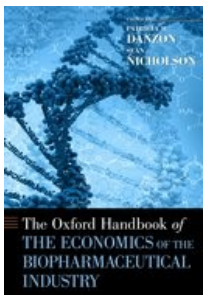

The Oxford Handbook of the Economics of the Biopharmaceutical Industry

Patricia M. Danzon (ed.), Sean Nicholson (ed.)

https://doi.org/10.1093/oxfordhb/9780199742998.001.0001

Published: 2012 **Online ISBN:** 9780199968749 **Print ISBN:** 9780199742998

CHAPTER

5 The Regulation of Medical Products

Anup Malani, Tomas J. Philipson

https://doi.org/10.1093/oxfordhb/9780199742998.013.0005 Pages 100–142

Published: 18 September 2012

Abstract

This article synthesizes and extends, in a nontechnical manner, recent research on the Food and Drug Administration (FDA). The aim is to shed light on whether the policies of the agency itself are safe and effective when measured in terms of economic efficiency. The first section provides an overview of the role of the FDA in regulating pharmaceutical drugs and medical devices. The second section surveys the existing efficiency rationales for government regulation of the information about and the quality of medical products, and then canvasses the literature for empirical studies on the effects of FDA regulation on innovation and costs. The final section examines the growing role of tort law—specifically, products liability litigation—in supplementing FDA regulation of drug quality.

Keywords: Food and Drug Administration, drug policy, economic efficiency, FDA regulation, medical device regulation, tort law, product liability, drug quality

Subject: Health, Education, and Welfare, Public Economics, Economics

Series: Oxford Handbooks

Collection: Oxford Handbooks Online

Improvements in health have been a major component of overall gain in economic welfare in the last century (Cutler and Richardson 1998; Murphy and Topel 2006). Part of these health gains are attributable to medical research (Cutler and McClellan 2001; Murphy and Topel 2003; Lichtenberg 2003; Cutler et al. 2007). At the same time, the cost of health care has tripled in real terms since 1965 and now accounts for more than 17 percent of the gross domestic product (GDP). Evidence exists to suggest that a large share of this cost growth is driven by technological innovation (Newhouse 1992), including the cost of medical products such as drugs. The large role of innovation in explaining both improvements in health and health care cost growth suggests that it is important to understand the medical research and development (R&D) process and how it is regulated.

In virtually all developed countries and many developing countries, the government provides regulatory oversight over the quality of products generated by medical innovation. In the United States, this oversight is conducted by the US Food and Drug Administration (FDA), which regulates drugs, medical devices, biologics (products made from living organisms, such as vaccines and blood products), cosmetics, radiation-emitting electronic products, veterinary products, and foods. According to the FDA, the products it regulates account for more than one-fifth of US consumer spending. In the area of medical products, the agency reviews whether drugs and many devices are safe and effective both before and after they have been cleared for sale.

The manner in which the FDA regulates the quality of drugs and devices has a substantial impact on the cost
p. 101 of their development. The FDA requires that ↳ companies conduct clinical trials to demonstrate that their medical products are safe and effective. These trials account for a large portion of the total development costs of these products (DiMasi et al. 2003; Adams and Brantner 2006). In addition, completion of trials does not guarantee that a product will be approved. This risk of nonapproval compounds the cost of product development (DiMasi et al. 2003).

Despite the central role of the FDA in regulating the quality and R&D costs of medical products, economists have conducted relatively little theoretical or empirical research on the efficiency of FDA policies. Ironically, if a product application were presented to the FDA with the scant amount of evidence that currently exists on the efficiency of the policies of the agency itself, such an application would likely be rejected on the basis of insufficient evidence. In this chapter, we synthesize and extend, in a nontechnical manner, recent research on the FDA. Our aim is to shed light on whether the policies of the agency itself are safe and effective when measured in terms of economic efficiency.

The first section provides an overview of the role of the FDA in regulating pharmaceutical drugs and medical devices. The second section surveys the existing efficiency rationales for government regulation of the information about and the quality of medical products and then canvasses the literature for empirical studies on the effects of FDA regulation on innovation and costs. The final section examines the growing role of tort law—specifically, products liability litigation—in supplementing FDA regulation of drug quality.

To understand the relationship of this chapter to others in this book, it is helpful to break government influence on medical product innovation into five parts. The first is the use of property rights over innovation to encourage investment in R&D. The second is direct government spending on R&D. The third is premarketing screening of new drugs and devices by the FDA. The fourth is postmarketing regulation of medical products by courts (and perhaps the FDA). The last is demand for medical products by government-run health insurance systems such as Medicare and Medicaid. This chapter focuses on the third and fourth categories; that is, the role of quality and safety regulation in directing innovation.

FDA Regulation of Drugs, Devices, and Biologics

Drugs

The US government began regulating drugs with the 1906 Food and Drug Act. That statute prohibited
p. 102 companies from selling misbranded or adulterated ↳ drugs. Using a modern example, aspirin would be misbranded if its label stated that it was ibuprofen. A drug is adulterated if it is tainted with a poison or some other dangerous substance. The 1906 Act was targeted mainly at quack patent medicines and regulated drugs largely after they had already been sold to consumers.

The United States government did not screen drugs before permitting them to be sold to consumers until the enactment of the Federal Food, Drug, and Cosmetic Act (FD&C Act) in 1938. That statute required that

drugs be proved safe before they could be sold to consumers. Companies did not have to prove drugs were effective, though the FDA could prevent marketing of a drug if the agency could demonstrate it did not work. In practice, however, there was no premarket screening of drugs for efficacy. In 1962, following a public health scare in which a number of children were born with serious birth defects after their mothers used thalidomide to reduce morning sickness during pregnancy, the US Congress amended the FD&C Act to require that a company provide substantial evidence that its new drug is both safe and effective before it could sell that drug (FD&C Act, §§505(a) and (d)).[1]

Since 1962, drug development can be broken down into two parts. The first is in vitro and animal testing. If this proves successful, a company may apply to the FDA for permission to conduct clinical trials on humans. The application is called an Investigational New Drug (IND) application (FD&C Act, §505(i)). The FDA uses the power to approve human testing to regulate the nature of trials conducted by a company. The FDA works with companies to determine appropriate endpoints for clinical evaluation and the sample size required for trials. It may also provide guidance on the level of efficacy or safety that would help clear the way for drug approval.

Clinical testing in turn has three phases. Phase 1 involves studies to determine the toxicology of a drug. A small number of healthy humans[2] are gradually given increasing doses of a drug to determine how toxic it is. If the dosage can be raised to pharmacologically effective doses with tolerable side effects, the drug passes phase 1. Phase 2 involves medium-sized studies to identify early evidence that the drug is effective in humans. These studies often do not have a control group. If the drug demonstrates efficacy relative to historical controls or the expectations of the company and the FDA, the drug passes phase 2. In phase 3, the
p. 103 company must typically conduct two large, well-controlled trials of a drug's safety and efficacy. ↳ The trials are typically randomized and blinded and use placebos or conventional treatments as controls.

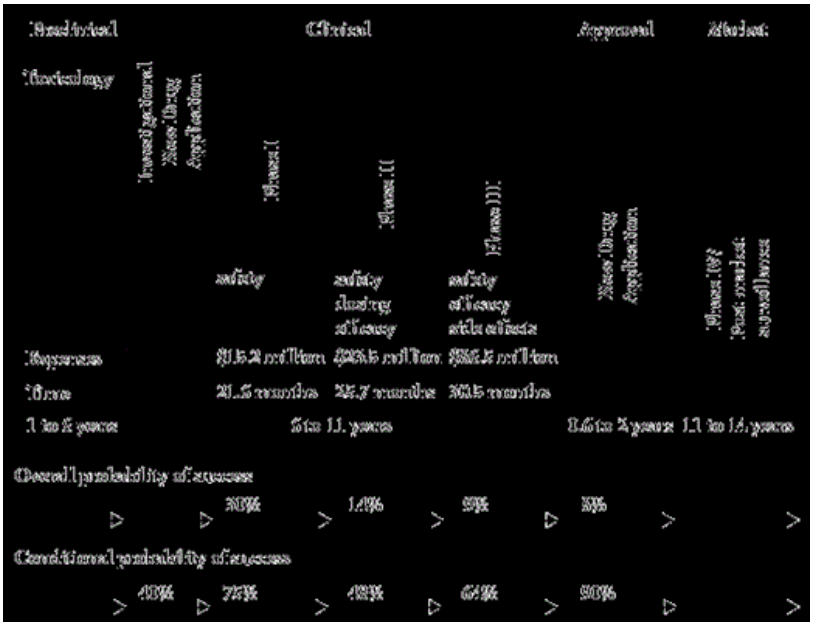

Figure 5.1 **Overview of the drug development process**.

Source: From Philipson and Sun, 2008.

If a new drug passes phase 3 with significant evidence of efficacy and a reasonable safety profile, the company may file a New Drug Application (NDA) to obtain approval from the FDA to market the drug. The FDA reviews the data to ensure that there is substantial evidence of efficacy and that the drug has been demonstrated to be safe. In practice, the FDA balances the value of efficacy from the drug against the cost in

terms of side effects when deciding to approve a drug. For example, the agency may tolerate greater side effects from a drug that treats a more serious medical condition or a drug that dramatically improves health outcomes. The overall review process, from preclinical testing to NDA approval, is illustrated in Figure 5.1, which is reproduced from Philipson and Sun (2008).

Drugs are approved for particular uses. This does not prevent doctors from prescribing an approved drug for alternative uses. Such "off-label" uses are quite common, for example, in oncology. Use-specific approval does, however, prevent companies from advertising unapproved uses to doctors and patients.[3] If a company seeks to market a drug for as yet unapproved uses, it must conduct studies to demonstrate that the drug is safe and effective for those uses.

p. 104 The incomplete regulation of off-label use highlights an important limitation in the federal regime for drug regulation. A drug need only be approved for one use to be available for almost any use that doctors may pursue. This affects a drug company's decision about the specific use for which to seek approval from the FDA. A company will consider the use for which proving safety and efficacy is easiest as well as the use for which marketing is most important. It also affects the value of the FDA's regulation. If only a fraction of uses are screened by the FDA, the drug market is only partially regulated by the FDA.[4]

The FD&C Act does not merely require a drug to be safe and effective; it also requires that the drug be labeled so that a doctor can prescribe the drug to patients in a manner that will ensure proper sorting of drugs to patients. The FDA screens a new drug's label before the drug is approved for marketing and may require that the label be amended as the agency receives new information about a drug after approval. The label must describe not only indications for use but also *inter alia* contraindications and side effects.

Clinical testing does not end once a drug has been approved for marketing. Although the FDA does not have the formal authority to require clinical testing after a drug is approved, it may hold up drug approval unless the drug company "voluntarily" agrees to conduct so-called phase 4 clinical trials. Although some of these phase 4 obligations involve placebo-controlled trials, many require long-term and large-scale observational studies. The purpose of these studies is to identify side effects that are serious but so rare that even phase 3 trials are not sufficiently large to detect them.

One problem identified soon after passage of the 1962 Amendments to the FD&C Act was that companies stopped developing and fully testing drugs for rare conditions. The financial burden of clinical testing exceeded the small expected revenue from these drugs. Congress addressed this problem in 1983 with the Orphan Drug Act. That statute provided companies that developed drugs for rare conditions, defined as diseases that affect fewer than 200,000 people annually, tax credits to offset their clinical testing costs and seven years of market exclusivity (FD&C Act §§526, 527).

A second set of problems with the 1962 Amendments was that they both reduced effective patent life for pioneer new drugs and hampered entry of generic versions of these drugs into the market. The long time needed for the FDA to review NDAs reduced effective patent life, defined as the number of remaining years a company has to market a drug under its patent monopoly following FDA approval. This in turn reduced the
p. 105 incentive to innovate. Once a patent expired, ↳ however, the Amendments required a generic drug company to perform the same clinical trials that were performed to gain approval for the pioneer drug. The low expected profits on generic drugs made such testing unprofitable. Although the second problem offset the first, it was not the how Congress intended the patent system and the FD&C Act to interact.

In 1984 Congress addressed both problems with the so-called Hatch-Waxman Act. This statute extended a drug company's patent on a pioneer drug to partly offset the time required for clinical testing and FDA review. The extension was capped at 5 years and could never extend effective patent life beyond 15 years.[5] In return, generic drug companies were not required to conduct trials demonstrating that their drugs were safe

and effective. Instead, they could market their drugs so long as they demonstrated that their drugs were bioequivalent to a pioneer drug whose patent had expired or had been invalidated (FD&C Act §505(j)).

Congress once more took aim at long FDA review times in 1992 with the Prescription Drug Users Fee Act (PDUFA). That Act required drug companies to pay a user fee to fund resources the FDA could use to speed up the drug approval process. In return, the Act required the FDA to commit to deadlines for completing its drug reviews. PDUFA was written to expire in five years, but it has repeatedly been renewed.

Devices

Medical device regulation has lagged behind drug regulation in the United States. Whereas the 1906 Food and Drug Act prohibited the misbranding or adulteration of drugs, devices did not have such restrictions until passage of the 1938 FD&C Act. Likewise, whereas the 1938 Act and then the 1962 Amendments introduced premarket screening of drugs, there was no premarket screening of devices until 1976.

The 1976 Medical Device Amendments introduced three tiers of regulation that finally provided for premarket screening of certain devices. This new regime first sorted devices into three classes and then prescribed different levels of regulation for the difference classes (FD&C Act §513). Class I devices are subject only to prohibitions on misbranding and adulteration. Class II devices are subject to these prohibitions and also to certain performance standards developed by the FDA. For example, for a device to be labeled an x-ray machine it must meet certain guidelines for the quality of its beam and the amount of
p. 106 radiation that the ↳ machine may leak (21 CFR 1020.30). Finally, class III devices are regulated much like new drugs. Manufacturers must provide data that demonstrate both safety and effectiveness. The Amendments sort devices into a class based largely on whether the FDA believes that the level of regulation implied by that class is sufficient to reasonably ensure that the device is safe and effective for consumers. Roughly 30 percent of devices fall in class I, 60 percent in class II, and 10 percent in class III (Hutt et al. 2007).

Since the 1976 Amendments, the development of device regulation has largely tracked that of drug regulation. The Orphan Drug Act and the Hatch-Waxman Act included provisions for devices. And in 2002, Congress extended PDUFA to apply to medical devices.

Evaluation of FDA Regulation

Rationale for Government Regulation

FDA regulation of medical products has two basic components. The first is production and dissemination of information and analysis about the quality of drugs and devices. This is embodied in the FDA's requirements that companies produce data to demonstrate safety and efficacy and that drug and device labels report indications as well as contraindications and adverse effects. The second is a minimum quality regulation. This is directly embodied in the FDA's rule that drugs and certain devices cannot be sold until they have been demonstrated to be safe and effective. It is indirectly implemented in the FDA's rule that class II devices must meet certain performance standards.

Production of Information

In order to justify the government provision of information, one must explain why markets do not produce enough information and why the government should step in to do so. Perhaps the production of information is a public good and free-riding causes inefficiently low levels of information to be produced (Musgrave 1959). Alternatively, the government may be more efficient at providing information. Perhaps the government has more expertise, or perhaps patients (and doctors they might hire to help them) have cognitive limitations.[6] Or perhaps there are ↳ economies of scale in data analysis. It may be more efficient
p. 107 for one entity to analyze the data than for each patient to do it on her own.

Even if one concludes that the government should provide information about a product, there remain important questions about which party—the medical product company, a third private party, or the government—is best suited to produce that information and who should pay for the production of that information. The producer of the medical product may have private information about the product that places it in the best position to test the product. For example, it knows which patient group is most likely to benefit. The drug producer and its consumers are also entities who gain from information that the product is safe and effective.

An alternative to testing by the producer is testing by a third private party or the government. One concern with permitting medical product companies to test their own products is that their studies may be biased in favor of their products. Studies have documented that clinical trials conducted by authors with financial conflicts from drug companies tend to report larger treatment effects than studies without such conflicts (Bekelman et al. 2003; Lexchin et al. 2003). However, government–run studies may take longer to complete, and this is a problem because delay is costly to producers, who lose monopoly profits under their patents, but not to the government. Another problem is that government studies may have a bias toward finding that drugs have side effects. Indeed, the FDA has been criticized for being overly cautious with drugs, because although the agency is not given credit for new effective drugs, it is blamed when it permits an unsafe drug to enter the market. (To be fair, the FDA has also been criticized for being captured by the drug industry.)

The question of who should fund the FDA's mission to produce information is in part a question about efficiency and in part a question about equity. There is a large optimal taxation literature that discusses which sorts of taxes are able to raise a given quantum of revenue with minimum distortions. The lessons there apply to the FDA context. However, one consideration that favors taxing medical product companies to fund clinical research is that doing so may bypass the administrative costs of collecting taxes from third parties. Questions about equity are perhaps beyond the ken of economists. But one factor implying that a company and its consumers should not be disproportionately taxed to pay for research on the company's product is that information suggesting that a product is not safe or effective benefits individuals who do not consume the company's product and perhaps the company's competitors. This favors a broader tax, although perhaps not so broad as to extend to individuals who are not in the market for a medical product.

Whereas efficient and fair production of information is important, the central concern with the government's decision to produce information is whether the benefits of that information are worth the
p. 108 costs.[7] In chapter 2 of this volume, ↳ DiMasi and Grabowski document the large and growing costs of clinical testing. According to DiMasi et al. (2003), a commonly cited study, the cost of completing three phases of clinical testing is roughly $200 million. Of course many drugs are abandoned before testing is completed. Typically only one in five completes testing and is approved by the FDA for sale. If the cost of testing is normalized by the number of successful drugs, one obtains estimates of clinical testing costs in the range of $800 million per approved drug.

Whichever estimate is used, this cost must be balanced against the benefits of clinical testing. Note that this is not the same question as whether the revenue from a medical product exceeds the cost of clinical testing.

The revenue reflects both the value of the medical product and the value of information about the product. To the extent that doctors would use the medical product without the information, the revenue does not reflect the value of information. If, however, doctors would use the product on a different population of patients when provided with the information, then the additional benefit to this new population of users minus the avoided costs to the old users is a benefit of information.

In an early, seminal paper on the value of FDA regulation, Peltzman (1973) attempted to capture some of these ideas in a graph along the lines of that in Figure 5.2, panels A and B. Suppose that a typical consumer's demand for a drug before the FDA provides any information on the drug is represented by the line DD. The price of the drug is *p*, so the equilibrium quantity is *q*.[8] (These initial demand and supply curves are drawn in bold to distinguish them.) The graph does not capture all facets of the drug market, specifically the fact that there may be a monopoly producer that charges a price higher than marginal cost and that changes price in response to demand shifts. But it is a useful simplification that makes the following important point: because the consumer may not have accurate information about the value of the drug, we cannot conclude from the graph that the surplus is the triangular area under DD but above p.[9]

After the FDA provides the consumer information, demand might shift either in (panel A) or out (panel B) to line D′D′ (Peltzman 1973; McGuire et al. 1975). A shift inward could occur if the FDA demonstrated that the drug is less safe and effective than previously thought. A shift outward could reflect either an FDA
p. 109 assessment that ↳ the drug is more safe or effective than previously thought,[10] patients' increased confidence in estimates about a drug's value, or (looking past our model of a representative patient) a better sorting of patients to drugs. These shifted demand curves can be used to calculate welfare because they reflect "true" information. Whichever way demand shifts, the price of the drug may rise to p' because of the cost of either producing information or conveying that information to consumers.[11]

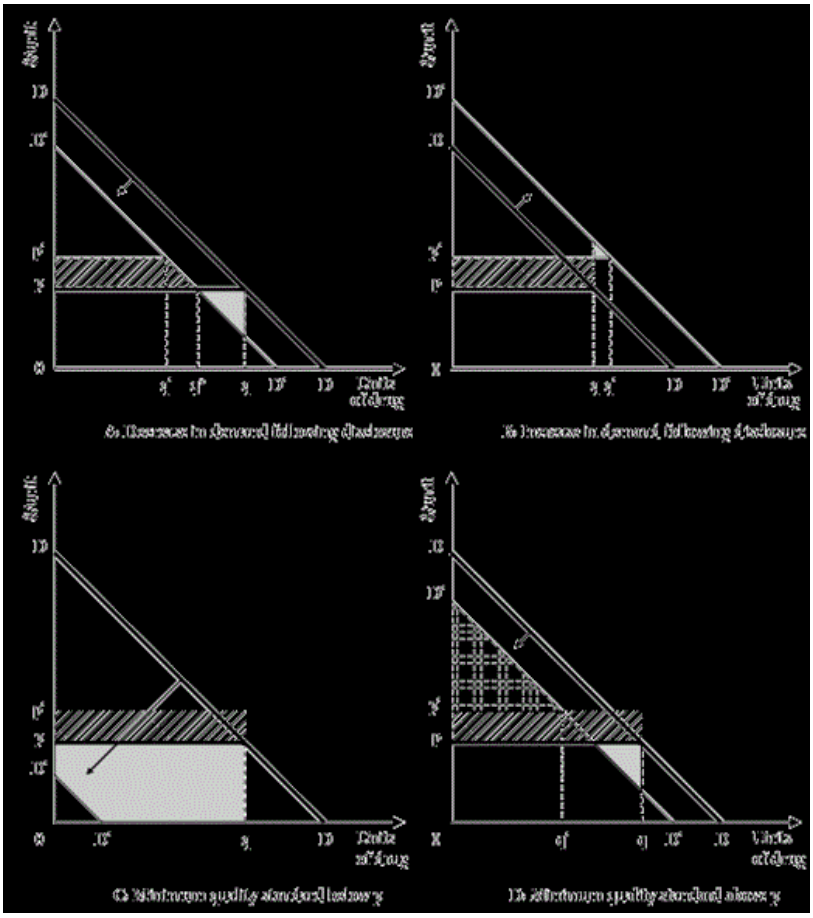

Figure 5.2 **Welfare effect from production of information and minimum quality standards**.

p. 110 If demand shifts inward to D′D′ as in panel A, the consumer's surplus before production of information ("preregulation") has two parts. The first part is the area under D′D′ and above p. This triangle represents the "true" surplus from consumption up to quantity q*. The second part is actually a loss, equivalent to the gray triangle. This is consumption between q^* and q that costs more than it is truly worth to the consumer. After the production of information ("postregulation"), the consumer's surplus is the area under D′D′ but above p'. The higher price reflects the cost of producing information. The net gain from the production of information is the gray-shaded triangle minus the line-shaded area. The former represents the avoidance of excessively costly consumption, and the latter reflects the cost of information. The net gain could be positive and thus increase welfare or negative and thus decrease it.

If, instead, demand shifts out to D′D′ as in panel B after the production of information, the preregulation surplus is the area under D′D′ and above p, but it is bounded on the right by the quantity q. This area represents the true surplus from preregulation consumption of quantity q. The postregulation surplus is the area under D′D′ and above p'. This is consumer surplus at the elevated level of demand after accounting for higher prices due to the cost of producing information. The net gain from regulation is the gray-shaded triangle minus the line-shaded rectangle. The former is extra surplus extracted due to greater quantity of consumption. The latter is the cost of information.

Minimum Quality Standards

An alternative view of FDA regulation is that it sets minimum standards for the quality of medical products. The minimum threshold for quality depends on the specific evidentiary standard imposed by the FDA. Because the FDA requires that the randomized controlled trials show positive average treatment effects and "reasonable" side effects relative to control, the minimum threshold bars a product that is not effective and safe for the average patient enrolled in those trials.

There are two basic ways to justify the government's imposition of these minimum quality standards on consumers who would use their own money to pay for products. First, even if the government provides information that a product is not safe or effective, a consumer may still consume it because there is a cost to processing information from the government. This lost value need not be bounded by the individual's cost of processing information. Because consumers do not know the value of the information from the government, they might not pay the cost of processing it even if, in hindsight, the value of processing that information would exceed the cost of doing so. A second justification for minimum quality standards is that individuals are unaware of their best interests and—importantly—are unaware that they are unaware. Therefore simple paternalism may justify FDA screening.

If minimum quality standards are justified on efficiency grounds as opposed to paternalistic grounds, we
p. 111 can evaluate these standards using the same graphs ↳ we used to illustrate the value of information from the FDA.[12] If all patients are homogeneous—so that they have the same demand for a product—and the FDA sets minimum quality standards such that any product that generates zero or negative consumer surplus is barred from sale, then FDA regulation simply bars all medical products whose true demand curves lie at or below p in panel C of Figure 5.2. Because these products generate no surplus, the minimum quality standard avoids waste from consumption *q*. For a true demand curve such as D′D′, the savings are equal to the gray-shaded area. We must, of course, subtract from this gain the cost of producing information. That is difficult to graph because there are no consumers to purchase the product and thus pay for the cost of information. However, if the cost of information would have raised the price of the product to p′ given the old quantity of consumption q, then the line-shaded area represents the cost of providing information. The net gain is the gray-shaded area minus this line-shaded cost.[13]

If the FDA's minimum quality standards are more restrictive, or if patients are heterogeneous and some benefit more than p from the product, then the FDA might bar from sale products for which there is positive demand even at price *p*, as illustrated in Figure 5.2, panel D. In this case, regulation has an additional cost equal to the loss of consumer surplus under true demand D′D′. Given the higher price of the product due to the production of information, this lost surplus is equal to the check-shaded triangle. Moreover, the savings achieved by avoidance of wasteful consumption falls to the smaller gray-shaded triangle. Assuming that the cost of producing information is indifferent to the quality standard chosen by the government, the cost of information remains equal to the line-shaded area from panel C. Thus the net gain is the gray-shaded triangle minus both the line-shaded rectangle and the check-shaded triangle. An important lesson from panels C and D is that welfare effects of minimum quality standards increase according to the extent to which information increases welfare (the net of the gray- and line-shaded areas) and decrease according to
p. 112 the stringency of the government's minimum quality standard (which may push D′D′ outward). ↳

Effects on Innovation

The discussion in the last two sections highlights the possible justifications for FDA regulation. Starting with early papers by Wardell (1973) and Peltzman (1973), however, critics have been concerned about the impact of FDA regulation on the expected profits of medical product companies and thus on their incentive to innovate.

If the information provided by the FDA suppressed expected demand for medical products or if the FDA barred lower-quality medical products from the market, producer surplus would fall, even if consumer welfare was improved. Moreover, FDA regulation would surely increase the cost of production by requiring the generation of extra information through costly clinical trials. Together these factors would reduce the return to, and thus the amount of, R&D investments intended to generate new drugs. This would reduce the resources available for R&D expenditures. If companies knew at the start which products would be of high quality and which would not, only R&D into low-quality products would be affected. However, this information is typically not available until companies conduct clinical trials. Therefore, the reduction in R&D might curtail both ultimately high- and low-quality innovation. Because producers do not capture the full surplus generated by drugs, even low-quality drugs, this indiscriminate reduction in innovation may reduce overall welfare.

Framework for Dynamic Cost-Benefit Analysis

Figure 5.2 provides a relatively static description of the welfare effects of FDA regulation. It ignores some important dynamic issues such as the impact of regulation on innovation. It also ignores the time required for the FDA to produce information about a medical product or to screen out low-quality products. In this section we extend a framework initially proposed in Philipson et al. (2008) to incorporate these effects into welfare evaluations of FDA policies.

The first step in describing the dynamic welfare effects of FDA regulation is calculating the present value of costs and benefits of the products being regulated. We begin with an individual product that has already been approved for sale by the FDA. Because medical products are mainly intellectual property, costs are largely incurred before benefits are realized. Let R be the fixed cost of developing the product. Suppose that compliance with FDA regulation adds α years to the time required to develop and begin marketing the product. This period α captures both the time required to do additional clinical testing that the company would not have done in the absence of FDA regulations and the time required for FDA review of the company's new drug or device application. For simplicity, we assume that each year of FDA review has a variable cost of r to the company. This might be attributed to the cost of clinical trials, to user fees, or to legal expenses associated with shepherding a product through the approval process. Putting these pieces together, the present value of costs from a medical product can be written $-R-\sum_{t=0}^{\alpha-1}\beta^t r=-R-F$ where α is a discount factor and F is the delay cost of FDA regulation.

p. 113 Turning to the benefits of the product, let w_t be the sum of consumer surplus S_t and producer surplus α_t for a given product during a year, t, in which it is marketed. Total welfare from the product may vary over time as marketing spreads news about the drug or the drug comes off patent. Finally, suppose that the medical product is withdrawn from the market after M years, either because it is overtaken by a superior drug or because consumer use has revealed serious side effects.[14] If the drug is never withdrawn, $M = \infty$. An important benefit of FDA regulation in this framework is to screen for side effects and thus, by selection, increase the amount of time (M) that an approved drug is on the market. With these assumptions, the present value of benefits from a medical product is $\beta^{\alpha}\sum_{t=0}^{M}\beta^t\pi_t+\beta^{\alpha}\sum_{t=0}^{M}\beta^t S_t=\Pi+S$. Total welfare is the sum of the costs and benefits: $W=-R-F+\Pi+S$

Next, we aggregate welfare from all approved products on the market. Let N be the number of new drugs introduced to the market and $G(R,\{w_t\})$ the joint distribution of development costs and surplus levels across products. Given N, the expected aggregate social welfare A can be written as $A=\int NW dG$[15]

So far, the focus has been on drugs that are approved by the FDA. Now we turn to drugs that do not make it onto the market. One reason for not reaching the market may be that the FDA does not approve a drug. We formalize this situation by assuming that the FDA approves a product only if it meets come condition—for example, that static surplus $\bar{s}$ exceeds some critical value, v.[16] Now more stringent regulation may be

described as an increase in v, which reduces the probability that $\bar{S} > v_1$ v. Another reason a drug may be missing from the market is that FDA regulation reduces innovation. To capture this effect, we let the number of new drugs, N, depend on producers' returns to innovation: $N = f(E_G[-R + F + \Pi])$, where $f' > 0$ (i.e., higher expected profits encourage more innovation). More stringent FDA approval may reduce a producer's returns, and thus innovation, by increasing delay costs (F) and the probability of approval.[17]

The second step in describing the dynamic welfare effects of FDA regulation is to conduct a discrete policy
p. 114 experiment. We can characterize minimum quality ↳ regulation by the FDA at baseline by a vector of parameters $(a\downarrow 0, v\downarrow 0, N\downarrow 0, M\downarrow 0)$. More stringent regulation can be described as a vector of parameters $(a\downarrow 1, v\downarrow 1, N\downarrow 1, M\downarrow 1)$, where $a_1 > a_0$ captures the information demands and delay from regulation, $v_1 > v_0$ captures a reduced probability of approval, $N_1 < N_0$ captures reduced innovation due to regulation, and $M_1 > M_0$ captures the value of regulation in extending the life of approved products. The aggregate welfare effect of stringent regulation is $A_1 - A_0 = \int N_1 W_1 dG - \int N_0 W_0 dG$

This difference can be broken down into additional costs and benefits, $A_1 - A_0 = C + B$ The costs are

$$C = \overbrace{(N_0 - N_1)}^{(i)} \overbrace{E[W \mid v_1, a_1, M_1]}^{(ii)} + \overbrace{N_0 \Pr[s > v_1]}^{(iii)} \{\beta^{M_1 - a_1} / (\beta(1-\beta))\}r$$
$$+ \overbrace{N_0(\beta^{M_1} - \beta^{M_0}) E\left(\sum_{t=1}^{M_1} w_t \mid v_1\right)}^{(iii)}$$
$$+ \overbrace{N_0 \beta^{M_0}\left[E\left(\sum_{t=1}^{M_1} w_t \mid v_1\right)\right] - E\left(\sum_{t=1}^{M_0} w_t \mid v_0\right)}^{(iv)}$$

The first component of costs (i) is the reduction in new drugs that are developed. The second component (ii) is the additional cost of developing and testing drugs that are approved.[18] The third component (iii) is the reduction in the present value of social surplus from approved drugs because their marketing is delayed. The final component of costs (iv) is the loss of consumer and producer surplus from drugs that are not approved.

The benefits of more stringent minimum quality regulation (B) are

$$B = \overbrace{[E(N_{0R} \mid v_0) - E(N_0 R \mid v_1)]}^{(i)} + \overbrace{N_0 \beta^{\alpha 0} E\left(\sum_{t=M_0}^{M_1} \beta^t S_t \mid v_1\right)}^{(ii)}$$

The first component of benefits (i) is the research expenditures avoided because of the reduction in innovation. The second component (ii) is the larger consumer surplus due to the additional time a drug may be marketed before it is withdrawn.

The policy experiment just described evaluates minimum quality regulation. It can be modified to account for the value of information generated by the FDA. Better information has two effects: it increases consumer
p. 115 surplus, $S_{1t} > S_{ot}$, and ↳ it concomitantly reduces the producer surplus, $\pi_{1t} < \pi_{0t}$, during each period an

approved drug is sold. The increase in per-period consumer surplus adds a third term to the benefit side of the ledger:

$$N_0\beta^{\alpha_0}\left[E\left(\sum^{M_0} S_{1t}\,|\,\nu_1\right)-E\left(\sum^{M_0} S_{0t}\,|\,\nu_0\right)\right].$$

The reduction in per-period producer surplus adds a corresponding term to the cost side of the ledger. In addition, it may increase the magnitude of the reduction in innovation due to FDA regulation, component (*i*) in the cost equation presented earlier.

Empirical Evaluation of FDA Regulation

Estimating the welfare effects illustrated in Figure 5.2 or in the dynamic framework from the previous subsection is no trivial matter. It requires observing changes in

- The number of chemical entities introduced
- Development costs
- Development and review times
- Withdrawal rates
- Demand and supply curves for drugs

in order to measure changes in consumer and producer surplus.[19] In this section we review the methodology used to estimate these parameters and estimates of these parameters from the prior literature.

There are two basic challenges to identifying how FDA regulation affects parameters such as new drug development. The first issue is how to "quantify" FDA regulation. Researchers have taken two basic approaches. One is to look at adoption of any premarket clearance regulation, such as the 1962 Amendments to the FD&C Act in the United States (e.g., Peltzman, 1973). This treatment is coded as a dummy variable and is set to "0" before 1962 and "1" after.[20] The other approach is to use the time it takes for the FDA to review an NDA as a proxy for regulation (Wiggins 1981; Jensen 1987; Berndt et al. 2005a, 2005b; Carpenter et al. 2008; Philipson et al. 2008). This average approval time has varied substantially over the years. In 1960, it
p. 116 was roughly 5 months. After passage of the 1962 Amendments, ↳ approval time rose dramatically, reaching 20 months in 1970. For most of the 1980s, it hovered between 30 to 35 months. Approval time declined substantially after the passage of PDUFA in 1992. By 1998, it was approximately 12 months, which is roughly where it stands today. This rise and fall in approval time is illustrated in Figure 5.3, which is reproduced from Olson (2004).

The second challenge is constructing a baseline against which to judge the effect of the FDA regulation. Because the FD&C Act is a national statute, researchers cannot use, for example, differences in outcomes across US states that regulate drugs and states that do not regulate drugs. This makes it difficult to separate effects of the statute from underlying time trends. Researchers have used two basic methods to overcome this problem. One is to assume a parametric structure for outcomes that would have occurred in the absence of the 1962 Amendments. This could be as simple as including a time trend in the regression, or it could involve something more elaborate. For example, Peltzman (1973) used pre-1962 data to estimate a model of

new drug introductions and then predicted baseline new drug introductions after implementation of the Amendments by inserting post-1962 data into his estimated model. When he plotted actual introductions of new chemical entities (NCEs) against his predicted introductions, the result was a striking plot that became popular among critics of the FDA. It is reproduced in Figure 5.4.

The other approach researchers have used to construct a baseline is to examine the development of drug markets in countries that are similar to the United States but either did not pass strict drug regulation in 1962 or take less time to review new drug applications. The primary candidate is the United Kingdom, which passed premarket clearance for safety in 1963 but did not require proof of efficacy before sale until 1971 (Grabowski et al. 1978). The United Kingdom also had shorter review times than the United States, at least until the passage of PDUFA in 2002. For example, in 1980, average total development time (including preclinical testing, clinical testing, and regulatory review times) was 145 months in the United States versus just 70 months in the United Kingdom (Thomas 1990).

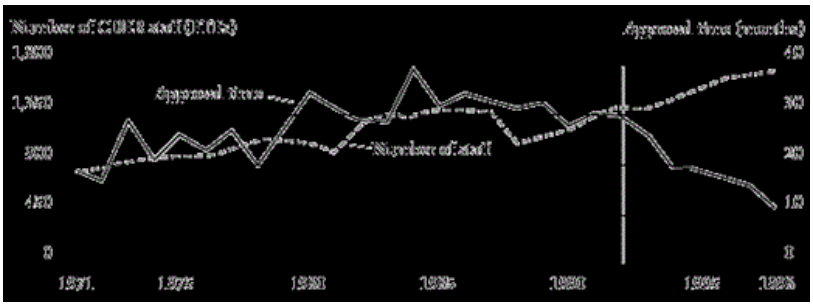

Figure 5.3 **FDA approval times, 1971–1998**.

Source: From Olson (2004).

p. 117 Tables 5.1 through 5.3 summarize the effects of FDA regulation on three important sets of parameters. One set focuses on innovation and includes outcomes such as new drug introductions and the productivity of R&D expenditure (i.e., new drug introductions/R&D expenditures). The second set examines drug development and FDA approval times. The third set considers the effect of FDA regulations on safety. The main outcomes are involuntary drug withdrawals. The tables report not only report findings but also the data employed, how FDA regulation is measured (e.g., 1962 dummy or review time), and how the counterfactual or baseline is constructed (e.g., parametric time trend or international comparison).

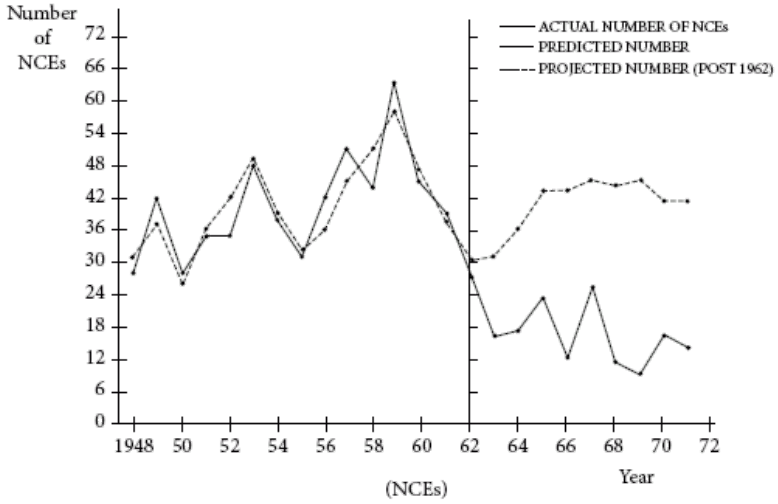

Figure 5.4 **Introduction of new chemical entities (NCEs), 1948–1971**.

Source: From Peltzman, 1973, Fig. 1.

Innovation

The initial papers studying the effect of FDA regulation on innovation used the 1962 Amendments as a treatment and the number of NCEs introduced each year as the outcome. Whether they used the UK data (Wardell 1973) or a model of introductions fitted to pre-1962 data (Peltzman 1973) as the control, they found large reductions in NCE introductions associated with the legislation. The chart from Peltzman (1973), reproduced in Figure 5.4, is illustrative.

The Peltzman paper was criticized, however, for overestimating the reduction in NCEs.[21] First, it examined only the quantity of drugs approved, not their quality. Perhaps only relatively unimportant drugs were held
p. 118 back in the 1960s. Second, drug companies may have voluntarily reduced NCE introductions even without
↳
p. 119 ↳
p. 120 ↳
p. 121 ↳
p. 122 ↳
p. 123 ↳
the 1962 Amendments. They may have interpreted the thalidomide controversy as evidence of increased consumer demand for safety and stopped developing drugs that had substantial side effects. Coupled with the great advance in the ability of the pharmacological sciences to detect side effects from drugs, companies may have held back drugs for fear of losing good will or facing legal liability. Third, given the high value of drugs developed in the 1950s and 1960s, it is possible that the returns to drug development had simply diminished by the 1960s (Grabowski et al. 1978).

Table 5.1 Review of literature concerning the effect of FDA regulation on innovation.

Source	Data (usually the dependent variable)	Measure of FDA Regulation	Baseline/Control	Finding
Wardell (1973)	NCE introductions, 1962–1971	1962 Amendments	UK	Annual NCE flow fell 54% due to 1962 Amendments.
Peltzman (1973)	NCE introductions, 1948–1971	1962 Amendments	Model of NCEs using pre-1962 data	Annual NCE flow fell 66% due to 1962 Amendments.
Grabowski et al. (1978)	NCE flow/R&D expenditures, 1960–1974	1962 Amendments and NDA approval times	UK	1962 Amendments increased average cost of NCE by factor of 2.3 (using 1962 dummy) or 1.9 (using approval times).
Wiggins (1981)	NCE introductions, 1970–1976	NDA approval times	Therapeutic classes with shorter approval times	Increase in approval times due to 1962 Amendments decreased NCE introductions 52%, holding R&D expenditures constant; accounting for effects of longer approval times on expenditures, reducing delay to pre-1962 levels would increase NCE introductions by 135%.
Wiggins (1983)	R&D expenses (from PhRMA) by therapeutic class, 1965–1968, 1971–1976	NDA approval times	Therapeutic classes with shorter approval times	Approval times reduced R&D expenditures during 1971–1976 but not 1965–1968, possibly because it took time for drug companies to determine how stringent FDA regulation would be after 1962.
May et al. (1983)	Number of NCEs tested on humans and NDA approvals, 1958–1979	1962 Amendments	Pre-1962 period	NCEs tested annually on humans fell from 89 to 17 in 1979; NDA approvals fell by 49%.
Cullen (1983)	190 drug product launches across 18 countries during 1961-1976	Surveyed 6 companies for their views of "regulatory tightness" in different counties in 1982. Ratings from 1 (most stringent) to 5 (least stringent).	17 countries other than US	Countries rated as having tighter regulations had (1) a larger increase in lag between first introduction in any country and introduction in that country from the 1960s to the 1970s and (2) a smaller increase in the number of products introduced in that country from 1960s to 1970s.
Jensen (1987)	NCE introductions by 26 firms 1969-1979	NDA approval times	Classes with shorter approval times, time trend	One-month decrease in approval times increased annual NCE introductions by 15%.
Thomas (1990)	NCE, sales and market cap of drug companies, 1960-1980	1962 Amendments and approval times	UK	FDA regulation did not affect NCEs at large firms but did substantially reduce NCE introductions by small firms. Due to reduced competition from small firms, sales rose at large firms in the US.

Abbreviations: FDA, Food and Drug Administration; NCE, new chemical entity; NDA, New Drug Application; PhRMA, Pharmaceutical Research and Manufacturers of America.

Table 5.2 Review of literature concerning the effect of FDA regulation on approval times.

Source	Data (usually the dependent variable)	Measure of FDA regulation (or other treatment variable)	Baseline/control	Finding
OTA (1989)	Effective patent length of drugs	*Ex post* commercial importance of drug		Drugs with greater *ex post* commercial importance have longer effective patent length.
Thomas (1990)	Preclinical testing, clinical testing, and NDA review times, 1960–1980	1962 Amendments	UK	US total development times grew from 35 months in 1960 to 120 months in 1970 to 145 months in 1980. The increases in preclinical testing, clinical, and NDA review times were 30, 60, and 20 months, respectively. In the UK, total development times increased from 30 to 70 months. Preclinical testing times were constant while the sum of clinical testing and review times increased by 40 months.
Kaitin et al. (1991)	Approval times	FDA ratings of novelty of drugs		FDA accelerated approval of more novel chemical entities.
Dranove & Meltzer (1994)	Time from drug patent application to NDA approval for 564 NMEs between 1950–1986	Various measures of importance of drug (e.g., FDA rating, commercial value, citations, worldwide introductions)		Development and approval times were lower for more important drugs.
Carpenter et al. (2003)	Approval times and FDA (CDER) staff, 1971-1998	PDUFA	Time trend	Funding for FDA staff had bigger influence on NDA review time than source of funding (user fees under PDFUA).
Olson (2004)	Approval times and FDA (CDER) staff, 1971-1998	PDUFA	Time trend	PDUFA reduced approval times by 34% by 1998. Different result from Carpenter et al. (2003) because Olson grouped approvals by approval year rather than NDA submission year.
Berndt et al. (2005a)	Clinical development and NDA review times, 1965-2003	PDUFA	Time trend	PDUFA reduced approval times by 7.6%/year during PDUFA I (1992-1996) and 3.6%/year during PDUFA II (1997-2001). PDUFA II may also have reduced clinical development times by 4.5%.

Abbreviations: CDER, Center for Drug Evaluation and Research; FDA, Food and Drug Administration; NDA, New Drug Application; NME, new molecular entity; OTA, Office of Technology Assessment; PDUFA, Prescription Drug User Fee Act.

Table 5.3 Review of literature concerning the effect of FDA regulation on safety.

Source	Data (usually the dependent variable)	Measure of FDA regulation	Baseline/Control	Finding
Bakke et al. (1984)	Drug discontinuations, 1964–1983	1962 Amendments	UK	Few discontinuations in either country, so no significant differences in discontinuations in US vs. UK
Bakke et al. (1995)	Drug discontinuations, 1974–1993	1962 Amendments	UK, Spain	More drugs discontinued in UK (20) and Spain (16) than US (10). Normalizing by number of drugs approved shrinks the difference: 4% in UK vs. 3% in US.
GAO (2002)	Drug withdrawals, 1986–2000	PDUFA	None	No significant effects of PDUFA on withdrawals. Withdrawals were 3.1% in 1986–1992 and 3.5% in 1993–2000.
CDER (2004)	Drug withdrawals, 1971–2004	PDUFA	None	No significant effects of PDUFA on withdrawals. Withdrawals were 2.7% in 1971–1993, 2.3% in 1994–Apr. 2004.
Berndt et al. (2005b)	Drug or biologic withdrawals, 1980–2000	PDUFA	None	No significant effects of PDUFA on withdrawals. Withdrawals (including biologics) were 2.8% in 1980–1992, and 2.2% in 1993–2000.
Carpenter et al. (2008)	FDA withdrawals, black-box warnings and voluntary withdrawals by drug companies, 1993–2004	PDUFA	Drugs approved well before or after PDUFA deadlines	PDUFA caused bunching of FDA approval during 2 months before deadlines. Drugs approved during that period had higher odds of being withdrawn by the FDA (OR = 5.5), getting blackbox warnings (4.4), and being voluntarily withdrawn (3.3) than drugs approved well before or after deadlines.

Abbreviations: CDER, Center for Drug Evaluation and Research; FDA, Food and Drug Administration; GAO, General Accounting Office; OR, odds ratio; PDUFA, Prescription Drug User Fee Act.

A second round of papers (Grabowski et al. 1978; Cullen 1983; Thomas 1990) therefore focused on the UK data as a control. The United Kingdom experienced the same increase in demand for safety after the thalidomide controversy and potentially diminishing returns in drug development. Yet premarket testing for safety was introduced only in 1963, and testing for efficacy not until 1973. Therefore, comparison of the United States and the United Kingdom in the 1960s would highlight the effect of premarket screening for efficacy. These UK comparisons also reveal significant reductions in research output associated with the increased US regulation.

One problem with studies that focused on the 1960s, according to Wiggins (1983), was it took some time for the FDA to decide how to implement the 1962 Amendments. Moreover, it also took drug companies some time to learn how cumbersome FDA regulation would ultimately be. Therefore, one can best assess the impact of the 1962 Amendments by examining how innovation responded in the 1970s. The difficulty with studying the 1970s is that the US and UK regulatory systems eventually converged, so the UK data did not obviously provide a valid control.[22] Therefore, investigators began quantifying FDA regulation by the amount of time it took for the FDA to review NDAs (Grabowski et al. 1978; Wiggins 1981, 1983; Jensen 1987; Thomas 1990).

Another issue that concerned economists was that, although NCE introductions fell in the 1960s, research expenditures rose. One interpretation was that the Peltzman finding underestimated the effect of FDA regulation because it focused on output rather than the productivity of research expenditures. A number of studies investigated this possibility by using NCE introductions/R&D expenditures as an outcome variable. For example, Grabowski et al. estimated that the 1962 Amendments increased the average cost of each NCE by a factor of 1.86 to 2.3. In addition, Wiggins (1981, 1983) examined whether FDA regulations reduced the amount companies invested in R&D and found that delays in FDA approval due to the 1962 Amendments reduced R&D expenditures in the 1970s. Holding these expenditures constant, NCE introductions fell 52 percent. Accounting for these reductions in R&D expenditures, NCE introductions fell a total of 135 percent after 1962.

p. 124 Whereas various studies introduced other improvements to the analysis of the effect of FDA regulation on innovation,[23] the most important of these was Thomas (1990), which suggested that FDA regulation might have had different effects on different companies. Specifically, regulation may have had a greater effect on small companies that were unable to afford the clinical testing required by the FDA and had less experience with the FDA process, compared with larger companies.[24] In addition, FDA regulation may have provided an indirect benefit to large companies by eliminating competition from smaller companies. Indeed, Thomas found that FDA regulation did not affect NCE introductions by large firms but did dramatically reduce NCE introduction by small firms. Moreover, because of reduced competition, sales (and market valuations) at large firms actually rose after FDA regulation.

Approval Times

A second important parameter in evaluating FDA regulation is its effect on approval time. Early work by Wardell demonstrated that US drug development times grew in comparison with UK times after adoption of the 1962 Amendments (Wardell 1973). This gap became known as the "drug lag." Thomas (1990) showed that the lag grew fastest in the 1960s but still grew in the 1970s, despite the fact that formally the UK and US regulatory systems had converged by 1973. For example, the lag between US and UK approval times grew from 5 months in 1960, to 70 months in 1970, and then to 75 months by 1980.

The remaining papers that examined approval times fall into two categories. One investigated heterogeneity in approval times for different drugs and the other looked at the role of PDUFA in lowering approval times. One criticism of the early literature on drug lag was that it might overestimate the cost of FDA delay if the delay affected only less valuable drugs. Of the studies that examined this issue, the best was Dranove and Meltzer (1994), which showed that drug approval times are shorter for more important drugs, where importance was measured by FDA ranking of a drug's novelty, its commercial value once approved, its citations in the academic literature, and in subsequent patents.[25]

p. 125 In 1992, Congress took note of the drug lag and passed PDUFA, which imposed deadlines on the FDA's review of NDAs and provided the FDA with more resources (from user fees imposed on NDA applicants) to evaluate NDA applications more quickly. The question academics asked was whether PDUFA actually lowered approval time and, if so, whether this was due to the deadline or the resources provided by Congress, or both.[26] Carpenter et al. (2003) and Olson (2004) come out on opposite sides of this debate. The difference is that Carpenter's group assigned a drug to the year in which its NDA application was filed,[27] whereas Olson assigned it to the year in which its NDA was approved. Because PDUFA was a national (rather than state) law, studies have used a dummy for the period after 1992 to code the treatment variable. This makes the year of assignment critical to the findings. Olson's findings were confirmed and extended by Berndt et al. (2005a), who showed that PDUFA I (1992–2006) reduced the approval time by 7.6 percent annually, whereas PDUFA II reduced it by only 3.6 percent annually. They also showed that, whereas PDUFA I had no effect on clinical development times, PDUFA II did lower these times by 4.5 percent. This is not

surprising, because one of the goals of subsequent versions of PDUFA was to streamline the regulatory process between the time of the IND application and that of the NDA application (Hutt et al. 2007).

Withdrawals

Early work on how FDA regulations affect the rate or the time at which drugs are withdrawn from the market focused on comparing the US and UK experience. These studies implicitly used the 1962 Amendments as the treatment variable. Bakke et al. (1984) looked at withdrawals from 1963 to 1983 and found no difference between the two countries. However, this can largely be explained by the small number of withdrawals in each country and thus low power to detect any difference in withdrawal rates. Bakke et al. (1995) revisited the question with data from 1974–1993 and found a larger difference between the United States and the United Kingdom. As predicted, the United States, which had relatively strict regulation (at least as measured by approval time), had both fewer drug withdrawals (10 versus 20 in the United Kingdom) and a lower withdrawal rate (2 versus 3 percent, respectively).

More recent work on withdrawal rates has focused on approval time as a measure for FDA regulatory intensity. Some relatively simple papers by the General Accounting Office (2002), the Center for Drug
p. 126 Evaluation and Research (2004), ↳ and Berndt et al. (2005b) compared the probability of withdrawal of a drug during the periods before and after the statute's adoption. They uniformly found somewhat lower, but not significantly lower, withdrawal rates before PDUFA.

Carpenter et al. (2008) used a more sophisticated approach to identify the effect of PDUFA. Instead of conducting a before-and-after PDUFA comparison, they demonstrated that PDUFA caused the agency to compress the timing of decisions on drugs to the two months just before PDUFA deadlines (months 11 and 12 for standard review drugs, months 9 and 10 for priority review drugs). The study then compared drugs approved close to the deadline with drugs approved well before or after the deadline. They found that drugs approved near deadlines had higher odds of being withdrawn (odds ratio [OR] = 5.5). Moreover, these drugs also had higher odds of having a blackbox warning (OR = 4.4) and of being voluntarily withdrawn by the drug company (OR = 3.3). Of course, these estimates only show that earlier deadlines increase withdrawal rates. They must be divided by the change in approval time implied by the early deadline to generate a regulatory dose-response curve. In effect, the timing of decisions (and withdrawals) during PDUFA needs to be compared with the timing before PDUFA.

Development Costs

A number of studies since the early 1970s have estimated the cost of drug development. These studies have been spaced roughly a decade apart and have generally covered the period between studies. DiMasi and Grabowski (2010) review this literature in chapter 2 of this volume. In early years, these studies relied on a small sample of drugs from a single firm (Schnee 1972; Sarett 1974) or on aggregate data (Mund 1970; Baily 1972). More recent studies relied on drug-level data from a sample of drug companies (e.g., Hansen 1979; DiMasi et al. 1991; Adams and Brantner 2006).[28] The latter studies attempted to separate, on the one hand, the cost of preclinical testing from that of clinical testing and, on the other hand, the direct out-of-pocket costs of research from the opportunity costs of that research. The last component, opportunity costs, is driven largely by delay and the real cost of capital. In order to account for the fact that many drugs ultimately fail to demonstrate value in trials or are not approved by the FDA, these studies divided total costs by the number of drugs approved, resulting in an estimate of the cost per approved drug rather than, say, cost per drug ever tested.

Together, these studies paint a picture of steadily increasing drug development costs. This is illustrated in Figure 5.5 (reproduced from DiMasi et al. 2003), which reports estimates from Hansen (1979), DiMasi et al.

(1991), and DiMasi et al. (2003) that roughly cover the 1970s, 1980s, and 1990s, respectively. Total costs per
p. 127 approved drug rose from $138 million in the 1970s to $802 million in the 1990s. ↳ More recent estimates have suggested that total costs might now be as high as $1.6 billion per drug.

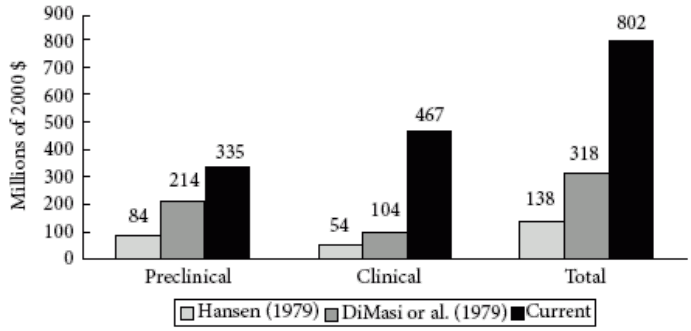

Figure 5.5 **Estimates of preclinical, clinical, and total costs of drug development from three studies.**.

Source: From DiMasi et al., 2003.

An important limitation of the literature on development costs is that it only demonstrates that costs have grown. The studies do not show that FDA regulation has been responsible for this growth. Although the dramatic increase in development costs after the enactment of the 1962 Amendments and during the run-up in approval times through the 1980s suggests that the FDA is responsible, the continued growth of development costs even after the decline of approval times in the 1990s raises some questions. Has the diminishing returns of drug development been the main driver of growing R&D costs in recent decades? Is approval time an adequate measure of the intensity of FDA regulation, or does the FDA offset shorter approval times with a higher standard for minimum required drug quality?

Consumer and Producer Surplus

The final parameters required to evaluate FDA policies are consumer and producer surplus. Only three papers have attempted to estimate these parameters (Table 5.4). An important limitation of these papers is that they each narrowly applied their estimates to just one or two components of the dynamic welfare framework laid out earlier in this chapter.

The first paper, Peltzman (1973), estimated the demand curve for new drugs by regressing the market share of newly introduced drugs in a therapeutic class on the ratio of new and old drug prices in that class. Peltzman included a dummy for the 1962 Amendments as a demand shifter. He used his estimate of the demand for new drugs only to estimate the value of information provided by the FDA minus the reduction in innovation due to FDA regulation. Peltzman's static framework for valuing the information produced by the FDA (see earlier discussion and Figure 5.2) suggests that the pre-1962 demand curve may not identify the "true" demand for new drugs because the FDA was not yet producing information about the quality of
p. 128 drugs. However, Peltzman argued that before 1962 consumers learned about ↳

p. 129 ↳
the true quality of drugs through experience, so demand during that period was still "true" demand. He estimated that the 1962 Amendments reduced demand for new drugs and thus the surplus from those drugs by roughly $420 million per year through 1970. He concluded that the loss of innovation resulting from the 1962 Amendments offset the value of any information they provided.

Table 5.4 Review of literature concerning the social surplus from FDA regulation.

Source	Data	Measure of FDA regulation	Methodology	Finding
Peltzman (1973)	Quantity and price of prescriptions of newly introduced and old drugs, by therapeutic class and year, 1960–1962, 1964–1970	1962 Amendments	Regress market share of new drugs on ratio of new and old drug prices. Surplus is $0.5^{*}(a - p)q$, where a is the y-intercept of the estimated demand curve, p is price, and q is quantity.	Consumer surplus for each year's NCEs was $51.9 million per year before the 1962 Amendments, $9.9m/year after the Amendments. Assuming 10% rate of return, discounted loss from Amendments was $420m/year.
Philipson et al. (2008)	Sales for all drugs, 1998–2002; PDUFA fees	PDUFA	Regress sales on age of drug to construct age-profile of sales. Producer surplus is PV of sales – user fees – variable costs, which are ¼ to ½ of sales. Social surplus calculated as different fractions of sales (before patent expiration: all sales, ½ sales, 0; after expiration: all sales) Change in surplus from PDUFA is benefit of starting sales earlier.	Additional producer surplus from PDUFA was $8–13 billion and additional total surplus from PDUFA is $13–30 billion, assuming a 9% rate of return.
Philipson et al. (2009b)	Survival probabilities for HIV, certain cancer patients by year; annual patient expenditures on key HIV, cancer drugs	N/A	Use Murphy-Topel framework to estimate WTP for improved survival. WTP minus patient expenditures is measure of consumer surplus (CS). Producer surplus (PS) is 80% of patient expenditures (assuming marginal costs are 20% of expenditures) Examine effect of 1-year acceleration of drug entry on social surplus.	CS from introduction of HAART in 1996 was $364 billion; PS was $38 billion. Entry 1 year earlier would have increased CS by $19 billion and PS by $4 billion. CS from introduction of Rituxan in 1998 was $12 billion and PS was $4 billion. Entry 1 year earlier would have increased CS $310 million and PS $330 million. CS from introduction of Receptin in 1999 was $149 and PS was $12 billion. Entry 1 year earlier would have increased CS $8 billion and PS $1 billion.

Abbreviations: HIV, human immunodeficiency virus; N/A, not applicable; PDUFA, Prescription Drug User Fee Act; PV, present value; WTP, willingness to pay.

The second paper to examine the FDA's impact on social surplus was Philipson et al. (2008). This differs from Peltzman (1973) in a number of respects. Instead of studying the effect of the 1962 amendments, this paper examines PDUFA and the value of reducing FDA approval time. Moreover, they used a substantially different methodology to identify surplus. Instead of estimating demand curves, they simply used sales data to bound the annual social surplus from all drugs on the market during 1998–2002. They then used drugs of different ages to estimate the stream of social surplus from a new drug over its life cycle. Finally, they used prior estimates of how much PDUFA accelerated drug introductions to estimate the value of accelerating

these streams of social surpluses. The paper concluded that PDUFA, by accelerating drug approvals, increased social surplus by $13–30 billion, assuming a 9 percent cost of capital.

The last paper was Philipson et al. (2009). Like Philipson et al. (2008), the focus was on identifying the value of accelerated introduction of drugs. The main difference is that the later paper used the effect of new drug introductions on survival probabilities of patients (combined with a value of life-years) to estimate a willingness to pay for a drug. Subtracting the price of the drug from this willingness to pay yielded the individual patient's consumer surplus. Producer surplus was estimated as 80 percent of sales revenue (assuming marginal costs of 20 percent of revenue). After estimating the stream of aggregate social welfare from three drugs—highly active antiretroviral therapy (HAART) for HIV patients, Rituxan for Hodgkin's lymphoma patients, and Herceptin for breast cancer patients—the authors calculated the value of accelerating this stream by one year. Using a 9 percent cost of capital, they estimated, for example, that introducing HAART one year earlier would have increased consumer surplus by $19 billion and producer surplus by $4 billion.

Summary

Comparing the empirical literature we have just reviewed to the dynamic cost-benefit framework described earlier in this chapter, one can draw two conclusions. First, the literature has attempted to estimate almost all the parameters required for welfare analysis. Second, the estimates for many important parameters are either dated or otherwise imperfect.

The first cost of FDA regulation is the reduction in innovation. Our best estimate of the welfare effect of that reduction is from Peltzman (1973), who examined only the period shortly after 1962. Indeed, no major study of innovation has taken place since that of Thomas (1990). The second cost is the higher development costs.
p. 130 Although we have strong, recent estimates of development costs from ↳ DiMasi et al. (2003) and Adams and Brantner (2006), neither of those studies attempted to estimate the causal relationship, let alone the correlation, between FDA regulation and development costs. The third cost is reduction in the present value of social surplus due to delay of drugs sales caused by FDA regulation. Philipson et al. (2008, 2009) made substantial progress in estimating this cost.

The fourth cost is the loss of social surplus from products that are not approved. Peltzman (1973) implicitly estimated this cost, because his paper did not distinguish between the reduction in surplus resulting from lower innovation and that due to nonapproved drugs. However, as we previously noted, that paper examined only the period shortly after the introduction of the 1962 Amendments. FDA standards and the demand for drugs have changed substantially since then. The final cost of FDA regulation is the decrease in producer surplus that occurs if information from the FDA reveals that true demand is lower than previously thought. The literature has provided no estimate of this potential cost.

There are three possible benefits to FDA regulation. The first is foregone research expenditures. These are the upfront cost savings resulting from the reduction in innovation. There are no estimates of this value in the literature. The second benefit is that FDA screening generates additional consumer surplus because drugs remain on the market longer before being withdrawn for safety reasons. Although a number of recent articles, summarized in Table 5.3, have tackled how much longer products remain on the market, there are no recent estimates of the surplus from this extended product life. It should not be difficult, however, to extend the estimates of surplus from Philipson et al. (2008, 2011) to answer this question.

The final benefit is that information from the FDA increases consumer surplus by revealing the true demand for drugs. Peltzman (1973) claimed to address this issue but basically assumed it away. He supposed that, before the 1962 Amendments, consumers learned true demand from experience with a drug, so that even pre-1962 demand is true demand. Therefore, when Peltzman estimated a reduction in demand after 1962,

he concluded that FDA regulation strictly reduced surplus. If pre-1962 demand was distorted by imperfect information, however, then his welfare estimates are incorrect. More work is required to test Peltzman's assumption and to update the numbers to reflect the information provided by modern FDA regulation.

Evaluation of Tort Law

In many countries including the United States, medical products are jointly regulated by agencies such as the FDA, which screens products to ensure that they are safe and effective before they are sold, and the tort
p. 131 liability system, which allows patients to sue manufacturers after they have consumed these products. ↳ This section examines the efficiency of using both systems to ensure the safety of medical products.

Overview of Tort Liability and Current Law on Preemption

In general, there are three bases for holding the manufacturer of a product liable for tort damages in state courts. First, the manufacturer is liable if there is a defect in the manufacture of the product. Second, the manufacturer is liable if the design of the product—not merely its construction—is defective. Specifically, this means that the product could have been designed to eliminate a health risk at a cost that would have been less than the health risk. Third, the manufacturer is liable if there is a risk from the product that the manufacturer could reasonably have discovered but failed to disclose to the consumer. Lawsuits making this claim are called "failure-to-warn" suits.

Although medical devices, like any product, are subject to suit under all three theories, drugs are not subject to design defect suits.[29] US courts have judged that drugs are unavoidably dangerous; that is, they cannot physically be redesigned to eliminate side effects. Therefore, as a class, they are exempt from design defect suits (Restatement (Second) of Torts §402(a), comment k; Restatement of Products Liability §6). The implication is that, whereas tort suits against device makers may function as minimum quality standards, suits against drug makers function only as disclosure requirements. Therefore, although there is substantial overlap between FDA authority over devices and state tort liability for devices, there is only partial overlap between FDA authority over drugs and state tort liability for drugs.

The overlap between FDA authority and state products liability raises the question of whether the existence of a federal regime for regulating medical products "preempts" or bars suits alleging product liability in state court. To answer this question, we must look first at the scope of federal regulation and then what the federal statute, the agency, and the courts have said about preemption. These sources suggest that federal regulation of devices is more likely to preempt state tort suits than is federal regulation of drugs.

With respect to drugs, the statute says very little about preemption. Moreover, the FDA has long taken the stance that its regulation of drugs does not preempt state suits. In 2006, the FDA changed it stance and passed a regulation stating that the agency's decisions concerning drug labels establish both a "floor" and a "ceiling" on warning that may be contained in those labels. Because state failure-to-warn suits seek to raise the level of warnings on labels, they are in tension with the FDA's ceiling on warnings. The preamble to
p. 132 the 2006 regulations explains that the ↳ state tort liability is redundant with federal liability and, in any case, the FDA has greater expertise on the costs and benefits of drug labels. This argument was not persuasive to the US Supreme Court which, in *Wyeth v Levine*, 555 US 555 (2009), ruled that state failure-to-warn suits were not preempted by federal regulation of drug labels. The Supreme Court observed that the FDA has limited resources and that state suits can complement the FDA's enforcement efforts.[30]

In contrast to the case of drugs, the FD&C Act has a provision explicitly preempting state suits against makers of medical devices: no State ... may establish or continue in effect with respect to a device intended

for human use any requirement— (1) which is different from, or in addition to, any requirement applicable under this Act" (FD&C Act, §521). This provision applies to both state statutes and tort suits (*Riegel v Medtronic, Inc.*, 552 US 312 [2008]). However, a state suit against a device is preempted only if the FDA has specifically approved that device. Suits against class III drugs that the federal Act permits to be marketed under §510(k) even before FDA has judged them safe and effective[31] are not preempted (*Medtronic, Inc. v Lohr*, 518 US 470 [1996]). Nor are suits against class I or class II devices, because the FDA has never determined whether they are specifically safe and effective.

Rationale for Tort Liability

There is both a general theoretical literature on the merits of *ex ante* regulation versus *ex post* liability and a specific debate—played out mainly in the legal literature and in court cases—about the merits of FDA regulation versus products liability suits against drug companies. The context for the specific debate is the legal cases, just summarized, addressing whether FDA regulation should preempt tort suits against medical products companies.

A central theme of the general literature is that regulation and tort liability are substitute methods of enforcement (Wittman 1977). The implication is that heavier *ex post* liability justifies weaker *ex ante* regulation, and vice versa. When liability is taken as given, Shavell (1984a, 198b) and Kolstad et al. (1990) have suggested that the regulator should set the minimum level of safety below the socially optimum level, because *ex post* liability will take up the slack. Conversely, if regulation is taken as given and the regulator
p. 133 chooses the optimal (or an excessive) level of safety, any ↳ tort liability can only lead to inefficient expenditure on safety and higher prices (Philipson et al. 2011).

So how much weight should be placed on regulation versus litigation, and specifically, should the FDA be given exclusive authority to regulate medical products, or should that authority be given to courts? Shavell (1984a, 1984b) highlighted three factors that influence these decisions. The first is knowledge and expertise. Regulators tend to oversee a small class of products and develop expertise about them. The FDA also employs pharmacologists, doctors, and biostatisticians who are specifically trained to judge the safety and efficacy of medical products. In contrast, courts employ generalist judges or lay juries who only occasionally evaluate medical products.

In some respects, however, courts may make better judgments than the FDA. This is not because a court takes action only after a product has been approved for sale and therefore has access to the product's performance in the market (Shavell 1984a). The FDA also has access to data on a medical product's performance in the market (through the Adverse Event Reporting System) and can withdraw approval for a drug even after it has been sold for some time.[32] The better reason courts may complement the FDA is that the agency has long complained about having insufficient resources to oversee all medical products (*Wyeth v Levine*). This is evident in the long time it took for the FDA to reevaluate the thousands of drugs and devices approved before the 1962 and 1976 Amendments, respectively (Hutt et al. 2007).

A concern related to the relative knowledge or capacity of the FDA compared with courts is that of bias. Supporters of tort liability allege that, because the FDA is now substantially funded by user fees from the drug industry, it has been captured by that industry and is incapable of setting safety standards in an unbiased manner. Opponents of tort liability observe that courts consider only the side effects from products, not their benefit to consumers (*Riegel v Medtronic*). Therefore, juries do not truly maximize consumer welfare, which depends on both safety and efficacy.[33]

The second factor relevant to the choice between regulation and litigation is the remedy available in each regime. If there is a positive probability that the manufacturer of an unsafe product may not be sued or if the
p. 134 manufacturer does not have to pay damages that exceed its assets, then expected tort liability may ↳ be less

than the expected harm from an unsafe product. In this case, tort liability will lead to suboptimal levels of safety (Shavell 1984a, 1984b). On the other hand, unlike regulation, tort liability not only regulates the quality of a product but also provides compensation to victims injured by that product. This compensation has insurance value.

The third factor is administrative costs. The advantage of the tort system over regulation is that, instead of incurring the cost of regulating every product, whether it is safe or not, it incurs cost only after a product has proved to be unsafe.[34] The disadvantage of the tort system is that it is very expensive to administer. For example, it is widely reported that only one third of each dollar of compensation paid by defendants actually reaches the pockets of plaintiffs (Polinsky and Shavell 2010). This limits the compensatory function of tort law.

Shavell's three factors (1984a, 1984b) focus on the choice between *ex ante* regulation and *ex post* liability. In reality, the FDA is not going away. In practice, therefore, the choice is between regulation by the FDA alone or both FDA regulation and tort liability together. Philipson et al. (2011) provide a framework for evaluating the efficiency of single versus dual government interventions to raise product quality or safety. Their main argument is that the standard efficiency implications of product liability change when there is dual regulation of safety through another government agency such as the FDA. In particular, increasing liability (through preemption or other means) in the presence of the FDA can reduce welfare under conditions when it would raise welfare in the absence of the FDA.

The critical issues are whether consumers are fully informed and capable of making sound judgments about medical product risks after FDA regulation. If consumers are not fully informed or making sound judgments, products liability may improve welfare by producing more information or, if consumers underestimate risks, by increasing costs and reducing consumption. If consumers are well informed but do not underestimate risks after FDA regulation, then products liability may simply impose an inefficient tax on the informed and rational market. Even if consumers are not well informed by the FDA, products liability may reduce welfare if the FDA's minimum quality regulations already fully protect uninformed consumers.

The cost from dual government intervention is illustrated in Figure 5.6. The x-axis shows the level of safety (s); the top U-shaped curve, $C(s)$, is a firm's cost as a function of safety. The optimal level of safety chosen by the firm in the presence of optimal product liability, s^{PL}, is the bottom of this curve, where the marginal cost of producing safety equals the marginal benefit in terms of reduced liability costs. Under a regime that lowers product liability, the cost curve shifts downward to the lower curve, $C^{O}(s)$, which differs from the initial cost curve in two ways. First, costs are lower under reduced liability, because firms pay lower liability
p. 135 costs. Second, ↳ the optimal level of safety is reduced to s^{PLo}. The firm's costs under reduced product liability are therefore given by point B.

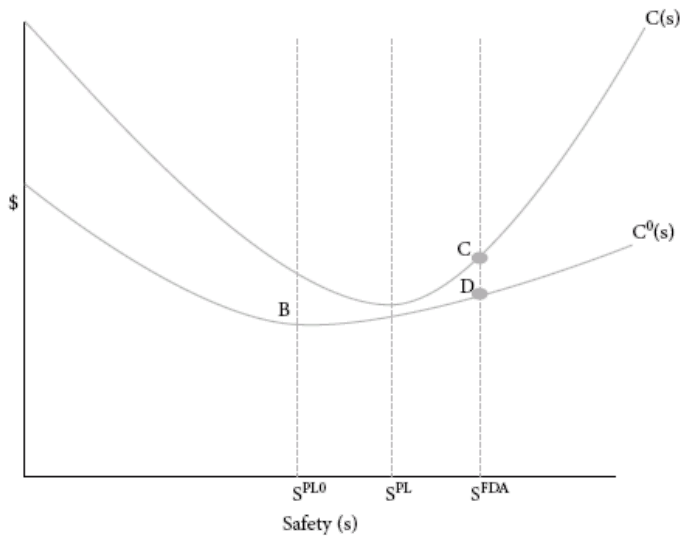

Figure 5.6 **Welfare gain from preemption**.

The level of safety mandated by the FDA, s^{FDA}, may lie to the left or to the right of the level of safety induced by liability, s^{PL}, depending on whether FDA safety levels are binding. Consider the likely case when the safety mandated by the FDA lies to the right of s^{PL}, as shown in Figure 5.6. In this case, the FDA-mandated level of safety is binding on firms; they will provide s^{FDA} with or without the reduced level of liability. The preemption raises welfare by lowering marginal costs from point C to point D, while having no effects on safety. The reduction in prices that results from this cost reduction raises welfare because it enhances access without any safety effects.

This analysis suggests that, in theory, preemption doctrine has the potential to increase welfare when the presence of the FDA is binding on firms. Intuitively, product liability in general has two opposing effects on welfare. It positively affects welfare by inducing the firms to provide safe drugs but negatively affects welfare by increasing marginal costs and price. When the level of safety mandated by the FDA is binding, the second effect dominates, because product liability has no *additional* effect on the level of safety firms choose to provide but does raise prices and thus restricts access.

Helland et al. (2010) have presented an alternative model of tort liability that starts from the assumption that FDA regulation does not remove all risky products from the market. In addition, it takes seriously the legal theory that underlies most product liability suits against drugs, namely that the manufacturer is liable only for failure to warn about defects about which it knew or should have known. This is an important point,
p. 136 because it implies that tort suits cannot change the side ↳ effects from drugs; they can only change the amount of information on side effects provided about drugs.

Therefore, in the Helland et al. (2010) model, the key issues are how consumers assess the safety of drugs for which no side effects are disclosed ("nondisclosing drugs") and how changes in liability affect the number of nondisclosing drugs. If a drug company fails to disclose a side effect from its product, rational consumers would infer that the drug is no safer than the average drug among the set of drugs whose side effects are not disclosed. Given this inference, an increase in tort damages has two effects. The first is to increase costs of nondisclosing drugs, which shifts the supply curve for these drugs inward. This in turn causes some companies to leave the drug market and changes the number of companies that disclose side effects. This shift in the composition of nondisclosing companies triggers a second effect from tort liability among rational consumers, which is to shift demand. If the number of high side effect drugs that drop out because of the negative supply effect exceeds the number of low side effect drugs that now decide to disclose their side effects, the average side effect among the set of nondisclosing drugs that remain decreases. This

implies that the demand effect will be positive. If a positive demand effect outweighs the negative supply effect, equilibrium quantity of sales will be larger. Because rational consumers do not purchase a drug unless it has positive expected consumer surplus for them, this increase in equilibrium quantity suggests an increase in social welfare.

Empirical Evaluation of Tort Liability

There is scant empirical evidence on the value of product liability for medical products. Two of the three papers on the topic examined the case of vaccines for diseases such as diphtheria, pertussis (whooping cough), and tetanus. Before the 1980s, there was little product liability litigation over vaccines.[35] The boom in liability during that decade caused many companies to leave the market. Partly in response to this problem, Congress established the National Vaccine Injury Compensation Program (NVICP) in 1988, which sharply reduced vaccine manufacturers' legal liability and created a public patient compensation fund funded by excise taxes on vaccine manufacturers. In effect, over the course of three decades, the United States went from little tort liability, to lots of tort liability, and then to something like a workers' compensation system paired with virtually no tort liability. These changes did not affect the role of the FDA, which still, throughout this period, determined whether vaccines could be sold. They affected only the level of tort liability.

p. 137 Both Manning (1994) and Philipson et al. (2011) studied the effects of tort liability on vaccine prices. They compared the price of the diphtheria, pertussis, and tetanus (DPT) vaccine with that of the diphtheria and tetanus (DT) vaccine. The only difference between these two products—the pertussis component in the former—is known to cause neurological side effects in roughly 1 of every 330,000 cases. This led to a large number of suits against DPT but not DT. Thus, the price of DT vaccine is a good control for the price of DPT vaccine in the absence of liability. Figure 5.7, which plots the prices of the DT and DPT vaccines over time, illustrates the main result. The prices of the DT and DPT vaccines were quite similar before 1982, when lawsuits were rare. After 1982, when the number of lawsuits for adverse events for the pertussis component began to rise sharply, the price of the DPT vaccine increased significantly compared with that of the DT vaccine. Because the two vaccines had similar prices prior to 1983, one may interpret the post-1982 difference in price as the cost of liability for the pertussis component. At its peak in 1986, the difference in price between the two vaccines was $14.04, and liability costs accounted for almost 96 percent of the DPT vaccine's price. Not only did prices of DPT rise with increased liability; they also fell after introduction of the NVICP in 1988, from $11.78 to $7.73—a decrease of 34 percent. The difference between the DPT and DT prices that persisted after the NVICP reflects the larger excise tax on DPT because of the risk of neurological side effects from the pertussis component.

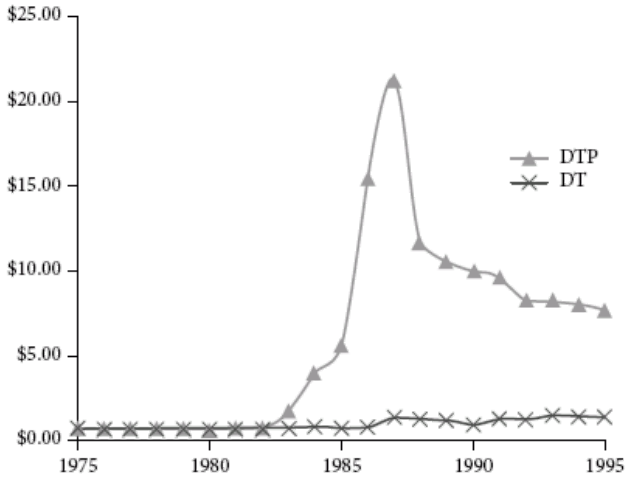

Figure 5.7 **Prices for the diphtheria, pertussis, and tetanus (DPT) vaccine and the diphtheria and tetanus vaccine, 1975–1995.**

Manning (1994) extended his analysis by estimating the implicit load on health insurance that the tort system provided to victims of the pertussis vaccine. Comparing data on tort damages to the price of vaccines, he estimated that $1 of compensation for victims cost manufacturers $1.60 to $22, with a median
p. 138 estimate ↳ of about $5.[36] That administrative cost is larger than typically estimated for tort liability.

Philipson et al. (2009a) supplemented their analysis of price effects with a comparison of the safety of vaccines introduced before the NVICP to that of vaccines introduced after that law. Presumably, if reducing tort liability made products less safe, vaccines introduced after the law should have a higher rate of side effects. Looking at a dozen or so vaccines, they found no statistically significant difference in the rate of side effects. Of course, the sample size for this analysis was very small, and the crude before-and-after comparison cannot rule out fortuitous or secular technological improvements that might have reduced the side effects of new vaccines even with tort liability.

It is important, however, to extend these studies. Although vaccines, like devices, are subject to liability merely for causing side effects, most drugs are not. In practice, drug manufacturers are subject to liability only if they fail to warn patients or doctors about side effects they know about. They are not liable for those side effects once the user is warned. Makers of DPT, in contrast, were liable for neurological side effects even though they warned users about those side effects. Therefore, the cost of product liability for drugs may be much lower than implied by the analysis done for vaccines.

Only one published paper has studied the effect of failure-to-warn suits on drug prices.[37] Manning (1997) examined the difference between the US and the Canadian price for 119 different drugs in 1991. Whereas there is substantial tort litigation against drugs in the United States, there is limited scope for such litigation in Canada. Therefore, Canada can serve as a control for liability-free drug prices in the United States, much as DT was a control for the nonliability price of DPT in the vaccine studies. When Manning regressed the US/Canada price differences on the degree of liability exposure of each drug (approximated by the number of successful tort suits against it), he estimated that half the price difference could be explained by tort liability.[38]

p. 139 Manning's (1997) result was dramatic but also of limited value. The analysis was cross-sectional, so one does not know whether some unobservable feature of the drug led to both more suits and higher prices. For example, perhaps the FDA required more clinical testing for some drugs but not others and it was this testing that led to both a higher price and more tort suits. Indeed, the causation might be reverse: perhaps

the higher price of some drugs signaled higher profits for manufacturers and invited suits. Finally, Canada is not a perfect control. Although it does have less liability, it does not have zero liability.

Ultimately, there is need for more empirical research. Like the literature on the costs and benefits of FDA regulation, the literature on tort liability has largely overlooked medical devices. More importantly, there is no persuasive evidence on the effect of tort liability on drug safety. Therefore it is hard to assess whether the costs of liability from either the vaccine studies or the US/Canada comparisons represents money well spent.

Conclusion

The FDA controls which medical products may be sold and, together with product liabilty law, the information that must be packaged with these products. These levers have potentially large impacts on the welfare benefits from medical products. Like the products they regulate, the FDA and the courts can have both beneficial effects (limiting the sales of unsafe products) and costly side effects (discouraging the development and sale of efficacious medicines). In this chapter, we reviewed the basic theoretical models for evaluating the costs and benefits of government intervention and empirical studies that estimate parameters required by these models. On the margin, it appears that empirical work lags behind the theory. It is too early to pronounce judgment on the value of either FDA regulation or tort law. Given the importance of the topic, however, we hope that substantial progress may be made by the time the next review chapter on government intervention in drug markets is due to be written.

References

Adams, C.P., and V.V. Brantner. 2006. “Estimating the Cost of New Drug Development: Is It Really $802 Million?” *Health Affairs* 25(2): 420–428.
WorldCat

Baily, M.N. 1972. “Research and Development Costs and Returns: The US Pharmaceutical Industry.” *Journal of Political Economy* 80(1): 70–85.
WorldCat

Bakke, O.M., M. Manocchia, F. de Abajo, K.I. Kaitin, and L. Lasagna. 1995. “Drug Safety Discontinuations in the United Kingdom,
p. 140 the United States, and Spain from 1974 ↳ Through 1993: A Regulatory Perspective.” *Clinical Pharmacology & Therapeutics* 58: 108–117.

Bakke, O.M., W.M. Wardell, and L. Lasagna. 1984. “Drug Discontinuations in the United Kingdom and the United States, 1964 to 1983: Issues of Safety.” *Clinical Pharmacology & Therapeutics* 35(5): 559–567.
WorldCat

Bekelman, J.E., Y. Li, and C.P. Gross. 2003. “Scope and Impact of Financial Conflicts of Interest in Biomedical Research: A Systematic Review.” *Journal of the American Medical Association* 289(4): 454–465.
WorldCat

Berndt, E.R., A.H.B. Gottschalk, T.J. Philipson, and M.W. Strobeck. 2005a. “Assessing the Impacts of the Prescription Drug User Fee Acts (PDUFA) on the FDA Approval Process.” *Forum for Health Economics & Policy* 8: Article 2. Available at http://www.bepress.com/fhep/8/2 (accessed November 2011).

Berndt, E.R., A.H.B. Gottschalk, T.J. Philipson, and M.W. Strobeck. 2005b. “Industry Funding of the FDA: Effects of PDUFA on Approval Times and Withdrawal Rates.” *Nature Reviews Drug Discovery* 4(7): 545–554.
WorldCat

Carpenter, D., M. Chernew, D.G. Smith. and A.M. Fendrick. 2003. “Approval Times for New Drugs: Does the Source of Funding for FDA Staff Matter?” *Health Affairs* doi: 10.1377/hlthaff.w3.618.

Carpenter, D., E.J. Zucker, and J. Avorn. 2008. “Drug-Review Deadlines and Safety Problems.” *New England Journal of Medicine* 358(13): 1354–1361.
WorldCat

Center for Drug Evaluation and Research. 2004. *2003 Report to the Nation: Improving Public Health through Human Drugs*. Washington, DC: US Department of Health and Human Services, Food and Drug Administration.
Google Scholar Google Preview WorldCat COPAC

Cullen, R. 1983. “Pharmaceuticals Inter-Country Diffusion.” *Managerial & Decision Economics* 4(2): 73–82.
WorldCat

Cutler, D., G. Long, E.R. Berndt, J. Royer, A.A. Fournier, A. Sasser, and P. Cremieux. 2007. “The Value of Antihypertensive Drugs: A Perspective on Medical Innovation.” *Health Affairs* 26(1): 97–110.
WorldCat

Cutler, D., and M. McClellan. 2001. “Is Technological Change in Medicine Worth It?” *Health Affairs* 20(5): 11–29.
WorldCat

Cutler, D., and E. Richardson. 1998. “The Value of Health: 1970–1990.” *American Economic Review Papers & Proceedings* 88(2): 97–100.
WorldCat

DiMasi, J.A., N.R. Bryant, and L. Lasagna. 1991. “New Drug Development in the United-States from 1963 to 1990.” *Clinical Pharmacology & Therapeutics* 50(5): 471–486.
WorldCat

DiMasi, J.A., R.W. Hansen, and H.G. Grabowski. 2003. “The Price of Innovation: New Estimates of Drug Development Costs.” *Journal of Health Economics* 22(2): 151–185.
WorldCat

Dranove, D., and D. Meltzer. 1994. “Do Important Drugs Reach the Market Sooner?” *RAND Journal of Economics* 25(3): 402–423.
WorldCat

General Accounting Office. 2002. “*Food and Drug Administration: Effect of User Fees on Drug Approval Times, Withdrawals, and Other Agency Activities.*” Report to the Chairman, Committee on Health, Education, Labor, and Pensions, U.S. Senate. GAO-02-958. Washington, DC: US GAO.
Google Scholar Google Preview WorldCat COPAC

Grabowski, H.G., J.M. Vernon, and L.G. Thomas. 1978. “Estimating the Effect of Regulation on Innovation: An International Comparative Analysis of the Pharmaceutical Industry.” *Journal of Law and Economics* 21(1): 133–163.
WorldCat

Hansen, R.W. 1979. “The Pharmaceutical Development Process: Estimates of Current Development Costs and Times and the Effects of Regulatory Changes.” In *Issues in Pharmaceutical Economics*, ed. R.I. Chien, 151–187. Lexington, MA: Lexington Books.
Google Scholar Google Preview WorldCat COPAC

p. 141

Helland, E., D. Lakdawalla, A. Malani, and S. Seabury. 2010. *The Effect of Tort Liability on the Market for Prescription Drugs*. Working paper. Available on request from Anup Malani: amalani@uchicago.edu.

Hutt, P.B., R.A. Merrill, and L.A. Grossman. 2007. *Food and Drug Law: Cases and Materials*. 3rd ed. New York: Foundation Press.
Google Scholar Google Preview WorldCat COPAC

Jensen, E.J. 1987. “Research Expenditures and the Discovery of New Drugs.” *Journal of Industrial Economics* 36(1): 83–95.
WorldCat

Kaitin, K.P., P.A. DiCerbo, and L. Lasagna. 1991. “The New Drug Approvals of 1987, 1988, and 1989: Trends in Drug Development.” *The Journal of Clinical Pharmacology* 31(2): 116–122.
WorldCat

Kolstad, C.D., T.S. Ulen, and G.V. Johnson. 1990. “Ex Post Liability for Harm vs. Ex Ante Safety Regulation: Substitutes or Complements?” *American Economic Review* 80(4): 888–901.
WorldCat

Lakdawalla, D., T. Philipson, and Y.R. Wang. 2006. *Intellectual Property and Marketing*. NBER Working Paper 12577. Cambridge, MA: National Bureau of Economic Research.

Lakdawalla, D., and N. Sood. 2006. *Health Insurance as a Two-Part Pricing Contract*. NBER Working Paper 12681. Cambridge, MA: National Bureau of Economic Research.

Lakdawalla, D., and N. Sood. 2009. “Innovation and the Welfare Effects of Public Drug Insurance.” *Journal of Public Economics* 93(3–4): 541–548.
WorldCat

Lexchin, J., L.A. Bero, B. Djulbegovic, and O. Clark. 2003. “Pharmaceutical Industry Sponsorship and Research Outcome and Quality: Systematic Review.” *British Medical Journal* 326(7400): 1167B–1170B.
WorldCat

Lichtenberg, F. 2003. “Pharmaceutical Innovation, Mortality Reduction, and Economic Growth.” In *Measuring the Gains from Medical Research: An Economic Approach*, eds. K.M. Murphy and R.H. Topel,74–109. Chicago: University of Chicago Press.
Google Scholar Google Preview WorldCat COPAC

Malani, A., O. Benbom, and M. van der Laan. 2009. *Accounting for Differences Among Patients in the FDA Approval Process*. Working paper. Available on request from Anup Malani: amalani@uchicago.edu.
Google Scholar Google Preview WorldCat COPAC

Manning, R.L. 1994. “Changing Rules in Tort Law and the Market for Childhood Vaccines.” *Journal of Law & Economics* 37(1): 247–275.
WorldCat

Manning, R.L. 1997. “Products Liability and Prescription Drug Prices in Canada and the United States.” *Journal of Law & Economics* 40: 203–243.
WorldCat

May, M.S., W.M. Wardell, and L. Lasagna. 1983. “New Drug Development During and After a Period of Regulatory Change: Clinical Research Activity of Major United States Pharmaceutical Firms, 1958–1979.” *Clinical Pharmacology and Therapeutics* 33(6): 691–700.
WorldCat

McGuire, T., R. Nelson, and T. Spavins. 1975. “An Evaluation of Consumer Protection Legislation: The 1962 Drug Amendments. A Comment.” *Journal of Political Economy* 83(3): 655–662.
WorldCat

Mund, V.F. 1970. “The Return on Investment of the Innovative Pharmaceutical Firm.” In *The Economics of Drug Innovation*, ed. J.D. Cooper, 125–138. Washington, DC: American University.
Google Scholar Google Preview WorldCat COPAC

Murphy, K.M., and R.H. Topel. 2003. *Measuring the Gains from Medical Research: An Economic Approach*. Chicago: University of Chicago Press.
Google Scholar Google Preview WorldCat COPAC

Murphy, K.M., and R.H. Topel. 2006. “The Value of Health and Longevity.” *Journal of Political Economy* 114(5): 871–904.
WorldCat

Musgrave, R.A. 1959. *The Theory of Public Finance*. New York: McGraw-Hill.
Google Scholar Google Preview WorldCat COPAC

Newhouse, J. 1992. “Medical Care Costs: How Much Welfare Loss?” *Journal of Economic Perspectives* 6(3): 3–21.
WorldCat

Office of Technology Assessment. 1989. *Pharmaceutical R&D: Costs, Risks and Rewards*. Washington, DC: Office of Technology Assessment.
Google Scholar Google Preview WorldCat COPAC

Olson, M.K. 2004. “Explaining Reductions in FDA Drug Review Times: PDUFA Matters.” *Health Affairs*. Available at http://content.healthaffairs.org/cgi/content/full/hlthaff.w4.s1v1/DC1 (accessed November 2011).

Peltzman, S. 1973. “An Evaluation of Consumer Safety Protection Legislation: The 1962 Drug Amendments.” *Journal of Political Economy* 81(5): 1049–1091.
WorldCat

Peltzman, S. 1975. “An Evaluation of Consumer Product Legislation: The 1962 Drug Amendments. A Reply.” *Journal of Political Economy* 83(3): 663–667.

WorldCat

Philipson, T.J., E.R. Berndt, A.H.B. Gottschalk, and E. Sun. 2008. “Cost-Benefit Analysis of the FDA: The Case of the Prescription Drug User Fee Acts.” *Journal of Public Economics* 92: 1306–1325.
WorldCat

Philipson, T.J., and E. Sun. 2008. “Is the Food and Drug Administration Safe and Effective?” *Journal of Economic Perspectives* 22(1): 85–102.
WorldCat

Philipson, T.J., E. Sun, and D. Goldman. 2011. *The Effects of Product Liability Exemption in the Presence of the FDA*. In *Regulation vs. Litigation: Perspectives from Economics and Law*. ed. D. P. Kessler: 137–163. Chicago, IL: University of Chicago Press.
Google Scholar Google Preview WorldCat COPAC

Philipson, T., E. Sun, A. Jena, and D. Goldman. 2009. *A Re-examination of the Costs of Medical R&D*. Working paper. Available at http://www.nber.org/public_html/confer/2009/FHf09/Philipson_Sun_Jena_Goldman.pdf.

Polinsky, A. M., and S. Shavell. 2010. “The Uneasy Case for Product Liability.” *Harvard Law Review* 123(6): 1437–1492.
WorldCat

Sarett, L.H. 1974. “FDA Regulations and Their Influence on Future Research and Development.” *Research Management* 17(2): 18–20.
WorldCat

Schnee, J.E. 1972. “Development Costs: Determinants and Overruns.” *Journal of Business* 45(3): 347–374.
WorldCat

Shavell, S. 1984a. “Liability for Harm versus Regulation of Safety.” *Journal of Legal Studies* 13(2): 357–374.
WorldCat

Shavell, S. 1984b. “A Model of the Optimal Use of Liability and Safety Regulation.” *RAND Journal of Economics* 15(2): 271–280.
WorldCat

Thomas, L.G. 1990. “Regulation and Firm Size: FDA Impacts on Innovation.” *RAND Journal of Economics* 21(4): 497–517.
WorldCat

Wardell, W. 1973. “Introduction of New Therapeutic Drugs in the United States and Great Britain: An International Comparison.” *Clinical Pharmacology and Therapeutics* 14(5): 773–790.
WorldCat

Wiggins, S.N. 1981. “Product Quality Regulation and New Drug Introductions: Some New Evidence from the 1970s.” *Review of Economics and Statistics* 63(4): 615–619.
WorldCat

Wiggins, S.N. 1983. “The Impact of Regulation on Pharmaceutical Research Expenditures: A Dynamic Approach.” *Economic Inquiry* 21(1): 115–128.
WorldCat

Wittman, D. 1977. “Prior Regulation versus Post Liability: The Choice between Input and Output Monitoring.” *Journal of Legal Studies* 6(1): 193–211.
WorldCat

Notes

1. This is also the standard now applied to biologics, which are defined by the Public Health Service Act §351(i) as “any virus, therapeutic serum, toxin, antitoxin, vaccine, blood, blood component or derivative, allergenic product, or analogous product … applicable to the prevention, treatment or cure of a disease or condition of human beings.” That Act states that biologics are to be regulated under the FD&C Act, and the FDA has interpreted the term "drug" in the FD&C Act to cover biologics.
2. An exception is cancer drugs. Phase 1 studies of these drugs, because they are extremely toxic, are performed on terminally ill cancer patients.
3. Because of First Amendment protection for truthful commercial speech, however, companies can share peer-reviewed articles about unapproved uses, so long as they are not accompanied by nonacademic marketing (Washington Legal Foundation v Henney, 1999).
4. This gap in the regulatory regime is partly addressed by the tendency of insurance companies to refuse reimbursement for off-label prescription of drugs. But such off-label use is difficult to monitor and can often be evaded by administering drugs on an inpatient basis at a hospital (rather than in a physician's office), because hospital reimbursement is largely based on a patient's diagnosis, not the specific treatment provided.
5. The statute also guaranteed that the FDA would not permit any generic drug manufacturer to market a drug within 5 years. This was important to pioneer companies, because a generic maker is permitted to sue to invalidate a pioneer's drug patent before it formally expires. The delayed FDA approval ensured that, even if a generic won its patent suit, the pioneer company would enjoy market exclusivity for 5 years.
6. This concept is embedded in the average patient standard that the FDA employs. By barring drugs that may benefit some patients but not the average patient, the FDA is assuming that they are better at sorting different patients to a drug than doctors are (Malani et al. 2009).
7. Medical product companies would conduct R&D and thus produce some information about product quality even in the absence of any government regulation requiring them to do so. A more precise cost-benefit analysis would weigh the cost of the additional information the government requires against the benefit of that additional information.
8. To keep matters simple, we ignore complications from, for example, health insurance (Lakdawalla and Sood 2006, 2009) and pharmaceutical advertising (Lakdawalla et al. 2006). These features make the calculation of welfare more complicated.
9. We ignore philosophical issues such as whether welfare can be calculated by using subjective expected utility. With subjective utility, consumer perceptions are important to welfare and the area under DD but above p may provide some measure of welfare. Instead, for simplicity, we assume that only objectively verified benefits matter for welfare calculations.
10. Peltzman (1975) argued that this scenario is unlikely because companies have an incentive not to undersell their drugs prior to regulation. However, this does not rule out an increase in demand resulting from the other causes listed.
11. Producing information can entail both fixed and variable costs, which may not be reflected in price in a straightforward fashion. For example, the cost of information may depend on the quantity of the drug consumed in equilibrium. To keep matters simple, we ignore this complexity in Figure 5.2.
12. If paternalism is used to justify minimum quality standards, then one cannot use the graph in Figure 5.2. When consumers are unaware of their best interest (i.e., utility), their prior demand curves are not an appropriate baseline from which to calculate the benefits of a change in demand after the production and processing of new information.
13. There are two things that may complicate this model but that we ignore to simplify the exposition. First, the company may treat the cost of information as a fixed cost. If the company then prices at marginal cost, p′ will not lie above p. The gray-shaded area will be zero even though cost of information production is positive. Second, whereas the net surplus from information is accrued repeatedly each period the product is sold, the cost of information may be incurred only once, when the information is first produced. If we assume that the cost of information is amortized over each period the product is sold and that this is reflected in the average p′ over those periods, the line-shaded area below average p′ will reflect the average per-period cost of information.
14. This framework can easily be modified to account for drugs that lose patent protection before they are withdrawn from the market. Loss of patent protection would reduce producer surplus and increase consumer surplus.
15. Here we have assumed for simplicity that N and W are NW independent. This is possible only if products are randomly withheld from development, which obviously is incorrect. The marginal product will be of lower value than drugs that continue to be developed and of higher value than products that would not have been offered in either case.
16. Consumer surplus depends on the price of a product. The FDA does not, however, consider or even know the price of products they screen. This can be addressed by incorporating the price of a drug in the critical value v, v so that a higher v v represents a higher standard for approval irrespective of price.
17. If the FDA-required production of information shifts out consumer demand in Figure 5.2, it is possible that producer surplus rises rather than falls.

18. Philipson, Sun, Jena, and Goldman (2009) also provide a welfare framework that accounts for this effect. They also calculate the magnitude of this effect for certain HIV therapies and cancer therapies. That paper does not, however, account for the other welfare effects of FDA regulation presented in our text.
19. Specifically, a proper welfare calculation requires separately estimating the lost surplus from products that are not approved by the FDA due to minimum quality regulations and the demand curve for products had the FDA not provided more accurate information on quality.
20. Similarly, in studies that examine the corresponding UK code, the treatment dummy variable is set at "0" before 1973, when the United Kingdom adopted premarket screening for efficacy, and "1" after that (Grabowski et al. 1978).
21. Wardell's papers (e.g., Wardell 1973), were widely cited but did not receive serious attention in the economics literature. This may be because the papers did not employ any serious statistical analysis to probe the findings.
22. Because the United Kingdom still had shorter approval times, Grabowski et al. (1978) were still able to use the UK data as a control, although they used approval times as a measure of FDA regulatory rigor.
23. For example, May et al. (1983) examined the number of NCEs that reached the stage of clinical testing. Cullen (1983) used companies' ratings of different countries' regulatory systems so that countries other than the United Kingdom might be used as controls.
24. Carpenter et al. (2008) provided another form of disparate impact of FDA regulation. They showed that the FDA takes longer to approve later drugs, giving early entrants a regulatory advantage. They found that an increase of one standard deviation in the log order of entry increased FDA approval time by 3.6 months. This gradient was increased by the 1962 Amendments but was unaffected by PDUFA.
25. Another important insight in the Dranove and Meltzer study (1994) was that FDA regulation might affect not only approval time but the amount of time required for drug development. The higher the FDA standard, the more time companies have to spend investigating a drug to prove whether it meets the higher standard. Therefore, Dranove and Melzer looked at the total time from patent filing to approval for more and less important drugs.
26. Hutt et al. (2007) reported, however, that Congress reduced its funding for the FDA as user fees grew, so that total funding did not grow as fast as user fees did.
27. The deadline clearly had some effect. Carpenter et al. (2008) showed that PDUFA caused the FDA to make many more judgments on drugs during the two months before the statutory deadline.
28. The two major sources of data are The Tufts Center on Drug Development and PharmaProjects.
29. Moreover, although both drugs and devices are subject to suits alleging manufacturing defects, these suits rarely win because the FDA regulates and inspects manufacturing processes at drug and device plants. This FDA regulation eliminates much of the risk from manufacturing defects.
30. Wyeth is not the final word on preemption. For example, a state suit that required a label with which the FDA directly disagreed might be preempted. The case does, however, stand as a significant hurdle to preemption.
31. Under §510(k), a new class III device that is "substantially equivalent" to a device that was marketed before the 1976 Medical Device Amendments to the FD&C Act may be marketed until the FDA sets guidelines for judging the safety and efficacy of both the pre-1976 device and the new device.
32. Indeed, Kolstad et al. (1990) commented that the delay between when a product is made or sold and when a court decides whether the product is safe may be a disadvantage. This delay may generate uncertainty about the court's standard for liability, which may induce the manufacturer to take insufficient care.
33. Another concern is that different state courts may arrive at different judgments about a medical product. This patchwork of regulations increases production costs. A manufacturer typically must comply with the most restrictive state court judgment to market its product economically to all states. Yet there is no reason to think the most restrictive state court judgment is sounder than less restrictive state court judgments, even if one believes that state court judgments are superior to FDA judgments.
34. This is akin to the tradeoff between prevention of a disease and ex post treatment of that disease.
35. One hurdle may have been the so-called privity rules that allowed only the direct buyer of a vaccine (usually a medical provider or insurance company) to sue the seller (the manufacturer).
36. Manning adjusted for the effect of market power on price by controlling for the number of manufacturers in his price regressions.
37. Helland et al. (2010) have a working paper that examines the effect of punitive damages caps on drug prices and side effects from drugs. The period of analysis is 1996-2007. Unlike Manning (1997), they based identification on variation in tort exposure of individual drugs due to variation in the geographic distribution of sales across US states and the timing of tort reforms in different states. The variation in geographic distribution in sales is in turn caused by variation in the geographic distribution of diseases. Because the analysis is preliminary, the results are not reported here.
38. He found no effect when he approximated liability with the number of suits, whether or not successful. He also found no effect on price differential when he approximated liability with survey or published results on the side effect profile of

each drug. The latter result is consistent with the lack of failure-to-warn liability when users already know about the side effects of drugs.