic patents. As discussed in Part I, such “rational ignorance” is cost-effective provided that most patents have little economic importance, and the set of important patents cannot be identified early on. Instead, we should focus evaluation resources on those few patents that turn out to be valuable. The Paragraph IV challenge process is perhaps the most
vigorous example of that ex post analysis, subjecting patent protection for economically important drugs to intensive ex post scrutiny. It also suggests that incentives to challenge might have a useful role to play in other industries too.

Whether this scrutiny is effective, however, depends upon what happens after the challenge is filed. In ongoing work, we are examining the dynamics of litigation, settlement, and entry that follow these challenges. If these processes resulted in low-cost and rapid invalidation of dubious patents (and early generic entry on the associated drugs), this might suggest the Hatch-Waxman regime is working as intended, and ex post testing is making up for “rational ignorance” ex ante. There is evidence, moreover, that some drug makers “game” the post-challenge process (Bulow 2004; Hemphill 2006, 2009). If we find large-sample evidence that this behavior is widespread, the assessment is quite different. Accordingly, while the results reported in this paper provide new insights into the determinants of challenges, more work is needed before we can make strong claims about their welfare impact.

There is a broader question about whether, absent the incremental market life provided by dubious patents, brand-name firms would have sufficient incentives to invest in socially valuable research. Answering that question is beyond the scope of this paper. Instead, our normative baseline is that patents that do not meet the PTO’s patentability standards should not be used to exclude generic competitors. We are open to the argument, however, that patentability standards are a poor fit for the incentives needed to generate valuable innovation in pharmaceuticals (Eisenberg 2007; Roin 2009). We are skeptical, however, that the ability to assert dubious patents—or the equilibrium implied by the listing, challenge, and litigation process described in this
paper—is a move towards the first-best outcome. Here again, more information on which patents are challenged, and how the outcomes of challenges vary with this, may help us say more about this question.

Note on Data Collection

This note contains additional details about our data collection and related institutional details of the Hatch-Waxman Act.

Drug approvals. We collected data for every non-injection prescription drug approved by the FDA. We limited our data to prescription drugs, following FTC 2009, on the view that a generic drug maker’s decision to offer an over-the-counter drug is quite different from prescription drugs. Drugs that were initially approved as prescription drugs, and only later had a further OTC approval, are included. We also omitted injection drugs, again following FTC 2009, as these have a distinctive method of administration and sales (i.e. hospitals).

Patents listed in the Orange Book: We collected patent data for each brand-name drug from the FDA’s compendium of Approved Drug Products with Therapeutic Equivalence Evaluations, commonly known as the Orange Book. The Orange Book also lists, for each brand-name drug, any unexpired regulatory exclusivity and approved therapeutically equivalent generic drugs. For early editions, we collected the data by hand, augmented by the results of a FOIA request to the FDA. For data from 2000 to the present, we rely on archival electronic versions of the Orange Book.

The Orange Book contains a comprehensive, though not perfectly exhaustive, account of patent protection relevant to each drug. For patents issued before NDA
approval, a brand-name drug maker is required to list any patent containing at least one claim that covers the drug’s AI, its formulation, or any method of use pertaining to an approved indication. For patents issued after NDA approval, listing is not required, but there is a strong incentive to do so. If the patent is not listed, the generic firm filing an ANDA need not certify that the patent is invalid or not infringed as a condition for FDA approval. Nor can the unlisted patent provide a basis for the automatic stay of approval. The drug maker is prohibited from including other types of patents in the Orange Book, such as methods for manufacturing the drug. Some brand-name drug makers, however, tend to err on the side of inclusion. Brand-name drug makers are free to assert unlisted patents against generic drug makers, but these instances are rare.

Our coding strategy separates patents into two categories, AI and non-AI. The AI category includes both basic compound patents and other patents that are not so basic. For example, some AI patents claim a variant of an existing molecule, permitting the generic firm to argue that the new drug is obvious and hence unpatentable in light of the prior art. In other instances, the basic AI patent may have expired already, and the relevant AI patent may only cover a particular form of the drug, such as a particular crystalline structure. In that case, the generic entrant might try to market a distinct, noninfringing form that is nevertheless bioequivalent to the brand-name drug, or, alternatively, argue that the patent is invalid. Finally, the coding sweeps in a few patents

---

18 An example is enantiomers, which are compounds that are non-superimposable mirror images of one another, like left and right hands. Sometimes, a drug will be discovered and marketed first as a mixture of multiple enantiomers. But only one enantiomer is really doing the therapeutic work, and later it may be purified and marketed separately as a new drug. The question is whether the single enantiomer can receive separate protection, given the earlier disclosure of the mixture and knowledge about how to accomplish the separation and purification of a single enantiomer.
claiming aspects of the drug, such as particular salt forms of previously patented substances, that are not AI claims at all.

**ANDAs with Paragraph IV certifications:** We limit our examination to ANDAs containing a Paragraph IV certification. There are three alternative certifications, in which the generic drug maker asserts that there are no Orange Book-listed patents, that any such patents have expired, or that the generic drug maker is content to wait until the expiration of the last listed patent. See 21 U.S.C. § 355(j)(2)(A)(vii)(I)-(IV). Only Paragraph IV certifications provide for generic entry prior to patent expiration, or potential eligibility for the 180-day exclusivity period. The exclusivity period is available only to the first ANDA filer, though, as noted in the text, if multiple ANDAs are filed on the same day, they share in the exclusivity.\(^{19}\)

**Determining the date of first Paragraph IV challenge.** We determine the date of first Paragraph IV challenge for all drugs that become eligible for challenge in January 2000 or later. The FDA reports the date of first challenge only for first challenges starting in March 2004; before that date, the FDA reports only that at least one challenge occurred. We have filled in the gaps for drugs that were first challenged between January 2000 and March 2004. For challenges between August 2000 and March 2004, we determined the date by comparing archived versions of the Paragraph IV List, and augmented these results with reports about new challenges written by equity analysts at financial firms. For challenges between January 2000 and August

\(^{19}\) The FDA reached this conclusion in 2003 (FDA 2003b), and this view was codified by statute later that year, § 355(j)(5)(B)(vi)(I) (sharing exclusivity among first applicants); § 355(j)(5)(B)(iv)(II)(bb) (defining first applicant). There is a second route to shared exclusivity. In some cases, the filing of an additional patent in the Orange Book gave generic firms a fresh opportunity to share in the exclusivity period. Under “patent-by-patent” exclusivity, multiple generic firms, each first to a file a Paragraph IV certification for a different patent, could potentially share in the exclusivity. This interpretation, which applies to a substantial number of drugs, was ended by a statutory change in December 2003.
2000, we inferred that the challenge occurred during that period from the facts that the drug became eligible for challenge in 2000 or later, and was not challenged during the range for which we have more detailed data (January 2000 to present). A similar approach allows us to determine whether a drug was challenged within the one-year window or three-year window, as described in Part III.

_Determining the date of challenge eligibility._ In our regression analyses, we examine drugs that first became eligible for challenge between 2000 and 2008. A key distinction is between NCEs and non-NCEs. NCEs are drugs for which no “active moiety”—roughly speaking, an active ingredient—has been approved by the FDA. NCEs receive special regulatory protection under the Hatch-Waxman Act: no ANDA containing a Paragraph IV certification may be filed during the first four years after drug approval. (The relevant data exclusivity is sometimes referred to as “five-year exclusivity,” but the period is shorter in the case of a Paragraph IV certification. 21 U.S.C. § 355(j)(5)(F)(ii).) For example, Razadyne tablets were first approved by the FDA in February 2001. This was the first approval for the drug’s active ingredient, galantamine hydrobromide, so Razadyne tablets are first available for challenge in February 2005. Thus, our dataset includes NCEs approved between 1996 and 2004. An NCE approved in 1996 is eligible for challenge in 2000. An NCE approved in 2005 or later is not eligible for challenge until 2009 or later, and is omitted from our data.\(^\text{20}\)

\(^{20}\) One aspect of NCE exclusivity is not directly reflected in our analysis. If a generic firm files an ANDA with a Paragraph IV certification, and the brand-name firm files suit within 45 days of receiving notice of the certification, an automatic statutory stay blocks FDA approval for 30 months, measured from the receipt of notice, or a district court ruling in the case, whichever comes first. In addition, if the drug is an NCE, the stay may be lengthened so that it expires seven-and-a-half years after NDA approval. § 355(j)(5)(F)(ii). This constraint, however, almost never binds.
Most non-NCE drugs are eligible for challenge as soon as the drug is approved.\footnote{For certain non-NCE drugs, an ANDA cannot be approved for three years after NDA approval, but this restriction does not limit the filing of an ANDA.} An exception is drugs that contain an active ingredient giving rise to NCE protection, which also receive whatever is left of the four years of protection. For example, an extended-release formulation of galantamine hydrobromide, Razadyne ER, was approved in December 2004. That drug, like Razadyne tablets, became eligible for challenge starting in February 2005.

Delayed eligibility is also the result, under a highly unusual formula, for certain antibiotics. Prior to 1997, the FDA approved antibiotics under a special regime that did include patent listing or Paragraph IV challenges. A statutory change in that year subjected antibiotics to the regular procedure on a going forward basis, but exempted “old antibiotics,” that is, drugs containing an active ingredient that had been applied for under the old regime. The exemption extended to new, later-approved drugs with old antibiotic ingredients. In October 2008, these old antibiotics were subjected to the Orange Book listing requirements and made subject to Paragraph IV challenges (FDA 2008). As a result, 35 antibiotics had listed patents, of which four attracted a Paragraph IV challenge.
Figures and Tables

Figure 1: Share of Drugs with Pre-Expiration Challenge by Sales Quintile

![Bar chart showing share of drugs with pre-expiration challenge by sales quintile.](image)

NMEs approved between 1996 and 2002: challenges within three years.

Figure 2: Average Relative Price of Prozac Before and After Generic Entry

![Line chart showing average relative price of Prozac before and after generic entry.](image)
Figure 3: Approval Timeline for a New Drug

Figure 4: The Effect of a Guaranteed Bounty

Litigation

<table>
<thead>
<tr>
<th>G: 125</th>
<th>C: 4750</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: 5125</td>
<td></td>
</tr>
</tbody>
</table>

Settlement: Minimum Accepted by Incumbent

<table>
<thead>
<tr>
<th>I: 5125</th>
<th>G: 250</th>
<th>C: 4625</th>
</tr>
</thead>
</table>

Settlement: Result of Bargaining

<table>
<thead>
<tr>
<th>I: 9250</th>
<th>G: 250</th>
<th>C: 500</th>
</tr>
</thead>
</table>
Figure 5: Patents Per Drug

- Patents
  - Mean
  - 75th percentile
  - Median

- NCEs
  - Mean
  - 75th percentile
  - Median

- Non-NCEs
  - Mean
  - 75th percentile
  - Median
Figure 6: Share of Drugs with Non-AI Patent

![Graph showing the share of drugs with non-AI patent across different approval cohorts from 1985-1987 to 2000-2002.]
Figure 7: Years of Nominal Patent Life

![Graph showing the mean and median years of nominal patent life for different approval cohorts over time.](image-url)
Figure 8: Share of Drugs with Paragraph IV Challenge

The graph shows the share of drugs with Paragraph IV challenge over different approval cohorts from 1985-1987 to 2000-2002. The share is categorized into All Drugs, NCEs, and Non-NCEs. The share increases over time for all categories, with NCEs and Non-NCEs showing a more consistent trend compared to All Drugs.
Figure 9: Share of Drugs with Early Paragraph IV Challenge

**NCEs**

- 1 year (blue line)
- 3 years (red line)

**Non-NCEs**

- 1 year (blue line)
- 3 years (red line)

Approval Year: 1994 to 2004 for NCEs, 1998 to 2008 for Non-NCEs.
Figure 10: Predicted Values of Hazard of Challenge by Drug Type

Figure 11: Predicted Values of Hazard of Challenge by Patent Portfolio
Table 1: Antitrust Challenges to Pay-for-Delay Settlements

<table>
<thead>
<tr>
<th>Year</th>
<th>Drug</th>
<th>Plaintiff</th>
<th>Outcome</th>
<th>Jurisdiction</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1]</td>
<td>2000</td>
<td>Hytrin</td>
<td>Public</td>
<td>Consent</td>
</tr>
<tr>
<td>[2]</td>
<td>2001</td>
<td>Cardizem CD</td>
<td>Public</td>
<td>Consent</td>
</tr>
<tr>
<td>[3]</td>
<td>2002</td>
<td>Procardia XL</td>
<td>Private</td>
<td>Yes</td>
</tr>
<tr>
<td>[4]</td>
<td>2003</td>
<td>Buspar</td>
<td>Public</td>
<td>Consent</td>
</tr>
<tr>
<td>[9]</td>
<td>2006</td>
<td>Tamoxifen</td>
<td>Private</td>
<td>No</td>
</tr>
<tr>
<td>[12]</td>
<td>2010</td>
<td>Androgel</td>
<td>Public</td>
<td>No</td>
</tr>
<tr>
<td>[14]</td>
<td>2010</td>
<td>Provigil</td>
<td>Public</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Year:** Year of significant decision. **Drug:** Brand name listed, with the exception of Nolvadex, which is better known by its trade name tamoxifen. **Plaintiff:** Public indicates FTC and/or state attorneys general. Private indicates private only. **Outcome:** Yes indicates liability; No indicates no liability; Consent is a consent decree with the FTC. Mixed: after the 11th Circuit announced a narrow liability standard, on remand, the district court found that conduct violated that standard. Italicized outcomes are not final. Provigil is a district court refusal to dismiss; Androgel is a district court dismissal; Cipro (Second Circuit) is a Second Circuit opinion for which a motion to rehear en banc is pending; Cipro (California state) is on appeal to an intermediate state court.

The settlements challenged were final settlements, with the exceptions of Hytrin and Cardizem CD, which were interim agreements that barred entry during ongoing litigation.

Suits as to Plavix, Yasmin, and Ovcon are omitted. **Plavix**: Kroger v. Sanofi-Aventis, 2010 WL 1286311 (S.D. Ohio Mar. 26, 2010), dismissed an antitrust challenge to a settlement involving Plavix that was never fully consummated. **Yasmin**: Sandoz, a later-filing generic firm, has alleged that the settlement between Bayer and Barr is part of an anticompetitive conspiracy. Sandoz alleges that Bayer and Barr agreed that Bayer would enforce a patent, which had not been asserted in litigation between Bayer and Barr, against other generic firms such as Sandoz. **Ovcon**: The FTC challenged a settlement over Ovcon that does not engage the Hatch-Waxman exclusivity provisions. See FTC v. Warner Chilcott Holdings Co. III, No. 05-2179, 2007 WL 158746 (D.D.C. Jan. 22, 2007).

Several appeals courts have noticed the pay-for-delay issue but not taken a strong stand. In Andrx Pharm., Inc. v. Biovail Corp. Int’l, 256 F.3d 799 (D.C. Cir. 2001), the D.C. Circuit opined in dicta that the Cardizem CD settlements raised troubling issues. In Kaiser Found. Health Plan v. Abbott Labs., 552 F.3d 1033 (9th Cir. 2009), the Ninth Circuit affirmed a lower court ruling involving a case maintained by a non-settling plaintiff after the Hytrin remand.
<table>
<thead>
<tr>
<th>Drug</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adderall XR</td>
<td>Barr Pharms., Quarterly Report (Form 10-Q) exhs. 10.1–10.3 (Nov. 9, 2006)</td>
</tr>
<tr>
<td>Altace</td>
<td>King Pharms., Inc., Current Report (Form 8-K) exhs. 10.1–10.2 (Jan. 8, 2008); King Pharms., Inc., Quarterly Report (Form 10-Q) exhs. 10.1–10.5 (May 9, 2006)</td>
</tr>
<tr>
<td>Effexor XR</td>
<td>Weyth Labs., Annual Report (Form 10-K) exh. 10.9 (Feb. 27, 2006)</td>
</tr>
<tr>
<td>Imitrex Injection</td>
<td>Spectrum Pharmaceuticals Inc., Annual Report (Form 10-K) exhs. 10.38, 10.40 (May 14, 2007)</td>
</tr>
<tr>
<td>Lexapro</td>
<td>Forest Laboratories Inc., Quarterly Report (Form 10-Q) exh. 10.21 (Feb. 9, 2006)</td>
</tr>
<tr>
<td>Mucinex</td>
<td>Adams Respiratory Therapeutics, Inc., Quarterly Report (Form 10-Q) exh. 10.1 (May 15, 2007)</td>
</tr>
<tr>
<td>Niaspan</td>
<td>Kos Pharms., Inc., Quarterly Report (Form 10-Q) exhs. 10.2–10.4 (Aug. 9, 2005)</td>
</tr>
<tr>
<td>Plavix</td>
<td>Bristol-Myers Squibb, Quarterly Report (Form 10-Q) exhs. 99.1–99.2 (Aug. 8, 2006)</td>
</tr>
<tr>
<td>Provigil</td>
<td>Cephalon, Inc., Quarterly Report (Form 10-Q) exhs. 10.1, 10.2, 10.4, (May 10, 2006) (Barr, Mylan); Cephalon, Inc., Annual Report (Form 10-K), exhs. 10.21, 10.22 (Mar. 13, 2006) (Teva, Ranbaxy); Cephalon, Inc., Quarterly Report (Form 10-Q) exh. 10.1 (Nov. 8, 2006) (Watson and Carlsbad)</td>
</tr>
<tr>
<td>Skelaxin</td>
<td>King Pharmaceuticals Inc., Current Report (Form 8-K) exhs. 10.1, 10.2 (Jan. 8, 2008)</td>
</tr>
<tr>
<td></td>
<td>King Pharmaceuticals Inc., Quarterly Report (Form 10-Q) exh. 10.1 (Jan. 8, 2008)</td>
</tr>
<tr>
<td>Zoloft</td>
<td>Stipulation of Filing of Redacted Settlement Agreement, Pfizer, Inc. v. Zenith Goldline Pharm., Inc., Nos. 00-0408, 01-6007 (D.N.J. June 14, 2002)</td>
</tr>
</tbody>
</table>
Table 3: Typology of Settlements

<table>
<thead>
<tr>
<th>Type</th>
<th>Drugs</th>
<th>Settlements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Payment and Delay</td>
<td>51</td>
<td>60</td>
</tr>
<tr>
<td>Acquisition</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Other settlements</td>
<td>48</td>
<td>81</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>101</td>
<td>143</td>
</tr>
</tbody>
</table>

Table 4: Settlements with Monetary Payment

<table>
<thead>
<tr>
<th>Year</th>
<th>Drug</th>
<th>Sales</th>
<th>Payment</th>
<th>Entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>Nolvadex</td>
<td>400</td>
<td>Cash</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Authorized generic sales</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>BuSpar</td>
<td>400</td>
<td>Cash</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Zantac</td>
<td>2950</td>
<td>Cash</td>
<td>2</td>
</tr>
<tr>
<td>1997</td>
<td>Sinemet CR</td>
<td>150</td>
<td>Cash</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Cipro</td>
<td>900</td>
<td>Cash</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>K-Dur*</td>
<td>250</td>
<td>Retained exclusivity</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Side deals (product licenses)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>Naprelan</td>
<td>50</td>
<td>Retained exclusivity</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Side deal (intellectual property (IP))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>Lamictal</td>
<td>1100</td>
<td>Generic sales (Lamictal CD)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Niaspan</td>
<td>450</td>
<td>Retained exclusivity</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Side deals (manufacturing (mfg), promotion)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Generic sales (Advicor)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Effexor XR</td>
<td>2750</td>
<td>Retained exclusivity (+ no authorized generic)</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Generic sales (Effexor IR)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>Provigil*</td>
<td>700</td>
<td>Retained exclusivity</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Side deals (IP, development, mfg, inventory)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Generic sales (Actiq)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Altace</td>
<td>700</td>
<td>Retained exclusivity</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Side deals (development, supply)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plavix</td>
<td>3400</td>
<td>Retained exclusivity (+ no authorized generic)</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Side deal (inventory)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deal sweeteners if settlement failed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Propecia</td>
<td>150</td>
<td>Retained exclusivity</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Generic sales (Zocor, Proscar)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adderall XR*</td>
<td>900</td>
<td>Retained exclusivity (+ no authorized generic)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Side deals (development, mfg, promotion)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Generic sales (Adderall IR)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AndroGel*</td>
<td>350</td>
<td>Side deals (mfg, promotion)</td>
<td>9</td>
</tr>
<tr>
<td>2007</td>
<td>Wellbutrin XL (150 mg)</td>
<td>850</td>
<td>Waived damages for 300 mg strength</td>
<td>1</td>
</tr>
<tr>
<td>2008</td>
<td>Nexium</td>
<td>3400</td>
<td>Retained exclusivity (+ no authorized generic)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Side deal (manufacturing)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Generic sales (Prilosec and Plendil)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lipitor</td>
<td>7200</td>
<td>Retained exclusivity</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Generic sales in Canada</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Caduet</td>
<td>400</td>
<td>Joint with Lipitor</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Aggrenox</td>
<td>300</td>
<td>Retained exclusivity</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Side deal (promotion)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Year*: Year of settlement, for Provigil, year of last settlement among four first-filers. *Drug*: * indicates monetary settlements with multiple first filers (Provigil) or with both first filer and later filer (K-Dur, Adderall XR, AndroGel). *Sales*: Annual U.S. sales, in millions of dollars, measured in the calendar year of settlement or the twelve months preceding settlement, or where unavailable, the closest available year. Totals were adjusted to constant 2008 dollars using the monthly Consumer Price Index prepared by the U.S. Bureau of Labor Statistics, and rounded to the nearest $50 million increment. *Payment*: “Retained exclusivity” is discussed in Part III. “No authorized generic” provisions are discussed infra Part V.I.C. “Side deals” entail possible overpayment by the brand-name firm; see infra Part V.A. “Generic sales” entail possible underpayment by the generic firm; see infra Part V.A. “Waived damages” are discussed infra Part V.I.C. *Entry*: Time between settlement and scheduled entry, rounded to the nearest year, except for Altace, where no date appears to have been disclosed. This figure does not include immediate authorized generic sales for Nolvadex, or unexpected six-month pediatric extensions for Nolvadex and Cipro. Details of these settlements are discussed in the Appendix.
Table 5: Settlements with Retained Exclusivity

<table>
<thead>
<tr>
<th>Year</th>
<th>Drug</th>
<th>Sales</th>
<th>Full?</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>Zoloft</td>
<td>3000</td>
<td>*</td>
</tr>
<tr>
<td>2004</td>
<td>Femhrt</td>
<td>50</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Estrostep</td>
<td>50</td>
<td>*</td>
</tr>
<tr>
<td>2005</td>
<td>Lamictal CD</td>
<td>50</td>
<td>*</td>
</tr>
<tr>
<td>2006</td>
<td>Duoneb</td>
<td>250</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Imitrex tablets</td>
<td>900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Imitrex injection</td>
<td>250</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>Mucinex</td>
<td>150</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Diastat</td>
<td>100</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Valtrex</td>
<td>1350</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Adenoscan</td>
<td>350</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Avandia</td>
<td>1200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Avandamet</td>
<td>300</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Avandaryl</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Keppra</td>
<td>900</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paxil CR</td>
<td>350</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Flomax</td>
<td>1200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cardizem LA</td>
<td>100</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Exelon</td>
<td>200</td>
<td>*</td>
</tr>
<tr>
<td>2008</td>
<td>Astelin</td>
<td>200</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Optivar</td>
<td>50</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Xopenex</td>
<td>500</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Miacalcin</td>
<td>150</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Depakote ER (500 mg)</td>
<td>700</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Mirapex</td>
<td>400</td>
<td>*</td>
</tr>
</tbody>
</table>

*Year and Sales:* As in Table 2. *Full:* Indicates whether the entry date was early enough to permit 180 days of sales prior to patent expiration.

Table 6: Pay-for-Delay Settlements Summarized

<table>
<thead>
<tr>
<th>Type</th>
<th>Drugs</th>
<th>Settlements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monetary</td>
<td>21</td>
<td>28</td>
</tr>
<tr>
<td>Retained exclusivity only</td>
<td>25</td>
<td>27</td>
</tr>
<tr>
<td>Interim agreement</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Supply agreement</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>51</td>
<td>60</td>
</tr>
</tbody>
</table>
Table 7: Description of Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paragraph IV challenge</td>
<td>=1 if a generic drug maker, in year $t$, officially asserts to FDA its plan</td>
</tr>
<tr>
<td></td>
<td>to launch a competing therapeutically equivalent drug, and that one or more</td>
</tr>
<tr>
<td></td>
<td>brand-name patents are invalid or not infringed</td>
</tr>
<tr>
<td>Sales</td>
<td>Sales of drug in year $t$ (annualized $$ billions)</td>
</tr>
<tr>
<td>Oral dosage form</td>
<td>=1 if drug offered in a tablet, capsule, or liquid form</td>
</tr>
<tr>
<td>New chemical entity</td>
<td>=1 if drug contains a novel active ingredient (see Appendix for details);</td>
</tr>
<tr>
<td></td>
<td>exempt from challenge during the first four years after approval</td>
</tr>
<tr>
<td>Therapeutic class dummies</td>
<td>Indicators for primary use of drug (e.g., cardiovascular)</td>
</tr>
<tr>
<td>Observation year dummies</td>
<td>Indicators for observation year (2001 through 2009)</td>
</tr>
<tr>
<td>Eligibility year dummies</td>
<td>Indicators for years since drug’s initial eligibility for challenge (1 through 9)</td>
</tr>
<tr>
<td>Future sales</td>
<td>Sales of all other drugs in therapeutic class in year $t+3$</td>
</tr>
<tr>
<td>Patent type</td>
<td>=1 if all patents pertain to the drug’s active ingredient (AI); excludes</td>
</tr>
<tr>
<td></td>
<td>patents on enantiomers or other isomers or polymorphs and other crystal</td>
</tr>
<tr>
<td></td>
<td>forms</td>
</tr>
<tr>
<td>Non-AI only</td>
<td>=1 if no patents pertain to the drug’s AI</td>
</tr>
<tr>
<td>Both types</td>
<td>=1 if at least one patent pertains to the AI, and at least one does not</td>
</tr>
<tr>
<td>Non-AI patents</td>
<td>=1 if both types of patent, and last non-AI patent expires later than last</td>
</tr>
<tr>
<td></td>
<td>AI patent</td>
</tr>
<tr>
<td>Extra years from non-AI</td>
<td>Years of difference between expiration of last-expiring non-AI patent and</td>
</tr>
<tr>
<td></td>
<td>last-expiring AI patent (=0 if no additional life)</td>
</tr>
<tr>
<td>Other patent measures</td>
<td></td>
</tr>
<tr>
<td>First quartile citations</td>
<td>Number of patents containing at least one patent in the first quartile of</td>
</tr>
<tr>
<td></td>
<td>forward citations, among patents issued in that year</td>
</tr>
<tr>
<td>First quartile family size</td>
<td>Number of patents containing at least one patent in the first quartile,</td>
</tr>
<tr>
<td></td>
<td>among patents issued in that year, of number of jurisdictions in which</td>
</tr>
<tr>
<td></td>
<td>patent was applied for</td>
</tr>
<tr>
<td>Early patents</td>
<td>Number of patents with application filed prior to drug approval</td>
</tr>
<tr>
<td>Late patents</td>
<td>Number of patents with application filed after drug approval</td>
</tr>
</tbody>
</table>
Table 8: Summary of Variables

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paragraph IV challenge</td>
<td>0.51</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Oral dosage form</td>
<td>0.66</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>New chemical entity</td>
<td>0.35</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Old antibiotics</td>
<td>0.07</td>
<td>0.26</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Approval year</td>
<td>2001.87</td>
<td>3.54</td>
<td>1990</td>
<td>2008</td>
</tr>
<tr>
<td>Sales ($ billion)</td>
<td>0.20</td>
<td>0.51</td>
<td>0</td>
<td>5.99</td>
</tr>
</tbody>
</table>

| Patents                  |       |           |      |     |
| AI only                  | 0.12  | 0.33      | 0    | 1   |
| Non-AI only              | 0.47  | 0.50      | 0    | 1   |
| Both AI and non-AI      | 0.41  | 0.49      | 0    | 1   |
| No extra time from non-AI| 0.08  | 0.28      | 0    | 1   |
| Extra time from non-AI  | 0.33  | 0.47      | 0    | 1   |
| Extra years from non-AI | 1.86  | 3.46      | 0    | 17.2|
| Bottom citations        | 0.98  | 1.27      | 0    | 8   |
| Bottom family size      | 0.79  | 1.43      | 0    | 9   |
| Early patents           | 3.26  | 3.06      | 0    | 25  |
| Late patents            | 0.36  | 1.07      | 0    | 9   |

\( N = 479. \)

Table 9: Data Structure of Duration Analysis

<table>
<thead>
<tr>
<th>Period</th>
<th>Challenge</th>
<th>No challenge</th>
<th>Censored</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>77</td>
<td>16.1%</td>
<td>362</td>
<td>75.6%</td>
</tr>
<tr>
<td>2</td>
<td>43</td>
<td>12.0%</td>
<td>292</td>
<td>81.8%</td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>10.8%</td>
<td>228</td>
<td>79.4%</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>14.2%</td>
<td>171</td>
<td>76.0%</td>
</tr>
<tr>
<td>5</td>
<td>19</td>
<td>11.4%</td>
<td>131</td>
<td>78.4%</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>12.0%</td>
<td>81</td>
<td>64.8%</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>11.2%</td>
<td>56</td>
<td>70.0%</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>5.4%</td>
<td>32</td>
<td>57.1%</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>6.4%</td>
<td>0</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Total 231 1353 223 1807
Table 10: Logit Model of Hazard of Challenge—Baseline and Patent Measures

(Dependent variable: challenge = 1 if drug \(i\) is challenged in year \(t\))

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>1.214***</td>
<td>1.194***</td>
<td>1.191***</td>
<td>1.208***</td>
<td>1.314***</td>
</tr>
<tr>
<td></td>
<td>(0.302)</td>
<td>(0.313)</td>
<td>(0.312)</td>
<td>(0.302)</td>
<td>(0.287)</td>
</tr>
<tr>
<td>Oral dosage form</td>
<td>1.090***</td>
<td>1.217***</td>
<td>1.222***</td>
<td>1.119***</td>
<td>1.120***</td>
</tr>
<tr>
<td></td>
<td>(0.285)</td>
<td>(0.293)</td>
<td>(0.293)</td>
<td>(0.292)</td>
<td>(0.292)</td>
</tr>
<tr>
<td>NCE</td>
<td>−0.024</td>
<td>0.108</td>
<td>0.114</td>
<td>0.152</td>
<td>0.184</td>
</tr>
<tr>
<td></td>
<td>(0.285)</td>
<td>(0.293)</td>
<td>(0.293)</td>
<td>(0.292)</td>
<td>(0.292)</td>
</tr>
<tr>
<td>Old antibiotics</td>
<td>0.485</td>
<td>0.689</td>
<td>0.683</td>
<td>0.588</td>
<td>0.815</td>
</tr>
<tr>
<td></td>
<td>(0.623)</td>
<td>(0.617)</td>
<td>(0.616)</td>
<td>(0.617)</td>
<td>(0.636)</td>
</tr>
<tr>
<td>Patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non–AI only</td>
<td>0.964***</td>
<td>0.974***</td>
<td>0.566**</td>
<td>1.210***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.360)</td>
<td>(0.360)</td>
<td>(0.253)</td>
<td>(0.384)</td>
<td></td>
</tr>
<tr>
<td>Both types</td>
<td>0.998***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.337)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra time from non-AI:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0.872**</td>
<td></td>
<td></td>
<td></td>
<td>1.051**</td>
</tr>
<tr>
<td></td>
<td>(0.437)</td>
<td></td>
<td></td>
<td></td>
<td>(0.456)</td>
</tr>
<tr>
<td>Yes</td>
<td>1.027***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.341)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra years</td>
<td></td>
<td></td>
<td></td>
<td>0.096***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.029)</td>
<td></td>
</tr>
<tr>
<td>First quartile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.633</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.561)</td>
</tr>
<tr>
<td>Second quartile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.396***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.414)</td>
</tr>
<tr>
<td>Third quartile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.433***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.396)</td>
</tr>
<tr>
<td>Top quartile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.953***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.463)</td>
</tr>
<tr>
<td>Constant</td>
<td>−7.350***</td>
<td>−8.368***</td>
<td>−8.382***</td>
<td>−7.980***</td>
<td>−8.731***</td>
</tr>
<tr>
<td></td>
<td>(1.367)</td>
<td>(1.450)</td>
<td>(1.452)</td>
<td>(1.443)</td>
<td>(1.538)</td>
</tr>
<tr>
<td>Eligibility year indicators</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Obs. Year indicators</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Therapeutic class f.e.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1807</td>
<td>1807</td>
<td>1807</td>
<td>1807</td>
<td>1807</td>
</tr>
</tbody>
</table>

Standard errors clustered by drug are in parentheses. Asterisks indicate statistical significance at the *** 1%, ** 5%, and * 10% levels.
Table 11: Logit Model of Hazard of Challenge—Alternative Patent Quality Measures

(Dependent variable: challenge = 1 if drug \( i \) is challenged in year \( t \))

<table>
<thead>
<tr>
<th></th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>1.206***</td>
<td>1.249***</td>
<td>1.178***</td>
</tr>
<tr>
<td></td>
<td>(0.306)</td>
<td>(0.306)</td>
<td>(0.301)</td>
</tr>
<tr>
<td>Oral dosage form</td>
<td>1.170***</td>
<td>1.150***</td>
<td>1.138***</td>
</tr>
<tr>
<td></td>
<td>(0.288)</td>
<td>(0.291)</td>
<td>(0.296)</td>
</tr>
<tr>
<td>NCE</td>
<td>–0.014</td>
<td>0.041</td>
<td>–0.088</td>
</tr>
<tr>
<td></td>
<td>(0.194)</td>
<td>(0.202)</td>
<td>(0.203)</td>
</tr>
<tr>
<td>Old antibiotic</td>
<td>0.586</td>
<td>0.574</td>
<td>0.331</td>
</tr>
<tr>
<td></td>
<td>(0.616)</td>
<td>(0.607)</td>
<td>(0.644)</td>
</tr>
<tr>
<td>Patents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First quartile citations</td>
<td>0.138**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First quartile family size</td>
<td>0.122**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early patents</td>
<td></td>
<td>–0.018</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.034)</td>
<td></td>
</tr>
<tr>
<td>Late patents</td>
<td></td>
<td>0.282***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.106)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>–7.517***</td>
<td>–7.608***</td>
<td>–7.766***</td>
</tr>
<tr>
<td></td>
<td>(1.362)</td>
<td>(1.395)</td>
<td>(1.376)</td>
</tr>
<tr>
<td>Eligibility year indicators</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Obs. Year indicators</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Therapeutic class f.e.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1807</td>
<td>1807</td>
<td>1807</td>
</tr>
</tbody>
</table>

Standard errors clustered by drug are in parentheses. Asterisks indicate statistical significance at the *** 1%, ** 5%, and * 10% levels.
Table 12: Logit Model of Hazard of Challenge—Non-Parametric Sales

(Dependent variable: challenge = 1 if drug $i$ is challenged in year $t$)

<table>
<thead>
<tr>
<th></th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second quartile</td>
<td>1.133**</td>
<td>1.134**</td>
<td>1.120**</td>
<td>1.144**</td>
</tr>
<tr>
<td></td>
<td>(0.489)</td>
<td>(0.488)</td>
<td>(0.483)</td>
<td>(0.481)</td>
</tr>
<tr>
<td>Third quartile</td>
<td>3.057***</td>
<td>3.055***</td>
<td>3.054***</td>
<td>3.056***</td>
</tr>
<tr>
<td></td>
<td>(0.489)</td>
<td>(0.488)</td>
<td>(0.481)</td>
<td>(0.480)</td>
</tr>
<tr>
<td>Fourth quartile</td>
<td>4.204***</td>
<td>4.203***</td>
<td>4.208***</td>
<td>4.191***</td>
</tr>
<tr>
<td></td>
<td>(0.486)</td>
<td>(0.485)</td>
<td>(0.479)</td>
<td>(0.477)</td>
</tr>
<tr>
<td>Oral dosage form</td>
<td>1.708***</td>
<td>1.711***</td>
<td>1.625***</td>
<td>1.630***</td>
</tr>
<tr>
<td></td>
<td>(0.367)</td>
<td>(0.367)</td>
<td>(0.367)</td>
<td>(0.368)</td>
</tr>
<tr>
<td>NCE</td>
<td>0.200</td>
<td>−0.194</td>
<td>−0.185</td>
<td>−0.154</td>
</tr>
<tr>
<td></td>
<td>(0.202)</td>
<td>(0.203)</td>
<td>(0.205)</td>
<td>(0.204)</td>
</tr>
<tr>
<td>Old antibiotic</td>
<td>0.244</td>
<td>0.239</td>
<td>0.151</td>
<td>0.270</td>
</tr>
<tr>
<td></td>
<td>(0.769)</td>
<td>(0.764)</td>
<td>(0.763)</td>
<td>(0.771)</td>
</tr>
<tr>
<td>Patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non–AI only</td>
<td>0.766**</td>
<td>0.779**</td>
<td>0.445*</td>
<td>0.772**</td>
</tr>
<tr>
<td></td>
<td>(0.347)</td>
<td>(0.348)</td>
<td>(0.256)</td>
<td>(0.324)</td>
</tr>
<tr>
<td>Both types</td>
<td>0.781**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.318)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No extra time from non-AI</td>
<td>0.654</td>
<td></td>
<td>0.609</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.415)</td>
<td></td>
<td>(0.403)</td>
<td></td>
</tr>
<tr>
<td>Extra time from non-AI</td>
<td>0.810**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.326)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra years from non-AI</td>
<td></td>
<td></td>
<td>0.072**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.030)</td>
<td></td>
</tr>
<tr>
<td>First quartile</td>
<td></td>
<td></td>
<td>0.584</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.464)</td>
<td></td>
</tr>
<tr>
<td>Second quartile</td>
<td></td>
<td></td>
<td>0.745**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.345)</td>
<td></td>
</tr>
<tr>
<td>Third quartile</td>
<td></td>
<td></td>
<td>0.809**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.352)</td>
<td></td>
</tr>
<tr>
<td>Top quartile</td>
<td></td>
<td></td>
<td>1.364***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.429)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>−10.510***</td>
<td>−10.508***</td>
<td>−10.108***</td>
<td>−10.366***</td>
</tr>
<tr>
<td></td>
<td>(1.497)</td>
<td>(1.497)</td>
<td>(1.484)</td>
<td>(1.483)</td>
</tr>
<tr>
<td>Eligibility year indicators</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Obs. Year indicators</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Therapeutic class f.e.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1807</td>
<td>1807</td>
<td>1807</td>
<td>1807</td>
</tr>
</tbody>
</table>

Standard errors clustered by drug are in parentheses. Asterisks indicate statistical significance at the *** 1%, ** 5%, and * 10% levels.
Table 13: Logit Model of Hazard of Challenge—Lagged Sales

(Dependent variable: challenge = 1 if drug \(i\) is challenged in year \(t\))

<table>
<thead>
<tr>
<th>Sales</th>
<th>(13)</th>
<th>(14)</th>
<th>(15)</th>
<th>(16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current</td>
<td>-0.490</td>
<td>-0.458</td>
<td>-0.487</td>
<td>-0.537</td>
</tr>
<tr>
<td></td>
<td>(0.936)</td>
<td>(0.943)</td>
<td>(0.980)</td>
<td>(0.989)</td>
</tr>
<tr>
<td>1 year lag</td>
<td>2.498</td>
<td>2.430</td>
<td>2.585</td>
<td>2.745</td>
</tr>
<tr>
<td></td>
<td>(1.656)</td>
<td>(1.670)</td>
<td>(1.726)</td>
<td>(1.788)</td>
</tr>
<tr>
<td>2 year lag</td>
<td>-1.222</td>
<td>-1.191</td>
<td>-1.261</td>
<td>-1.336</td>
</tr>
<tr>
<td></td>
<td>(1.334)</td>
<td>(1.338)</td>
<td>(1.342)</td>
<td>(1.447)</td>
</tr>
<tr>
<td>3 year lag</td>
<td>0.747</td>
<td>0.756</td>
<td>0.678</td>
<td>0.848</td>
</tr>
<tr>
<td></td>
<td>(1.338)</td>
<td>(1.340)</td>
<td>(1.283)</td>
<td>(1.253)</td>
</tr>
<tr>
<td>Oral dosage form</td>
<td>1.381***</td>
<td>1.388***</td>
<td>1.309***</td>
<td>1.290***</td>
</tr>
<tr>
<td></td>
<td>(0.321)</td>
<td>(0.321)</td>
<td>(0.320)</td>
<td>(0.319)</td>
</tr>
<tr>
<td>NCE</td>
<td>0.035</td>
<td>0.041</td>
<td>0.110</td>
<td>0.123</td>
</tr>
<tr>
<td></td>
<td>(0.234)</td>
<td>(0.233)</td>
<td>(0.233)</td>
<td>(0.229)</td>
</tr>
<tr>
<td>Old antibiotic</td>
<td>0.571</td>
<td>0.561</td>
<td>0.498</td>
<td>0.681</td>
</tr>
<tr>
<td></td>
<td>(0.620)</td>
<td>(0.619)</td>
<td>(0.622)</td>
<td>(0.641)</td>
</tr>
<tr>
<td>Patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non–AI only</td>
<td>0.855**</td>
<td>0.867**</td>
<td>0.571**</td>
<td>1.128***</td>
</tr>
<tr>
<td></td>
<td>(0.369)</td>
<td>(0.369)</td>
<td>(0.267)</td>
<td>(0.387)</td>
</tr>
<tr>
<td>Both types</td>
<td>0.988***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.346)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No extra time from non-AI</td>
<td>0.835*</td>
<td></td>
<td></td>
<td>1.003**</td>
</tr>
<tr>
<td></td>
<td>(0.446)</td>
<td></td>
<td></td>
<td>(0.467)</td>
</tr>
<tr>
<td>Extra time from non-AI</td>
<td>1.024***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.351)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra years from non-AI</td>
<td></td>
<td>0.120***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.030)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First quartile</td>
<td></td>
<td></td>
<td>0.567</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.538)</td>
<td></td>
</tr>
<tr>
<td>Second quartile</td>
<td></td>
<td></td>
<td>1.260***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.425)</td>
<td></td>
</tr>
<tr>
<td>Third quartile</td>
<td></td>
<td></td>
<td>1.479***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.411)</td>
<td></td>
</tr>
<tr>
<td>Top quartile</td>
<td></td>
<td></td>
<td>2.311***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.461)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-4.820***</td>
<td>-4.825***</td>
<td>-4.507***</td>
<td>-5.123***</td>
</tr>
<tr>
<td></td>
<td>(1.166)</td>
<td>(1.167)</td>
<td>(1.122)</td>
<td>(1.175)</td>
</tr>
<tr>
<td>Eligibility year indicators</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Obs. Year indicators</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Therapeutic class f.e.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1624</td>
<td>1624</td>
<td>1624</td>
<td>1624</td>
</tr>
</tbody>
</table>

Standard errors clustered by drug are in parentheses. Asterisks indicate statistical significance at the *** 1%, ** 5%, and * 10% levels.
Table 14: Logit Model of Hazard of Challenge—Future Sales

(Dependent variable: challenge = 1 if drug $i$ is challenged in year $t$)

<table>
<thead>
<tr>
<th></th>
<th>(17)</th>
<th>(18)</th>
<th>(19)</th>
<th>(20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>2.047***</td>
<td>2.011***</td>
<td>2.092***</td>
<td>2.014***</td>
</tr>
<tr>
<td></td>
<td>(0.702)</td>
<td>(0.711)</td>
<td>(0.740)</td>
<td>(0.716)</td>
</tr>
<tr>
<td>Future sales</td>
<td>0.681</td>
<td>0.660</td>
<td>0.709</td>
<td>0.586</td>
</tr>
<tr>
<td></td>
<td>(0.519)</td>
<td>(0.528)</td>
<td>(0.545)</td>
<td>(0.563)</td>
</tr>
<tr>
<td>Oral dosage form</td>
<td>0.753**</td>
<td>0.752**</td>
<td>0.591</td>
<td>0.579</td>
</tr>
<tr>
<td></td>
<td>(0.379)</td>
<td>(0.378)</td>
<td>(0.380)</td>
<td>(0.378)</td>
</tr>
<tr>
<td>NCE</td>
<td>0.291</td>
<td>0.316</td>
<td>0.352</td>
<td>0.422</td>
</tr>
<tr>
<td></td>
<td>(0.255)</td>
<td>(0.255)</td>
<td>(0.261)</td>
<td>(0.258)</td>
</tr>
<tr>
<td>Patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non–AI only</td>
<td>0.926*</td>
<td>0.963*</td>
<td>0.462</td>
<td>1.168**</td>
</tr>
<tr>
<td></td>
<td>(0.502)</td>
<td>(0.505)</td>
<td>(0.350)</td>
<td>(0.527)</td>
</tr>
<tr>
<td>Both types</td>
<td>1.126**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.475)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No extra time from non-AI</td>
<td>0.795</td>
<td></td>
<td>0.929</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.711)</td>
<td></td>
<td>(0.711)</td>
<td></td>
</tr>
<tr>
<td>Extra time from non-AI</td>
<td>1.197**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.481)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra years from non-AI</td>
<td></td>
<td></td>
<td>0.097**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.040)</td>
<td></td>
</tr>
<tr>
<td>First quartile</td>
<td></td>
<td></td>
<td>0.789</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.740)</td>
<td></td>
</tr>
<tr>
<td>Second quartile</td>
<td></td>
<td></td>
<td>1.578***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.528)</td>
<td></td>
</tr>
<tr>
<td>Third quartile</td>
<td></td>
<td></td>
<td>1.590***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.531)</td>
<td></td>
</tr>
<tr>
<td>Top quartile</td>
<td></td>
<td></td>
<td>1.853***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.606)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.168)</td>
<td>(1.175)</td>
<td>(1.134)</td>
<td>(1.237)</td>
</tr>
<tr>
<td>Eligibility year indicators</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Obs. Year indicators</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Therapeutic class f.e.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1171</td>
<td>1171</td>
<td>1171</td>
<td>1171</td>
</tr>
</tbody>
</table>

Standard errors clustered by drug are in parentheses. Asterisks indicate statistical significance at the *** 1%, ** 5%, and * 10% levels.
Appendix: Selected Agreements Between Drug Makers

This Appendix summarizes the major features of selected agreements between brand-name and generic drug makers, the subject of Chapters 2 and 3. The agreements are arranged by drug, in rough chronological order of first settlement. In many cases, the first-filing generic drug maker was eligible for the 180-day exclusivity period at the time of settlement, and retained eligibility after settlement. Some exceptions are noted below. Annual sales (for the United States, except where otherwise noted) are included for each drug, generally for the year of settlement or the year before.

In the interest of concision, some terms are omitted, as well as details such as divisions of—and changes in—ownership of a New Drug Application, its patents, or its marketing. Finally, it bears note that this Appendix is a synthesis of public information that is often fragmentary or conflicting, creating a significant possibility that the information is incomplete or even erroneous in significant respects.

1. Nolvadex

Tamoxifen citrate, marketed by Zeneca under the brand name Nolvadex, is a treatment for breast cancer. Sales in 1992 were $265 million.

Agreement: March 1993, with first filer Barr. Zeneca agreed to pay $66 million to Barr and Barr’s raw materials supplier. In addition to cash, Barr was permitted to make authorized generic sales of Zeneca’s product, sold at a 15 percent discount to Zeneca’s version. Barr agreed to enter with its own ANDA product upon patent expiration in August 2002. Pediatric exclusivity extended Zeneca’s exclusivity until February 2003, when Barr and other generic manufacturers received ANDA approval.
The parties also agreed to ask the Federal Circuit to vacate a district court ruling that the relevant patent was invalid, a request the Federal Circuit granted.

Upon settlement, Barr initially changed its Paragraph IV certification to Paragraph III, thereby certifying that it would wait to enter until patent expiration. It later reverted to a Paragraph IV certification and asserted a continued entitlement to the 180-day exclusivity period. Barr’s reversion raised doubts about the approvability of later filers.


2. BuSpar

Buspirone hydrochloride, marketed by Bristol-Myers Squibb under the brand name BuSpar, is a widely prescribed anti-anxiety drug. Sales in 1995 were $288 million.

3. Zantac

Ranitidine, marketed by Glaxo under the brand name Zantac, is a treatment for ulcers. Sales in 1995 were $2.15 billion.

Agreement: October 1995, with first filer Genpharm. Glaxo paid Genpharm $132.5 million, if the process of inference discussed below is correct. In addition, Glaxo granted Genpharm licenses and supply agreements to sell a generic version of Zantac in other countries. Genpharm also received additional compensation from its selective waiver of the 180-day exclusivity period, in favor of later filer Granutec, which permitted Granutec to enter the market earlier than would otherwise be possible. Genpharm agreed not to launch any ranitidine product until July 1997, when the basic Zantac patent expired, or to launch a “Form 2” product—the form actually marketed by Glaxo—until a second patent, covering Form 2 and also at issue in the patent litigation, expired in 2002.

As the Financial Times described the Zantac settlement at the time, “With so much at stake, the fact that Glaxo is having to pay Genpharm to turn it from a competitor into a distributor is money well spent.” Glaxo’s market valuation increased by nearly £2 billion immediately following settlement.

The amount of Glaxo’s payment was not disclosed at the time, but can be inferred from FTC 2002. In that report, FTC reports a list of drugs with cash settlements, including Product I, which entailed a payment of $132.5 million. Product I is the only listed drug whose sales, like Zantac’s, exceeded $1 billion in the year of agreement. The agreement involving Product I included a delay of one year, nine months, which matches the delay between the Zantac agreement (in October 1995) and the expiration of the first patent in issue in July 1997. The agreement involving Product I, like Zantac, resolved patent litigation in other countries, an unusual feature of the agreement. Similarly, the Glaxo-Genpharm pact settled Zantac litigation outside the United States. Finally, Product I fits none of the other cases that have received antitrust attention from the FTC or private parties. At the time of the settlement, Glaxo described the amount of settlement as “not material,” but a payment that size would likely not be considered material for a company as large as Glaxo.


4. Sinemet CR

A mixture of carbidopa and levodopa in a controlled-release form, marketed by Merck as Sinemet CR, is a treatment for Parkinson’s disease. Sales in 1998 were $125 million.
Agreement: December 1995, with first filer Purepac. Purepac received three equal annual payments, likely totaling $1.75 million. (A payment of that size is reported for “Product P” in FTC 2002. Product P is the only settlement with three equal payments in the report, and the time between the Sinemet CR agreement and patent expiration is, like Product P, ten years.) Purepac agreed to enter upon patent expiration in June 2006.


5. Cipro

Ciprofloxacin, marketed by Bayer under the brand name Cipro, is an antibiotic. The exclusivity and other provisions of the Hatch-Waxman Act do not apply to many antibiotics, but Cipro is an exception to the exception, to which the provisions do apply. Sales in 1996 were $680 million.

Agreement: January 1997, with first filer Barr. Bayer paid Barr $398 million: $49 million initially, followed by quarterly payments. Although the agreement provided for authorized generic sales, as in the Nolvadex agreement, this provision was not implemented. Instead, Bayer exercised its option under the agreement to pay cash. Barr agreed to enter upon patent expiration in December 2003, a date extended to June 2004 with pediatric exclusivity.
Barr later secured authorized generic sales during the last six months of Bayer’s exclusivity. These sales began in June 2003, six months before the expected December 2003 expiration of the relevant patent. Upon Bayer’s receipt of a pediatric extension, the arrangement was extended for another six months.

As with Nolvadex, Barr initially changed its Paragraph IV certification to Paragraph III, then later reverted to a Paragraph IV certification and asserted its continued entitlement to the 180-day exclusivity period.


6. K-Dur

Potassium chloride in an extended-release form, marketed by Schering-Plough under the brand name K-Dur, is a replacement for electrolytes lost from the body as a side effect of certain antihypertension drugs. Sales in 1997 were $190 million.

Agreement: June 1997, with first filer Upsher-Smith. Schering agreed to pay Upsher $60 million, described as payment for rights to five Upsher products. Schering also agreed to pay milestone royalty payments and a percentage royalty on sales, but the
project did not progress far and these payments appear not to have been made. Upsher agreed to enter in September 2001, prior to patent expiration in September 2006. Upsher agreed to delay marketing not only the particular drug at issue, but also any other “microencapsulated” extended-release potassium chloride product. (Upsher contested this interpretation of the relevant provision of the agreement.)

**Agreement:** January 1998, with later filer ESI Lederle, a division of American Home Products. Schering agreed to pay ESI $30 million, including a $5 million noncontingent payment, $10 million payment contingent on approval of ESI’s generic product, and $15 million for certain licenses. Schering received licenses from ESI pertaining to generic enalapril and buspirone, grants described as the basis for some (but not all) of Schering’s cash payments. ESI agreed to enter in January 2004, prior to patent expiration in September 2006. ESI agreed to delay marketing not only the particular drug at issue, but also any other generic versions of the product, even if noninfringing.

**Sources:** Legal: Schering-Plough Corp. v. FTC, 402 F.3d 1056 (11th Cir. 2005); In re Schering-Plough Corp., No. 9297, 2003 WL 22989651 (FTC Dec. 8, 2003); Complaint, In re Schering-Plough Corp., No. 9297 (FTC Mar. 13, 2001).

**7. Cardizem CD**

Diltiazem hydrochloride in a controlled-release form, marketed by Hoescht Marion Roussel under the brand name Cardizem CD, is a widely prescribed medication for hypertension. Sales in 1997 were $692 million.
**Agreement:** September 1997, with first filer Andrx. Hoescht agreed to pay Andrx $10 million per quarter, starting when Andrx received approval for its generic product, plus an additional $60 million per year *for the same period* upon a final, unappealable determination of noninfringement. Total payments equaled $90 million. The settlement did not provide relief from the prospect of losing the patent suit, which continued. Andrx agreed to delay marketing not only the particular drug at issue, but also other noninfringing products. In June 1999, Hoescht and Andrx entered into a stipulation settling the patent infringement case and terminating the agreement. The FDA eventually approved a reformulated Andrx product, and Andrx launched with exclusivity.

**Agreement:** May 1999, with later filer Purepac Pharmaceutical. Hoescht licensed Purepac to enter the market (and pay royalties) once Andrx’s exclusivity expired. At the time of settlement, entry was anticipated for mid-2000, but the earlier triggering of Andrx’s exclusivity permitted Purepac’s entry in December 1999.


8. **Hytrin**

Terazosin hydrochloride, marketed by Abbott under the brand name Hytrin, is a treatment for high blood pressure. Sales in 1998 were $542 million.
Agreement: April 1998, with first filer Geneva. Abbott agreed to pay Geneva up to $101 million. The formula was to pay $4.5 million per month until a district court judgment. If the generic firm won in district court, $4.5 million per month would go into escrow during appeal, payable if Geneva won. Geneva agreed to delay marketing not only the particular drug at issue, but other products containing terazosin hydrochloride. The agreement did not resolve the patent suit but only addressed entry during its pendency.

Agreement: March 1998, with later filer Zenith. Abbott agreed to pay Zenith up to about $45 million, $3 million up front and the rest periodically. This agreement, unlike the Geneva agreement, was a final settlement. Zenith agreed to enter in February 2000, and to delay marketing of any terazosin hydrochloride product.


9. Naprelan

Naproxen sodium in an extended-release form, marketed by Elan under the brand name Naprelan, is a non-steroidal drug used to manage pain, fever, and inflammation. Sales in 1999 were $59 million.

Agreement: May 1999, with first filer SkyePharma. Elan agreed to pay SkyePharma an undisclosed amount, described as a fee for licensing certain
SkyePharma patents. SkyePharma agreed to enter at an undisclosed date prior to patent expiration in 2014.

Naprelan appears to be Product O in FTC 2002. This is the only drug where the payment took the form of a royalty paid on the brand-name firm’s sales. The payments match, roughly, the “millions” asserted in a suit brought by another generic firm. (The FTC reported a $5 million payment in the first year for Product O.) Naprelan sales fit within the relevant sales range of less than $100 million. If this inference is correct, then Naprelan’s permitted entry date was November 2001, and the licensing conditions were an 8.5 percent royalty on Elan’s sales during the first two years of SkyePharma’s nonentry, followed by a 7.5 percent royalty during the subsequent six months.


10. Procardia XL

Nifedipine in an extended-release form, marketed by Pfizer under the brand name Procardia XL, is a medication to reduce blood pressure and manage angina. Sales in 2000 were $311 million.

**Agreement:** February 2000, with first filer Mylan. Different strengths can have different first filers, and Mylan was first filer only for the 30-milligram strength. Pfizer
granted Mylan a license to sell an authorized generic version, covering all three strengths of the drug. Pfizer also granted Mylan a nonexclusive license for its ANDA product, apparently with a royalty on net sales, but so far as appears this was never used.

The arrangement raised the issue of whether Mylan could maintain an approval bottleneck, despite the authorized generic sales. The FDA and a district court concluded no, because such sales trigger the exclusivity period.


11. Lamictal

Lamotrigine, marketed by Glaxo under the brand name Lamictal, is used in the treatment of epilepsy. Sales in 2005 were $1.03 billion.

Settlement: February 2005, with first filer Teva. Teva received an exclusive license to distribute a second product, a generic version of a minor variant of the product—a chewable version—starting in June 2005. On the main product, Teva agreed

22 The basis for this conclusion is that Procardia XL is likely the second “supply agreement” on p. 30 of FTC 2002. As with Procardia XL, the supply agreement described by the FTC involved a drug with sales between $250 million and $500 million; was reached ten years, nine months before patent expiration (compare the February 2000 Procardia agreement and U.S. Patent No. 5,264,446, which expires in November 2010); took the form of an exclusive distribution deal; included an option of choosing a patent license agreement instead of private-label sales; and covered strengths on which the generic firm had not filed an ANDA.
to enter in July 2008, six months prior to patent expiration (including pediatric exclusivity) in January 2009. Glaxo appears to have promised not to launch an authorized generic product during Teva’s exclusivity period. (Glaxo’s annual report described the grant as “exclusive.” In an analyst call, Glaxo declined to answer whether it had agreed to forego an authorized generic product.)

Sources: Legal: Complaint, SmithKline Beecham Corp. v. Teva Pharmaceuticals USA, No. 02-3779 (D.N.J. Aug. 5, 2002); Letter from Gary Buehler, Dir. of Office of Generic Drugs, FDA, to Philip Erickson, TEVA Pharmaceuticals, USA (Aug. 30, 2006). Other: GlaxoSmithKline, Fourth Quarter Report 2005; Teva Pharm. Indus., Registration of Securities of Foreign Private Issuers (Form 20-F) (Mar. 20, 2006); GlaxoSmithKline, Annual Report (2006); Press Release, TEVA Pharm. USA, Teva Announces Settlement of Lamictal Litigation with GlaxoSmithKline (Feb. 17, 2005).

12. Niaspan

Niacin in an extended-release form, marketed by Kos (later part of Abbott) under the brand name Niaspan, is a medicine to manage cholesterol, primarily by raising the level of HDL (“good”) cholesterol. Sales in the 12 months ending February 2005 were $403 million.

Agreement: April 2005, with first filer Barr. Kos agreed to pay Barr an upfront fee, plus payments to “stand ready” to produce Kos product if the need arose, plus additional payments to co-promote Niaspan to obstetricians and gynecologists. Barr agreed to enter in September 2013, about four years prior to expiration of the latest expiring patent.
The parties reached similar terms as to Advicor, a second Kos product containing a mixture of Niaspan and lovastatin. (Lovastatin lowers LDL (“bad”) cholesterol.) Worldwide sales of Advicor in 2005 were $116 million. At the time of settlement, Barr had not yet filed an ANDA on Advicor, but apparently planned to. The agreement posed no legal bar to Barr filing an ANDA, although its licensing terms from Kos were the same under its own ANDA or as an authorized generic.


13. Mircette

A mixture of desogestrel and ethinyl estradiol (two synthetic hormones), marketed at the time by Organon under the brand name Mircette, is an oral contraceptive. Sales in 2001 were about $150 million.
Agreement: December 2005, with first filer Barr, following a letter of intent signed in June 2005. Barr purchased the rights to Mircette for $142 million, and made an additional payment of $13.75 million to Organon’s patent licensor to settle the patent litigation. The product purchase resolved litigation in which Barr had filed an ANDA, been sued, won a judgment of noninfringement, and launched at risk, then suffered a reversal and remand at the Federal Circuit.


14. Effexor XR

Venlafaxine hydrochloride in an extended-release form, marketed by Wyeth under the brand name Effexor XR, is an antidepressant. Sales in 2005 were $2.6 billion.

Agreement: December 2005, with first filer Teva. Wyeth granted Teva an exclusive license to sell an older product, the immediate-release version of Effexor, starting in June 2006. Teva agreed to enter as to Effexor XR in July 2010, prior to the expiration of patents on extended-release formulations. (Teva did not challenge the compound patent, which expires in June 2008.) Wyeth appears to have promised not to
launch a competing authorized generic product during Teva’s exclusivity. The parties also settled Canadian litigation.

**Agreement**: July 2008, with later filer Impax. Impax agreed to launch, with a royalty on sales, in June 2011, or (in “limited circumstances”) in January 2011.


15. Provigil

Modafinil, marketed by Cephalon under the brand name Provigil, is a stimulant, what Cephalon calls “the first in a new class of wake-promoting agents.” Sales in 2006 were $685 million.

**Agreements**: Late 2005 and early 2006, with four first filers: Teva, Ranbaxy, Mylan, and Barr. In total, Cephalon committed to pay the first filers up to $181 million. Cephalon’s payments to Teva and Ranbaxy took the form of licenses for intellectual property, plus manufacturing and supply agreements. Its payments to Mylan took the form of a development agreement for new products. Its payments to Barr took the form of a supply agreement. (Cephalon also granted Barr the possibility of earlier entry on another product, Actiq.) All four firms agreed to enter in October 2011, or six months later if Cephalon secured a six-month pediatric extension, prior to patent expiration in October 2014 (or April 2015 with pediatric extension). Cephalon’s CEO stated that
thanks to the agreements, “we were able to get six more years of patent protection. That’s $4 billion in sales that no one had expected.”

**Agreement**: August 2006, with later filer Carlsbad (in conjunction with Watson). The agreement provides for entry in October 2011 or, in the event Cephalon secures pediatric exclusivity, April 2012. These are the same entry terms as the first filers, but assuming one or more first filers enjoy exclusivity upon entry, in practical terms Carlsbad’s entry is delayed by 180 days.


16. Altace

Ramipril, marketed by King Pharmaceuticals under the brand name Altace, is an acute coronary event (ACE) inhibitor used to treat hypertension. Sales in 2006 were $653 million.

**Agreement**: February 2006, with first filer Cobalt. King agreed to pay a Cobalt affiliate $110 million: $35 million upfront, plus $25 million during the last quarter of 2006 and the first two quarters of 2007. These payments were described as compensation for the development of new Altace formulations. In addition, Cobalt received the right to sell an authorized generic version at an undisclosed date. Cobalt
agreed to provide 30 days notice prior to launching a capsule product under its ANDA. Under the agreement, Cobalt remained free as a legal matter to enter the market.


17. Propecia

Finasteride in a 1-milligram dosage, marketed by Merck under the brand name Propecia, is a treatment for male hair loss. (Finasteride in a 5-milligram dosage is marketed under the brand-name Proscar as a treatment for prostate enlargement.) Sales in 2006 were $145 million.
**Agreement:** May 2006, with first filer Dr. Reddy’s. Merck made Dr. Reddy’s an authorized generic seller of Procard and Zocor, an agreement described as having been reached in January 2006. Dr. Reddy’s agreed to enter in January 2013, less than a year prior to the last-expiring patent.

**Sources:** Legal: Letter from Mary Graham to Judge Gregory M. Sleet, Merck & Co. v. Dr. Reddy’s Laboratories, No. 04-1313 (D. Del. Mar. 1, 2006); Stipulation of Dismissal, Merck & Co. v. Dr. Reddy’s Laboratories, No. 04-1313 (D. Del. May 18, 2006); Letter from Gary Buehler, Dir. of Office of Generic Drugs, FDA, to Kumara Sekar, Dir. of Global Regulatory Affairs and Compliance, Dr. Reddy’s Labs., Inc. (June 28, 2006). Other: Merck, Other Financial Disclosures, Fourth Quarter (Jan. 30, 2007); Merck, Annual Report (Form 10-K) (Feb. 28, 2007); Press Release, Dr. Reddy’s, Dr. Reddy’s Launches Authorized Generic Versions of Proscar and Zocor (June 23, 2006).

18. **Plavix**

Clopidogrel bisulfate, marketed by Bristol-Myers Squibb under the brand name Plavix, is an anti-platelet drug used to prevent blood clots. Sales in 2005 were $3.2 billion.

**Agreement:** March 2006, with first filer Apotex. Bristol granted Apotex certain launch protections if, as discussed below, regulators rejected the agreement: a contractual delay in Bristol’s ability to seek a preliminary injunction if Apotex launched without a district court ruling, and a cap on Apotex’s damages at a much reduced level. On the other hand, if regulators permitted the settlement, Apotex agreed to enter in
March 2011, a few months prior to patent expiration in November 2011, with a six-month delay in entry if Bristol secured a pediatric extension.

For the entry delay to take effect, the settlement agreement required approval by the FTC and state attorneys-general, under the terms of an earlier consent decree meant to address prior alleged anticompetitive activity by Bristol. Regulators rejected the March 2006 agreement, which led the parties to reach a revised agreement in May 2006. The May version was rejected too, and in August 2006, Apotex launched despite the lack of a district court ruling. Apotex quickly flooded the market with more than a billion dollars of generic clopidogrel before a district judge halted further sales pending resolution of the patent suit.

The agreements had several additional terms, including a payment described as compensation for the generic firm’s inventory. The March agreement included a breakup fee, payable to the generic firm if the agreement failed to receive regulatory approval, which increased with the length of delay in receiving a response from regulators. The May agreement omits mention of a breakup fee, but Apotex alleged that the fee remained an unwritten term of the deal that its bargaining partner failed to report to regulators. Apotex also alleged that there was a second unwritten term, a commitment by Bristol not to launch an authorized generic. Bristol later pled guilty to criminal charges that it had misled regulators about the content of the second agreement.

19. Adderall XR

Adderall XR is the brand name for a mixture of amphetamine salts in an extended release form, marketed by Shire and prescribed to treat attention deficit hyperactivity disorder. Worldwide sales in 2006 were $835 million.

Agreement: August 2006, with first filer Barr. Shire agreed to pay Barr a net amount of up to $102 million. The agreement had several components. Shire agreed to pay a Barr subsidiary up to $165 million—$25 million immediately and up to $140 million over eight years—described as compensation for product development and marketing a Barr contraceptive, Seasonique, in certain territories. The product development pertained to six proprietary Barr women’s health products. In connection with the development deal, there was also provision for Barr to supply product to Shire, in exchange for additional compensation. The marketing and promotion rights included five countries in Western Europe. Barr agreed to pay Shire $63 million, described as
compensation for transferring Shire’s rights to Adderall IR (immediate-release) tablets. Barr agreed to enter in August 2009, prior to expiration of the last-to-expire patent. Shire also agreed not to compete with Barr through an authorized generic.

**Agreement:** January 2006, with later filer Impax. Impax agreed to enter in January 2010, or earlier under certain circumstances. The terms included a provision that Impax can launch 181 days after Barr’s negotiated launch, making the entry date October 2009. The parties agreed that Impax would receive $40 million over three years, described as compensation for promoting Carbatrol, a Shire epilepsy drug, starting in July 2006. Impax is entitled to additional incentive payments based upon prescription growth.

**Additional agreements:** March 2008, with later filer Teva (entry 181 days after Barr’s entry); April 2008, with later filer Colony (181 days after Barr’s entry).

20. AndroGel

Testosterone gel 1%, marketed by Solvay under the brand name AndroGel, is a daily testosterone replacement therapy. Sales in 2005 were $330 million.

Agreement: September 2006, with first filer Watson. Solvay agreed to make annual payments projected to range from $15 million to $30 million, described as compensation for promotion of AndroGel to urologists. Watson agreed to enter in August 2015, prior to patent expiration.

Agreement: September 2006, with later filer Par. Solvay agreed to pay Par and a contracting partner $8 million annually, described as compensation for co-promotion and a “backup” manufacturing deal. Par agreed to enter in February 2016, prior to patent expiration.

21. **Wellbutrin XL**

Bupropion hydrochloride in an extended-release form, marketed by GlaxoSmithKline under the brand name Wellbutrin XL, is an antidepressant. Sales in 2006 were $1.47 billion.

**Agreement:** In March 2007, with first filer Anchen (plus Anchen’s partners, Impax and Teva, treated here as though they’re a single entity). By the time of settlement, Anchen had won summary judgment, received FDA approval, and launched as to the 300-milligram strength. It did not launch as to a second strength, 150 milligrams, of roughly equal economic importance. Under the agreement, Anchen received, as to the 300-milligram product, a damages waiver for past sales and a license for future sales. (There was also a provision for Biovail to supply product.) For the 150-milligram strength, Anchen agreed to enter in 2008, prior to patent expiration in 2018, albeit a patent that had already been judged by a district court to be not infringed by the generic product. Glaxo agreed not to launch an authorized generic product.

**Sources:** Legal: Biovail Laboratories v. Anchen Pharmaceuticals, No. 04-1468 (C.D. Ca. Aug. 1, 2006); Letter from Gary Buehler, Dir. of Office of Generic Drugs, FDA, to Margaret L. Choy, Vice President of Regulatory Affairs, Anchen Pharms., Inc. (Dec. 14, 2006); Letter from Gary Buehler, Dir. of Office of Generic Drugs, FDA, to Mark C. Shaw, Vice President of Regulatory Affairs and Compliance, Impax Labs., Inc. (Dec. 15, 2006). News: Teva Launches New Strength of Wellbutrin XL, Generic Line,
June 11, 2008; Teva Announces Launch of Generic Wellbutrin XL Tablets, 300 mg
Under Agreement with Anchen and IMPAX, Bus. Wire, Dec. 18, 2006; Scott Anderson,
(noting FDA approval of Anchen 150-mg and 300-mg products); Gustav Ando, Teva
Launches Generic Wellbutrin, Global Insight Daily Analysis, Dec. 19, 2006. Other:
GlaxoSmithKline, Income Statement for Year Ended 31st December 2006; Press
Release, Teva Announces Agreement With Biovail And Anchen Regarding Generic
Wellbutrin XL (Mar. 5, 2007); Press Release, Biovail Labs., Biovail Comments on
Wellbutrin XL Court Ruling (Aug. 2, 2006); Press Release, Impax Labs., Inc., IMPAX
Announces FDA Approval of Generic Wellbutrin XL 300 mg (Dec. 18, 2006); Press
Release, Impax Labs., Inc., IMPAX Settles Patent Litigation Related to Generic
Wellbutrin XL (Mar. 5, 2007); Press Release, Teva Pharm., Teva Announces Launch of
Generic Wellbutrin XL Tablets, 300 Mg under Agreement with Anchen and IMPAX
(Dec. 18, 2006); Press Release, IMPAX Laboratories, Inc., IMPAX Announces FDA
Approval of Generic Wellbutrin XL 300 mg (Dec. 18, 2006); Press Release, Biovail
Labs., Biovail Announce Comprehensive Settlement Related to Wellbutrin XL (Mar.
5, 2007); Biovail Labs., Annual Report (Form 20-F) (Mar. 22, 2007); Press Release,
Teva Announces Launch of Generic Wellbutrin XL Tablets, 150 mg (June 2, 2008).

22. Nexium

Esomeprazole magnesium, marketed by AstraZeneca under the brand name
Nexium, is a treatment for excessive stomach acid. Sales in 2007 were $3.4 billion.
**Agreement**: April 2008, with first filer Ranbaxy. AstraZeneca agreed to buy active ingredient from Ranbaxy starting to May 2009, and to take Nexium production from Ranbaxy starting in May 2010. AstraZeneca also made Ranbaxy its authorized generic distributor of 40-milligram omeprazole (generic Prilosec) and felodipine (generic Plendil). Ranbaxy agreed to enter in May 2014, apparently with no authorized generic. At the time of settlement, Ranbaxy estimated a revenue boost from the settlements of between $1.25 and $1.5 billion, between 2008 and 2014.


**23. Lipitor**

Atorvastatin, marketed by Pfizer as Lipitor, is a treatment for high cholesterol. Sales in 2007 were $7.2 billion.

**Agreement**: June 2008, with first filer Ranbaxy. The terms are complex and difficult to discern using public information. The agreement settled Lipitor litigation in seven countries, with varying entry dates, and “resolved” disputes in four others. The
parties also settled litigation on Caduet, a combination of Lipitor and a second drug; resolved a lingering dispute involving Accupril, in which Ranbaxy had earlier launched at risk before being enjoined and likely faced significant damages liability; and resolved litigation in Ecuador regarding Viagra. As for Lipitor, Ranbaxy agreed to enter in November 2011, after the expiration of the two patents mainly at issue, and before the expiration of two additional patents, later put at issue and not listed in the Orange Book, that expire in 2016.


Barr Labs., Inc. 2001. Amendment to a Previously Filed 10-K405 (Form 10-K405/A), May 15.


Downey, Bruce. 2007. Paying Off Generics to Prevent Competition with Brand Name Drugs: Should It Be Prohibited? Hearing Before the Senate Committee on the Judiciary, 110th Congress.


______. 2003. Examining the Senate and House Versions of the “Greater Access to Affordable Pharmaceuticals Act” Before the Senate Committee on the Judiciary, 108th Congress.


Pfizer, Inc. 2001. Motion to Dismiss the Complaint, Great Lakes Health Plan, Inc. v. Pfizer, Inc., No. 01-106. Northern District of West Virginia, July 30.


Cases Cited

Andrx Pharmaceuticals, Inc. v. Elan Corp., No. 00-3481 (S.D. Fl. Apr. 24, 2003)


Aventis Pharma Deutschland, GmbH v. Lupin, Ltd., 499 F.3d 1293 (Fed. Cir. 2007)

Bement v. Nat’l Harrow Co., 186 U.S. 70 (1902)

Berkey Photo, Inc. v. Eastman Kodak Co., 603 F.2d 263 (2d Cir. 1979)


Boise Cascade Corp. v. FTC, 637 F.2d 573 (9th Cir. 1980)


Ethyl Corp. v. FTC, 729 F.2d 128 (2d Cir. 1984)


In re Bristol-Myers Squibb Co. (Teva Pharmaceuticals USA, Inc.), No. C-4076, FTC (May 24, 2004)


In re Tamoxifen Citrate Antitrust Litigation, 466 F.3d 187 (2d Cir. 2006)


National Petroleum Refiners Association v. FTC, 482 F.2d 672 (D.C. Cir. 1973)


Pfizer, Inc. v. Ranbaxy Laboratories, 457 F.3d 1284 (Fed. Cir. 2006)

Pfizer, Inc. v. Teva Pharmaceuticals USA, Inc., 429 F.3d 1364 (Fed. Cir. 2005)


Simpson v. Union Oil Co., 377 U.S. 13 (1964)

SmithKline Beecham Corp. v. Apotex Corp., 247 F. Supp. 2d 1011 (N.D. Ill. 2003), affirmed, 365 F.3d 1306 (Fed. Cir. 2004), vacated on rehearing en banc, 403 F.3d 1328 (Fed. Cir. 2005), affirmed on other grounds, 403 F.3d 1331 (Fed. Cir. 2005)

Teva Pharmaceutical Industries Ltd. v. Crawford, 410 F.3d 51 (D.C. Cir. 2005)

United States v. General Electric, 272 U.S. 476 (1926)

United States v. Masonite Corp., 316 U.S. 265 (1942)

United States v. Microsoft, 253 F.3d 34 (D.C. Cir. 2001) (en banc) (per curiam)


Valley Drug Co. v. Geneva Pharmaceuticals, Inc., 344 F.3d 1294 (11th Cir. 2003)