The Generic Drug Trilemma

Daniel J. Hemel
University of Chicago Law School

Lisa Larrimore Ouellette
Stanford Law School

April 1, 2022

Abstract:
More than 90 percent of prescriptions dispensed in the United States each year are for off-patent drugs. Yet the bulk of scholarship on prescription drug policy focuses on patented-protected drugs. Discussions of prescription drugs are typically oriented around the “innovation-access dilemma”—the tradeoff between stronger patent-based incentives for innovators and higher prices for purchasers of patent-protected products. But for drugs in the “patent afterlife”—the period after patent protection and other forms of market exclusivity have expired—the innovation-access dilemma is not the fundamental policy tradeoff. Higher prices do not necessarily redound to the benefit of innovators, and price is not the only significant impediment to access: drug shortages—often triggered by safety concerns—also prevent patients from obtaining the medicines they need.

This chapter seeks to provide scholars and policymakers with a unifying framework for analyzing the variegated challenges of the pharmaceutical patent afterlife. We argue that the key tradeoffs in off-patent drug policy take the form of a trilemma—the three corners of which are price, quantity, and quality (encompassing efficacy as well as safety). The ideal for off-patent drug policy is to facilitate (1) low prices and (2) sufficient quantities of drugs that are (3) equivalent or similar to brand-name drugs approved by the Food and Drug Administration. But policies that satisfy one or two of those goals inevitably sacrifice the third.

The trilemma framing yields several analytical payoffs. First, it sheds light on the root causes of puzzling problems plaguing some off-patent drug markets, such as sudden price spikes and persistent mismatches between supply and demand. Second, it draws attention to the non-innovation costs of recent drug pricing reform proposals—costs that are often overlooked amid the emphasis on innovation versus access. Finally, it motivates a search for solutions that can transcend the trilemma and achieve all three goals of off-patent drug policy simultaneously. The search leads to no silver bullet, but it potentially points to a larger role for the federal government in generic drug procurement.


1 Introduction

The innovation-access dilemma is the dominant theme in discussions of patent law and policy (e.g., Landes and Posner 2003; Barnes 2010; Sampat and Williams 2019). Patent protection incentivizes innovation by offering a time-limited monopoly for novel and nonobvious inventions, but the high prices facilitated by patent monopolies limit access to knowledge goods. The stakes on both sides of the innovation-access tradeoff are particularly high in the prescription drug context: pharmaceutical innovation has fueled significant health and longevity gains in recent decades (Lichtenberg 2019), but high prices prevent millions of patients in and beyond the United States from accessing medicines they need (WHO 2016; Kearney et al. 2021).

In prior work, we have argued that the innovation-access dilemma is, in fact, less of a dilemma than the conventional framing lets on. First, policymakers can avoid the tradeoff by replacing patents with non-patent rewards, including grants, tax credits, and prizes (Hemel and Ouellette 2013). Second, even without eschewing intellectual property, policymakers can sidestep the innovation-access tradeoff by “matching” patent-based rewards with non-patent allocation mechanisms such as government procurement (Hemel and Ouellette 2019). The Covid-19 experience offers an illustration of “matching” in action: while preserving patent protection on the incentive side, the federal government provided all Americans with free access to vaccines through procurement contracts with Pfizer, Moderna, and Johnson & Johnson.

But whether one accepts or rejects the claim that the innovation-access dilemma is the fundamental policy problem in patent law generally or pharmaceutical patents specifically, it is clearly not the fundamental policy problem in the market for generic drugs, which constitute 90 percent of prescription drugs dispensed in the United States (Association for Accessible Medicines 2021).¹ Once drugs lose the protection of patents and other forms of market exclusivity—in the period we call the “patent afterlife”—many of the profits flow to firms that had no role in the development of the drug in question. While patients still often pay high prices for off-patent drugs, high prices paid to firms that did not develop the relevant drug are not part of society’s bargain for more innovation.

To be sure, drug policy in the patent afterlife is not free from tradeoffs. The tradeoffs, though, are different from—and potentially more difficult than—the innovation-access tradeoff in patent law. And while these tradeoffs have garnered much less attention than the innovation-access dilemma, they have significant implications for access to medicines in the United States.

¹ “Off-patent” drugs constitute even more than 90 percent of prescription drugs dispensed. The category of off-patent drugs includes not only generics, but also biosimilar versions of biologic products as well as branded drugs that are no longer patent-protected.
This chapter argues that the key tradeoffs in generic drug policy take the form of a trilemma—the three corners of which are price, quantity, and quality (encompassing efficacy as well as safety). As with other trilemmas in economics—e.g., the classic macroeconomic trilemma (Fleming and Mundell 1964) and the political trilemma of the world economy (Rodrik 2000)—the trilemma framing captures the impossibility of achieving all three goals at the same time (here, low prices, sufficient quantities, and a safety and efficacy profile that mirrors the profile of originator drugs). Policy solutions that satisfy one—or even two—of the objectives identified by the trilemma are feasible, but solutions that adequately address all three remain elusive. We can have a cheap and plentiful supply of generic drugs, a plentiful supply of high-quality generic drugs, or a cheap supply of high-quality generics, but the abiding challenge of generic drug policy is the challenge of solving for all three variables simultaneously.

Start with the corner of the trilemma that has recently attracted the most policy attention: price. Policies targeted at reducing prices, such as price caps, are also likely to have negative effects on quantity—including a higher risk of drug shortages—because they reduce incentives for generic entry and increase incentives for exit. (As discussed below, the U.S. and Canadian experiences largely bear out this prediction.) Reciprocally, policymakers could minimize the risk of shortages by offering to reimburse off-patent drug manufacturers at high rates, but those generous reimbursements would (of course) come at the expense of the first corner of the trilemma.

The third corner of the trilemma—quality—implicates a similar quandary. With respect to small-molecule drugs, the U.S. Food and Drug Administration (FDA) requires generic manufacturers to demonstrate “bioequivalence” with the brand-name product. Firms seeking to market the equivalent of generics for off-patent biologics must satisfy even more demanding requirements. The FDA also requires firms to adhere to rigorous manufacturing protocols. These prerequisites for approval raise entry barriers that reduce competition. That lack of competition, in turn, leaves the market for off-patent drugs more vulnerable to sudden price hikes and to persistent mismatches between supply and demand.

At least in theory, policymakers could achieve any two of the three trilemma objectives. For example, policymakers could achieve both low prices and sufficient quantities by reducing entry barriers for off-patent drugmakers. But those entry barriers exist to protect safety and efficacy, so reduced regulatory scrutiny would entail a tradeoff with drug quality. Similarly, policymakers could satisfy the price and quality corners of the trilemma by combining price caps with rigorous safety and efficacy evaluations. But the combination of low prices and high entry barriers would exacerbate the risk of shortages. To satisfy both quantity and quality objectives, policymakers could combine thorough regulatory scrutiny with price guarantees for manufacturers who meet those high standards—but then the goal of low prices would go unrealized.

Our trilemma framing is not, to be sure, an axiom of futility: some reforms will improve outcomes along two dimensions much more than they cause harm on the third. What the
trilemma framing accomplishes is to highlight the inevitability of compromise in off-patent drug policy. It thus helps to explain how some of the puzzling outcomes in off-patent drug markets—sudden price spikes, persistent shortages, and safety lapses—trace back to choices by policymakers to prioritize (rightly or wrongly) other corners. It also sheds light on the non-innovation costs of recent drug pricing reform proposals, including legislation that would limit prices paid by Medicare for certain drugs. And it inspires reflection on approaches that might alleviate some of the tension among the price, quantity, and quality objectives, such as a more active government role in prescription drug procurement.

In Section 2, we describe generic pharmaceutical markets in the United States, where we use “generic” in the broadest sense to include all products in the patent afterlife, including generic versions of traditional small-molecule drugs, the equivalent of generics for more complex biologic drugs (known as “biosimilars”), and off-patent brand-name drugs. Section 3 introduces the trilemma. Section 4 considers generic drug policies focused on antitrust enforcement and government procurement, explaining why both areas offer promise but neither provides a complete solution. Section 5 concludes.

2 The U.S. Pharmaceutical Patent Afterlife

When a firm has a new product that it thinks could treat or prevent diseases, it cannot just start selling the product to patients or doctors. Rather, the U.S. pharmaceutical market is strictly regulated by the FDA. If the product is a small-molecule drug—one with a simple chemical structure that can be well characterized—the drug sponsor seeks FDA approval through a New Drug Application (NDA). For biologic products—more complex structures that are often derived from living material, such as monoclonal antibodies and vaccines—the equivalent of an NDA is a Biologics License Application (BLA). The NDA or BLA must contain evidence that the product is safe and effective for its intended use, that its benefits outweigh its risks, and that it will be manufactured in a controlled way that will preserve its quality. To generate evidence in support of these applications, the sponsor must conduct costly clinical trials, with recent cost estimates ranging from $648 million (Prasad and Mailankody 2017) to $2.6 billion (DiMasi, Grabowski and Hansen 2016).

By contrast, a generic manufacturer that wants to enter these markets does not need to conduct new clinical trials from scratch. Instead, the firm may seek accelerated approval at the FDA under the 1984 Hatch-Waxman Act for small-molecule drugs and the 2009 Biologics Price Competition and Innovation Act (BPCIA) for biologics. Under these regulatory pathways, a generic entrant may demonstrate that its product is sufficiently similar to a brand-name product to rely on the original product’s clinical trial data, rather than conducting a full set of new clinical trials of its own.

For small-molecule drugs, a generic entrant typically files an abbreviated New Drug Application (ANDA). Because small-molecule drugs generally comprise only 20 to 100 atoms, laboratory measurements can be used to show that a generic is chemically...
equivalent to the brand-name drug; the regulatory standard is that the ANDA must contain evidence of “bioequivalence,” meaning that there is no “significant difference in the rate and extent to which the active ingredient … becomes available at the site of drug action.” Hatch-Waxman also created the 505(b)(2) NDA pathway for hybrid products that depend on both an existing drug’s clinical trial data plus new clinical trial data. These “brand-name non-originators” are typically drugs that have the same active ingredient as an already approved drug but a different formulation or delivery mechanism, which are eligible for three years of data exclusivity before other firms can rely on the new clinical trials.

For biologic drugs, the prospective entrant files an abbreviated Biologics License Application (aBLA). The greater complexity of biologics makes it more difficult to demonstrate that a second product will have the same clinical effect, which is why generic biologics are termed “biosimilars.” To show biosimilarity, an aBLA must provide evidence that the product’s characteristics are “highly similar” to the original biologic and that it has “no clinically meaningful differences” in clinical response studies (which are less costly than the original clinical trials). If the aBLA also demonstrates that there is little risk of switching between the original biologic and the new biosimilar, the biosimilar is deemed “interchangeable” and may be substituted at a pharmacy in the same manner as a generic drug. The scientific differences between small-molecule and biologic drugs are reflected in different development costs, with small-molecule generics typically taking 1 to 3 years and $1 million to $5 million (with no new clinical trials), and biosimilar development typically taking 8 to 10 years and around $100 million (including tens of millions in clinical trials to confirm the biosimilar’s clinical effects) (Atteberry et al. 2019).

FDA scrutiny does not end once a brand-name drug or generic is on the market. In addition to imposing the rigorous entry barriers described above, the FDA subjects drug companies to ongoing high quality standards through its Current Good Manufacturing Practices (CGMP) regulations. Among other requirements, pharmaceutical manufacturers must test each batch of medicine to ensure consistent production. The FDA also inspects some facilities and tests some drugs in its own labs.

To encourage brand-name firms to invest in drug development without the concern that the market will be immediately undercut by generic competitors, new pharmaceuticals are typically protected by a period of market exclusivity. Patents provide developers with a right to exclude generic competition for twenty years from the patent filing date—usually resulting in less than twenty years of actual market exclusivity because some of the patent term is consumed by the time spent on clinical trials and regulatory approval (Budish, Roin, and Williams 2015). Firms often prolong patent life through “evergreening,” or filing later patents on secondary innovations related to their drugs, although these patents tend to be lower quality (Frakes and Wasserman 2020) are more likely to be held invalid or not infringed during litigation (Hemphill and Sampat 2013). In the United States, Congress has also supplemented patents with additional regulatory exclusivity periods tied to FDA approval, including a five-year period for new small-molecule drugs with simple chemical
structures and a twelve-year period for more complex biologic drugs like monoclonal antibodies and vaccines (Thomas 2017).

Overall, actual market exclusivity for small-molecule drugs averages about thirteen years (Durvasula et al. 2022; Lietzan and Lybecker 2020). The exclusivity period for biologic drugs is more difficult to calculate due to the limited patent transparency, but in theory, it should last no longer than twenty years after the biologic is placed on sale (Price and Rai 2019). We are agnostic on what the optimal period is for exclusivity-based rewards; patent terms and regulatory exclusivity periods may not be well tailored for many pharmaceutical innovations. But in this chapter, we are focused on the period after the expiration of patents and other exclusivities: the patent afterlife.

Historically, most ANDAs targeted drugs that were no longer covered by patents (FTC 2002). However, to address the problem of patents that were improperly granted or that are being improperly asserted as covering the drug at issue, both Hatch-Waxman and the BPCIA provide incentives for generic and biosimilar developers to seek approval before a brand-name manufacturer’s patents have expired. A pre-expiration ANDA must include a “paragraph IV” certification stating that the relevant patents are invalid or that the brand-name manufacturer’s patents aren’t infringed by the generic product. Hatch-Waxman awards the first paragraph IV filer with 180 days of generic market exclusivity once its ANDA is approved, meaning that the FDA will not allow a second generic on the market during this time. The BPCIA similarly incentivizes market entry by providing a 12 to 42 month biosimilar market exclusivity period to the first interchangeable biosimilar (with the variation depending on the state of patent litigation).

The Hatch-Waxman Act is generally viewed as a success in increasing generic entry and lowering prices for small-molecule pharmaceutical drugs. In the Act’s first ten years, the FDA approved over 4000 generic versions of small-molecule drugs, with prices dropping by over 70 percent for drugs with four or more competitors (Heled 2021). But despite these successes, there are hundreds of small-molecule drugs that are no longer protected by patents or other forms of market exclusivity but that have no generic competitors (FDA 2021a), and many more with limited competition.

In contrast to Hatch-Waxman’s qualified success, the BPCIA is generally viewed as failing at its goal of increasing access to biologics in the United States. Heled (2021) has documented that in the BPCIA’s first ten years, only sixteen biosimilars were actually available on the market, and because some referenced the same biologic, only seven biologic drugs actually faced competition, with price declines of 10 to 37 percent. And for some biologics like vaccines, the FDA has not even issued regulations for how a firm would show that its vaccine is biosimilar to an existing vaccine. In other words, there is no such thing as a generic (or biosimilar) vaccine as a regulatory matter. Competition in vaccine markets thus requires an independent BLA based on new clinical trials.
Conti and Berndt (2020) examine unique markets for off-patent small-molecule and biologic drugs from 2004 to 2016 and conclude that after a drug loses patent protection or other exclusivity, the median number of manufacturers in each market is typically only two or three. Over half of generics only have one supplier (National Academies 2018). As a number of scholars have noted, the generic markets attracting the least competition—including small-molecule drugs with limited markets (Scott Morton and Boller 2017) and most biologics (Atteberry et al. 2019)—often have a particularly high ratio of entry costs to available profits, giving them characteristics of natural monopolies. And the low number of entrants in most generic markets limits price competition: more than two generic producers are typically needed to yield an average price drop of over 50 percent, and drops over 80 percent typically take five or more generic entrants (Conrad and Lutter 2019).

Table 1 summarizes the four main categories of drugs we consider in this chapter. We will use the term “generic” to refer (generically) to all four, though some definitions of “generic” would apply only to the first.

### Table 1. Types of Drugs in the Patent Afterlife

<table>
<thead>
<tr>
<th>Type of Drug</th>
<th>Regulatory Pathway</th>
<th>Exclusivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>small-molecule generic</td>
<td>Hatch-Waxman ANDA</td>
<td>180 days for first generic to successfully challenge patents</td>
</tr>
<tr>
<td>brand-name non-originator</td>
<td>Hatch-Waxman 505(b)(2) NDA</td>
<td>usually 3 years data exclusivity</td>
</tr>
<tr>
<td>biosimilar</td>
<td>BPCIA aBLA</td>
<td>12 to 42 months for first interchangeable biosimilar</td>
</tr>
<tr>
<td>off-patent brand-name</td>
<td>NDA or BLA with expired patents/exclusivities</td>
<td>no formal exclusivity</td>
</tr>
</tbody>
</table>

### 3 Three Goals of Generic Drug Policy

Here, we introduce three key goals of generic drug policy—low prices, adequate quantities, and high quality—and explain why they form a trilemma: policy reforms are able to satisfactorily achieve, at most, two out of these three goals. Finally, we distinguish trilemma-related concerns from other generic drug policy tradeoffs.

#### 3.1 Price

The most common complaint regarding generic drugs in the United States is their high price (e.g., Rosenthal 2014). In recent years, large and abrupt price increases for generic...
drugs have become front-page news (e.g., Pollack and Tavernise 2015). From the first quarter of 2010 to the first quarter of 2015, more than a fifth of established generic small-molecule drugs (315 out of 1441) experienced a sudden price increase of 100 percent or more (GAO 2016). In one of the most notorious examples, Turing Pharmaceuticals—then run by the since-convicted “pharma bro” Martin Shkreli—hiked the price of the six-decade-old anti-parasite drug Daraprim by 5000 percent (Lupkin 2019).

Yet cross-country comparisons complicate the narrative of high U.S. generic drug prices. A RAND Corporation study commissioned by the U.S. Department of Health and Human Services found that as of 2018, U.S. unbranded generic drug prices were 16 percent lower than prices in other countries within the OECD (Mulcahy et al. 2021). The same study found that unbranded generic drugs accounted for a much larger share of drug volume in the United States than elsewhere in the OECD (84 percent versus 35 percent). In other words, relative to other high-income countries, U.S. purchasers are paying less for unbranded generics and buying more of them.

The RAND study’s findings do not imply that high prices are a spurious concern for off-patent drug policy. First, the cross-country comparison excludes biologics, which now constitute more than two-fifths of total U.S. pharmaceutical spending (IQVIA 2020). (As noted above, biologic drugs do not technically have “generics”—a drug designed to have the same clinical effect as a branded biologic is a “biosimilar.”) Second, the comparison also excludes brand-name non-originators, such as drugs approved by the FDA through the hybrid section 505(b)(2) pathway. Once all small-molecule non-originator drugs are included in the cross-country comparison, U.S. purchasers pay—on average—21 percent more than OECD peers. Third, the category of unbranded generics does not include off-patent drugs that do not (yet) face a generic competitor, for which prices may remain high while effective market exclusivity persists. Finally, for some small-molecule drugs (like the estrogen derivative estradiol), prices remain high even after generics enter the market (Thomas 2018).

3.2 Quantity

A second area of concern regarding generic drugs in the United States is quantity. As of March 2022, the FDA classified more than 120 drugs as “currently in shortage” (FDA 2022). Although shortages can arise with respect to on-patent and off-patent drugs, the FDA reports that two-thirds of drugs that went into shortage during a five-year period from 2013 to 2017 were drugs with a generic version on the market (FDA 2020). Shortages are especially common with respect to sterile injectables, such as anesthetics and chemotherapy treatments (Yurukoglu, Liebman, and Ridley 2017).

From an economic perspective, the very idea of a “shortage” is somewhat puzzling: unless the supply and demand curves are both vertical lines, there must be some price at which they intersect. Within a Marshallian supply-demand framework, there are at least three ways to make sense of widespread reports of drug “shortages.” First, demand for the
relevant drug may be inelastic. For example, insulin is often cited in microeconomics textbooks as an example of a product with inelastic demand (e.g., Brown 1995; Cowen and Tabarrok 2009). If demand is entirely unresponsive to price, then a temporary supply shock may result in disequilibrium. Second, purchasing practices may introduce rigidities that prevent prices from rising to market-clearing levels. For example, most medical centers acquire drugs through group purchasing organizations (GPOs), which act as intermediaries between purchasers and manufacturers (Bruhn, Fracica, and Makary 2018). Long-term contracts between GPOs and manufacturers—including contracts that lock a GPO into using a particular manufacturer as its sole source for a drug—may prevent prices from responding rapidly to changes in demand and supply. Third and finally, the term “shortage” might be understood to refer broadly to supply shocks that cause quantity to fall and price to rise, even if there is no extended period of disequilibrium between supply and demand. Although “shortage” is arguably a misnomer (i.e., the market clears at the new, higher price), the result is still that patients cannot obtain drugs at prices that prevailed prior to the shock.

Whether or not drug shortages are “[r]eal shortages” (Stomberg 2018) in the economic sense, the health costs of generic drug supply shocks are potentially significant. In 2018, the American Medical Association adopted a policy declaring drug shortages to be an “urgent public health crisis” (American Medical Association 2018). In several surveys, physicians and pharmacists report higher rates of adverse drug outcomes and even patient deaths due to drug shortages (for a literature review, see Phuong et al. 2019). Retrospective cohort studies document negative clinical outcomes associated with specific shortages. For example, Vail et al. (2017) find that patients with septic shock in hospitals affected by a 2011 norepinephrine shortage experienced higher rates of in-hospital mortality. Gross et al. (2017) find higher rates of C. difficile infection among patients at hospitals that switched to other antibiotics during a nationwide shortage of piperacillin/tazobactam in 2014.

Not all shortages are associated with negative clinical outcomes. For example, Trifilio et al. (2013) find no reduction in remission rates among patients with acute myeloid leukemia who switched from the chemotherapy drug daunorubicin to an alternative (idarubicin) during a daunorubicin shortage. Indeed, there is some evidence that older patients benefited from the shortage-induced switch. Moreover, the number of new shortages appears to be on the decline. According to the FDA, new drug shortages fell from a high of 251 in 2011 to 43 in 2020, notwithstanding Covid-19-related supply chain disruptions (FDA 2021b).

---

2 Notwithstanding insulin’s utility as a textbook example of inelastic demand, many patients actually appear to be price-sensitive: a study of patients at the Yale Diabetes Center found that more than a quarter had cut back on insulin use due to cost (Herkert et al. 2019).

3 For a critique of the FDA’s methodology for identifying shortages, see Lutter (2022). The University of Utah Drug Information System reports a higher number of shortages but a similar trend: from a high of 267 new shortages in 2011 to 129 in 2020 and 114 in 2021 (ASHP 2022).
3.3 Quality

A final area of concern regarding generic drugs in the United States is quality, including both safety and efficacy. From a theoretical perspective, there are several reasons to be concerned about the quality of generic drugs. While these concerns may loom larger in theory than in practice, the theoretical perspective reminds us why the quality of generics should not be taken for granted.

One source of concern is the extreme quality uncertainty in the prescription drug market. Patients have no real way to detect manufacturing defects even after taking a drug. In the absence of any manufacturing defect, prescription drugs typically are less than 100 percent effective and produce adverse reactions in a subset of patients. A patient’s own experience is thus minimally informative about potential manufacturing flaws. Quality uncertainty potentially leads to a “market for lemons,” in which manufacturers of high-quality products cannot monetize their quality investments (Akerlof 1970). Those manufacturers may respond by diluting their quality standards or by exiting the market.

Second, reputational mechanisms—which address quality uncertainty in other markets—are particularly weak in the market for generic drugs. All states require or permit pharmacists to substitute a cheaper available generic when a doctor’s prescription specifies a brand-name drug (Sacks et al. 2021). Patients do not necessarily choose—and may not even know—the identity of the generic manufacturer. The low salience of reputation reduces incentives for generic manufacturers to invest in quality.

Third, no single generic firm internalizes the full costs of manufacturing defects. For example, the infamous Cutter incident—in which live polio virus contaminated 120,000 doses of the Salk polio vaccine in 1955—shook confidence in (and reduced uptake of) other vaccine manufacturers’ products as well (Oshinsky 2006). Acting on its own, a generic manufacturer may choose a level of investment in quality assurance that is lower than the optimal level from the entire industry’s (or society’s) perspective.

Fourth and finally, ex-post liability for manufacturing defects cannot resolve all the problems of quality uncertainty in the market for generic drugs. Some defects will likely go undetected, since patients and physicians will not be able to distinguish the consequences of manufacturing defects from the normal operation of a non-defective drug. Moreover, some of the social costs of manufacturing defects—for example, reduced uptake of other manufacturers’ products—likely won’t be recoverable in tort. And all this is on top of familiar shortcomings of limited liability—specifically, the fact that limited liability shields actors from the full costs of their torts (Leebron 1991).

Anecdotal evidence partly bears out these theoretical concerns, In 2013, generic manufacturer Ranbaxy accepted a $500 million fine for safety violations at two of its Indian factories, including delays in reporting “unknown impurities” detected in an epilepsy drug that eventually led to a 73-million-pill recall (Thomas 2013). Discoveries of the likely
carcinogen NDMA have led to major FDA-initiated recalls of off-patent drugs to treat heartburn (Johnson 2019), high blood pressure (Edney, Berfield, and Yu 2019), and diabetes (Blankenship 2020a). Investigative journalists have also uncovered anecdotes of safety violations at generic manufacturing plants, such as fabrication of data for regulators (e.g., Eban 2019; Stockman 2021).

Evaluating the actual health costs of these quality failures is more challenging: contaminants of unknown risk and reporting failures do not directly translate to patient harm, and these problems may be episodic rather than systemic. It is also unclear whether quality problems are substantially higher for generics than for brand-name drugs—quality concerns can also strike before pharmaceuticals enter the patent afterlife. Although critics of the generics industry point to the outsourcing of most generic manufacturing to India and China and problems with the FDA’s foreign inspection program, many brand-name drugs are also sourced from overseas factories. The FDA does not publicly disclose where a drug was made because this information is considered a trade secret, which makes it hard to research the impact of manufacturing origin.

Controlled studies generally have not found significant differences in the clinical efficacy of generic and brand-name products (Kesselheim et al. 2008; Kesselheim et al. 2010). But we think these studies do not resolve the quality debate, for two reasons. First, as noted above, brand-name products may also have quality problems, and understanding the health consequences of these problems would require a comparison of real-world marketed drugs with the idealized, defect-free version. Second, some studies have found that switching from brand-name to generic drugs was associated with worse clinical outcomes and more adverse events (for a systematic review, see Straka, Keohane and Liu 2017). These results may be explained from negative perceptions of generics (Colgan et al. 2015) and related psychosomatic effects (Goldszmidt et al. 2018), but issues of perception can still impact the market for generic drugs. We are agnostic on whether the FDA has struck the right balance in regulating drug quality—the key point is simply that quality is not costless.

### 3.4 The Generic Drug Trilemma

Each of these three goals—low prices, sufficient quantity, and high quality—theoretically lies within reach. But efforts to advance one or two of these goals inevitably conflict with the third. We first consider policies that would satisfy one corner of the trilemma (i.e., would achieve one of the three goals), and then consider policy combinations that would satisfy two corners.

**Price.** The most straightforward way to reduce prices of prescription drugs is to impose price caps. For example, in 1998, the government of Ontario—Canada’s most populous province—introduced the so-called “70/90” regulations under which the first generic entrant was prohibited from charging more than 70 percent of the brand-name product price and subsequent generic entrants were initially capped at 90 percent of the first generic price (Zhang et al. 2016). Since 2018, several U.S. states have created prescription drug
affordability boards authorized to establish price ceilings for prescription drugs (Williamson 2021). The Build Back Better Act, which passed the House in late 2021, would effectively cap the price of insulin and a limited number of single-source drugs, including some off-patent drugs without generic competitors.

Price caps, to be sure, are not certain to achieve their immediate goal of reducing prices. For example, Anis, Guh, and Woolcott (2003) find that price caps set by the Ontario 70/90 regulations served as focal points for generic manufacturers: firms offered the maximum price allowable under the regulation even when they might have set lower prices in an unregulated market. But even well-designed price caps cannot escape from the law of unintended consequences. Price caps reduce incentives for generic entry and increase incentives for exit, as evidenced by increased generic exit in Ontario (Zhang et al. 2016). Policies laser-targeted at reducing prices are therefore likely to have negative effects on quantity (Scott Morton 2001; Rye 2012).

**Quantity.** For the same reason that price caps negatively affect quantity, the most straightforward way to address drug shortages is to guarantee higher prices. When Congress asked the FDA to analyze the “drug shortage crisis” in 2018, the agency concluded that the root cause of shortages was economic forces that lead to a “race to the bottom” in generic drug pricing (FDA 2020). But by construction, price guarantees would run counter to the goal of lowering drug prices. Quality issues, meanwhile, are the immediate reason for most drug shortages in the United States: according to the FDA, 62 percent of new drug shortages from 2013 to 2017 were triggered by quality issues (FDA 2020). Relaxing quality standards would thus reduce the number of shortages—but of course, safety and efficacy would suffer as a result.

**Quality.** To address concerns about generic drug quality, regulators could impose more rigorous quality requirements on new entrants, coupled with strict liability regimes targeting firms whose generics have manufacturing defects. Heightened ex ante scrutiny or ex post liability, however, raise the costs of supplying the market and thus are likely to undermine the other two corners of the trilemma by decreasing quantity and raising prices. For example, Atal, Cuesta, and Sæthre (2019) examine Chile’s introduction of bioequivalence requirements for generic drugs and found that this “stronger quality regulation decreased the number of drugs in the market by 25% [and] increased average paid prices by 10%.”

**Price Plus Quantity.** Just as a rigorous quality standard will increase both prices and the risk of shortages, policymakers could achieve both lower prices and higher quantities by reducing ex ante regulatory barriers and ex post liability risks. For example, Sachs (2019) proposes that after a price spike exceeding a certain threshold for a particular drug, the FDA could “preclear” generic manufacturers to make and sell that drug before completing the full ANDA or aBLA approval process. (One could imagine a similar “preclearance” mechanism triggered by drug shortages.) Yet as Sachs acknowledges, shortcutting the generic drug approval process “creates serious problems” for safety and efficacy. Similarly,
Cohen et al. (2019) note that the FDA already has statutory authority to authorize importation from Canada, and they suggest that the agency use this authority to allow generic drug imports when prices rise suddenly in the United States. (They also argue for expanding the FDA’s importation authority to include a select group of other countries beyond Canada.) Yet as Bruser and McLean (2014) note, the FDA’s safety standards are often more stringent than those of Canada’s pharmaceutical regulator. While U.S. patients may be better off on balance if the FDA were to allow generic imports in response to price spikes, the choice would not be tradeoff-free.

**Price Plus Quality.** Like price and quantity, price and quality are mutually realizable goals. For example, policymakers could combine binding price caps with rigorous safety and efficacy regulations. But the combination of price caps and stringent quality standards would heighten the risk of shortages. Exacting quality standards would raise manufacturer costs, spurring exit and deterring entry without high prices as an inducement.

Arguably, the U.S. childhood vaccine experience in recent years reflects the consequences of policies that prioritize price and quality over quantity. The 1993 Vaccines for Children program provides eligible children with covered vaccines at no out-of-pocket cost to their families. Eligible children are those who are uninsured or eligible for Medicaid, as well as children who receive vaccines at Federally Qualified Health Centers and all children who are members of an Indian tribe. More than half of young children in the United States are eligible for vaccines through the program (HHS 2020). Covered vaccines are all pediatric vaccines included on a list maintained by the CDC Advisory Committee on Immunization Practices. One provision in the 1993 law imposes price caps on preexisting pediatric vaccines: the U.S. Department of Health and Human Services is prohibited from paying more through the program for any vaccine than the May 1993 price adjusted for subsequent increases in the Consumer Price Index.4

Congress’s choice to cap prices for a significant portion of the childhood vaccine market without making any downward adjustments to quality standards puts the trilemma framework to the test. Sure enough, the United States subsequently saw a rash of childhood vaccine shortages. By 2004, according to Ridley, Bei, and Liebman (2016), eight of eleven routine childhood vaccines had gone into shortage, with the probability of shortages correlated to low prices for the relevant vaccine. Thirty-five states had temporarily suspended or reduced immunization requirements for daycare and/or school programs—evidently in response to supply constraints (GAO 2002).

The childhood vaccine experience illustrates the perils of transplanting the framework of an innovation-access tradeoff into the generic drug space, where a different policy tradeoff dominates. As Ridley, Bei, and Leibman summarize, one of the reasons why Congress imposed price caps through the Vaccines for Children program was a belief that it is “no

---

longer necessary to worry about incentives for innovation for these vaccines because the costs of innovation were paid long ago by manufacturers.” But in the patent afterlife, prices affect other dimensions as well. Specifically, when prices are constrained to be low but quality standards raise production costs, manufacturers are less likely to invest in maintaining and expanding supply. Whether or not Congress made the “right” tradeoff among price, quality, and quantity as part of the 1993 Act, the choice to cap prices without conceding on quality clearly was a tradeoff—though one not immediately acknowledged at the time.

**Quantity Plus Quality.** Finally, policymakers could ensure sufficient quantities and maintain rigorous quality standards if they were willing to guarantee high prices for generic drugs. In this respect, “money answereth all things” (Ecclesiastes 10:19). Manufacturers would have strong incentives to invest in spare capacity—and in measures to protect safety and efficacy—if profit margins on generic drugs were wide enough. But of course, the one objective that cannot be satisfied by throwing money at a problem is economizing on price.

Figure 1 illustrates the three corners of the trilemma and some of the policy reforms that satisfy one or two corners.

![Figure 1. The Generic Drug Trilemma](image)

### 3.5 Other Tradeoffs

Just as the innovation-access dilemma does not capture every relevant aspect of patent policy (such as the role of disclosure), the price-quantity-quality trilemma does not capture all aspects of generic policy. The lack of robust generic competition for many drugs can extend a brand-name developer’s expected period of market exclusivity, which has two
potential social benefits: increased ex ante incentives to develop the drug in the first place, and increased incentives to encourage demand promotion while the drug is on the market. But we think these are secondary issues to the trilemma, and that each has ambiguous welfare effects.

First, consider innovation incentives. An extensive literature has demonstrated that expected profits affect innovation in the pharmaceutical industry (e.g., Finkelstein 2004; Acemoglu and Linn 2004; Giaccotto et al. 2005; Yin 2008; Blume-Kohout and Sood 2013; Dubois et al. 2015), so the ability to extend the period of economic rents into a drug’s patent afterlife likely increases incentives to create the drug in the first place. In some cases, this might be a feature of weak generic competition rather than a bug. For example, because of low expected profits, the current system offers underpowered incentives to drugs such as treatments for early-stage cancers (Budish, Roin, and Williams 2015) and vaccines (Kremer and Snyder 2015), and further reducing the reward for these innovations with more robust generic entry would likely exacerbate these problems.

However, we see little reason to believe that delays in generic entry for a drug will be correlated with the extent to which the drug is under-valued by existing rewards. To the contrary, the drugs that are most likely to have natural monopoly characteristics include those that affect small numbers of patients (resulting in low profits) and biologics (which have higher entry costs), which already receive additional incentives. The Orphan Drug Act provides a seven-year market exclusivity period on top of additional tax credits and grants for drugs that treat rare diseases, and the BPCIA offers a twelve-year regulatory exclusivity period for biologics, which provides greater protection than the corresponding five-year period for small-molecule drugs. In short, there are more targeted ways to promote innovation in underincentivized drugs like cancer preventatives than hoping generic entry is slow.

Second, lack of generic entry could provide some benefits for drugs while they are on the market. Although market exclusivity typically reduces utilization through a monopolist’s incentive to artificially restrict quantity and raise prices, it can also increase utilization through the incentive to promote demand for the product, such as through marketing expenditures. Through an empirical analysis of the U.S. market for small-molecule drugs whose patents expired between 1990 and 2003, Lakdawalla and Philipson (2012) document that “in the short run, patent expirations reduce output and consumer welfare by decreasing marketing,” while “[i]n the long run, patent expirations benefit consumers, but by 30 per cent less than would be implied by the reduction in price alone.” In a similar analysis of drugs on the market from 2000 to 2004, Duflos and Lichtenberg (2012) conclude that “[p]rice and marketing expenditure both decline by about 50–60% in the years immediately following generic entry, but the number of prescriptions remains essentially constant during those years.” The short-term welfare effect of these marketing efforts depends on the product at issue: Purdue Pharma’s unprecedented marketing of OxyContin
contributed to the opioid epidemic, but Merck’s efforts to promote the HPV vaccine may have saved lives (Hemel and Ouellette 2020; Hemel and Ouellette 2023).

4 Partial Solutions to the Trilemma

Although policymakers cannot escape the price-quantity-quality trilemma, certain solutions will be better than others. In some cases, policy reforms can achieve gains for one or two goals that far outweigh the losses for the remaining corners. The trilemma also helps illustrate that no single goal is worth pursuing at any cost. Rather, ideally, society should seek gains in any one corner of the trilemma only up to the point that marginal benefit of those gains for social welfare equals the marginal cost of welfare losses at the other corners. Reaching this optimum, however, is easier said than done.

In some circumstances, competitive markets that treat drugs like commodities may be the best mechanism for balancing these tradeoffs, and policies such as greater antitrust scrutiny may help facilitate such competition. But policymakers should remember that competition is not an end in itself. For many drugs, rather than incentivizing generic manufacturers to incur large fixed costs in order to force brand-name manufacturers to lower prices, it may be more effective for government planners to directly contract for generics. Here, we describe antitrust and central planning approaches to generic drug policy and explain why neither is a complete solution to the trilemma.

4.1 Antitrust Enforcement

One consequence of the United States’ choice to allow unregulated pricing with restricted entry into the off-patent drug market is to leave the market vulnerable to anticompetitive conduct. Robust enforcement of antitrust laws can address some anticompetitive practices, but it does not offer a total escape from the price-quantity-quality trilemma.

Much of the literature on generic drug entry focuses on the period before patent protection expires—the patent life rather than the patent afterlife. A significant policy concern during a drug’s patent life is that the relevant patents may have been erroneously granted by the U.S. Patent and Trademark Office (Frakes and Wasserman 2019), or that they may not actually cover the drug in question. One of the innovations of the Hatch-Waxman Act—the paragraph IV process described in Section 2—was to create a route for generic drug manufacturers to challenge a patent as invalid or not infringed. The BPCIA followed Hatch-Waxman in incentivizing patent challenges. In effect, Congress has sought to enlist generic drugmakers as “patent police.”

Patent policing by generic manufacturers does implicate the innovation-access tradeoff. The paragraph IV process is designed to reveal instances in which brand-name drugs do not merit the innovation incentive that comes with patent protection. However, paragraph IV has not fully lived up to its designers’ high hopes. One threat to the integrity of the process
is the practice of “pay for delay,” in which a brand-name manufacturer pays or otherwise compensates a generic firm in exchange for the generic firm dropping a patent challenge (Hemphill 2006). A further threat is common ownership: when generic drug manufacturers and brand-name counterparties have the same institutional shareholders, they are more likely to settle paragraph IV disputes (Xie and Gerakos 2020). Finally, when generic drug manufacturers also have patented products in their portfolios, they are less likely to pursue paragraph IV challenges to judgment—possibly for fear of establishing precedents that will undermine their own patents (Carrier, Lemley, and Miller 2020).

Antitrust enforcement could potentially help invigorate patent policing by generic manufacturers. Stricter scrutiny of pay-for-delay deals might deter generic and brand-name manufacturers from entering anticompetitive agreements. Policy interventions to reduce common ownership (Elhague 2016; Posner, Scott Morgan, and Weyl 2017) might encourage more aggressive patent challenges under Hatch-Waxman and the BPCIA. Merger review might play a role in preventing brand-name and generic manufacturers from combining.

Our primary focus here, though, is not on the ways in which generic manufacturers fulfill their patent policing function, but on outcomes after the relevant patents have expired or been invalidated. Anticompetitive practices continue into the patent afterlife, but they differ from the practices associated with on-patent drugs. Often, anticompetitive practices in the patent afterlife exploit elements of the regulatory infrastructure designed to ensure the safety and efficacy of generic drugs. And the market for off-patent drugs is more vulnerable to anticompetitive conduct as a result of entry barriers erected with safety and efficacy in mind.

Historically, one mainspring of anticompetitive conduct in the patent afterlife has been the requirement—borne out of concerns regarding generic drug quality—that generic drug manufacturers demonstrate bioequivalence. In order to satisfy the requirement, a potential generic drug manufacturer must test its own product against samples of the corresponding brand-name drug. Some brand-name manufacturers sought to prevent potential generic competitors from performing those tests by refusing to provide samples of the brand-name drug (Pear 2018)—what one former Federal Trade Commission official has described as a “sample blockade” (Kades 2021). Turing Pharmaceuticals—the company once led by Martin Shkreli, who was later convicted of securities fraud—used this strategy to maintain its monopoly over the anti-parasite drug Daraprim long after all the relevant patents expired.5

“Safety protocol filibusters” (Kades 2021) are another anticompetitive practice that brand-name manufacturers have used to block generic entry during the patent afterlife. These “filibusters” involve drugs that are part of the FDA’s risk evaluation and mitigation strategies (REMS) program. For drugs that carry a risk of serious adverse effects, the FDA requires manufacturers to develop safety protocols to reduce the risk of adverse outcomes. For example, a REMS protocol might require healthcare professionals to complete a

---

training module and obtain certification before prescribing or dispensing a particular drug. Historically, the FDA has required brand-name and generic manufacturers of the same drug to develop a single, shared REMS system. The generic manufacturer could not begin to sell the drug until the single, shared REMS was in place. That requirement encouraged brand-name manufacturers to engage in foot-dragging—or “filibustering”—during negotiations to establish the shared REMS, thus delaying generic entry. The FDA had the authority to waive the single, shared REMS requirement, but the agency rarely used that power (Dabrowska 2018).

The CREATES Act of 2019 seeks to address both the sample blockade and safety protocol filibuster issues. The statute allows a potential generic entrant to sue to obtain samples of the brand-name drug. The statute also gives the FDA greater flexibility to allow generic manufacturers to establish their own REMS. Anecdotal evidence suggests that these reforms have been successful (Kades 2021). However, these are not the only anticompetitive practices in the patent afterlife. For example, brand-name manufacturers still may—and do—use “citizen petitions” to challenge generic drugmakers’ assertions that the brand-name and generic products are bioequivalent (Carrier 2018). These citizen petitions can delay generic approval for months or more even when they are ultimately denied.

Sample blockades, safety protocol filibusters, and sham citizen petitions all are distinct anticompetitive practices, but they are connected by a common thread: in each case, brand-name manufacturers exploit the fact that prices of off-patent drugs are unregulated while entry remains restricted. Those entry restrictions, in turn, are designed with safety and efficacy in mind. In other words, efforts to protect quality enable brand-name manufacturers to take actions that inflate price.

Even when anticompetitive conduct does not involve direct exploitation of quality protections, the entry barriers erected for quality-related reasons leave the market for off-patent drugs more vulnerable to collusion. For example, two executives at the generic manufacturer Heritage Pharmaceuticals pled guilty in 2017 to charges that they conspired with competitors to fix prices (DOJ 2017). A third executive at Sandoz, the generic unit of Novartis, pled guilty to related charges in 2020 (Blankenship 2020b). An ongoing set of lawsuits filed by forty-eight states alleges “rampant” collusion among generic manufacturers of prescription topical products (Stuart 2020). Quality controls—such as rigorous requirements to establish bioequivalence—facilitate this type of collusion by limiting the number of players in the market for any given drug and making it harder for new entrants to undermine existing cartels. In largely denying a motion to dismiss the states’ antitrust claims, a federal district court emphasized that the states “plausibly outline a regulatory regime” that would provide a motive for price fixing because “[h]igh barriers to entry . . . make an industry more conducive to collusion.”

---

Robust enforcement of antitrust laws—as well as legislative changes such as the CREATES Act—may help to alleviate some of the costs of a market with unregulated prices plus significant barriers to entry. Ultimately, though, antitrust enforcement is treating a symptom of the underlying choice to prioritize quantity and quality over price. As long as prices are unregulated and barriers to entry are high, antitrust enforcement will continue to play an important role in the off-patent drug market—but even the most rigorous antitrust scrutiny will not escape the trilemma. For example, in an empirical study of high off-patent drug prices in the United States compared with other countries, Ganapati and McKibbin (2021) note the greater market power of U.S. drug suppliers as a factor limiting price decreases in U.S. generic markets. But they also conclude that given the high fixed costs of entry, price controls are likely needed to substantially lower generic prices. Of course, capping prices without adjusting quality requirements will implicate concerns about quantity.

4.2 Government Procurement

So far, the main U.S. strategy for lowering generic drug prices has been to encourage competitive entry by generic manufacturers, including by empowering generic firms to act as “patent police” and by increasing antitrust scrutiny of generic markets. But as explained above, even robust enforcement of antitrust laws for off-patent drugs will still lead to a highly imperfect solution to the generic drug trilemma. Many generic markets will still have natural monopoly characteristics such that they fail to attract sufficient competitors to have a measurable impact on price. In some cases, rather than trying to create conditions for a competitive market, it may be better for the federal government to control generic markets through direct procurement.

With generic drug procurement, the government could specify the desired drug quality and quantity, and then could seek bids on price (e.g., Kolchinsky 2020). By overshooting on quantity and sourcing from a diverse set of suppliers, the government could minimize the risk that a sudden exit by one manufacturer or a quality control issue triggers a shortage. Long-term contracts could offer firms sufficient expected profits to overcome entry costs. In theory, policymakers could achieve the same outcome using quality regulation plus price controls, but trying to set a price that leads to the right quantity likely entails a higher epistemic burden than setting a quantity and negotiating on price.

This approach would not be unprecedented. The federal government is already a large direct purchaser of pharmaceuticals through agencies including the Department of Defense (DoD), the Department of Veterans Affairs (VA), the Indian Health Service, the Federal Bureau of Prisons, and the Department of State. The two largest purchasers, the DoD and the VA, spent nearly $15 billion on pharmaceutical procurement in 2018, which accounted for nearly 5 percent of total U.S. drug expenditures (CBO 2021). Federal pharmaceutical contracting has directly impacted an even larger number of Americans during the Covid-19 pandemic.
pandemic, when all Americans could freely access vaccines that the government procured from Pfizer, Moderna, and Johnson & Johnson.

Recently, some commentators have suggested expanding government procurement of generic versions of patented pharmaceuticals as a way to reduce their cost (Kapczynski and Kesselheim 2016; Morten and Duan 2020). In particular, the federal government effectively has a compulsory license to patented inventions under 28 U.S.C. § 1498, so a patent owner could not obtain an injunction to block federal procurement of infringing generic drugs during the patent term—the sole remedy would be an action for damages in the U.S. Court of Federal Claims. Brennan et al. (2016) note precedents from the 1960s, such as DoD’s purchase of an antibiotic from a generic supplier in Italy (where drug patents were not allowed) because its product was 72 percent cheaper than the brand-name product sold by Pfizer, the U.S. patentholder. But this use of § 1498 directly implicates the innovation-access dilemma, raising thorny questions about the appropriate framework for assessing damages in the Court of Federal Claims. We think federal pharmaceutical procurement is even more attractive in the patent afterlife, where these issues of appropriately valuing medical innovations have faded.

To be sure, federal procurement is not a panacea for the trilemma. First, it imposes a high informational burden on the government. The bidding process would aggregate market information about price, but setting quality standards, quantity targets, and optimal contract terms is far from trivial. A social planner with full information would invest in quality and quantity only up to the point that the marginal benefit of these gains for social welfare equals the marginal cost, but there is a weak evidence base for each of these variables in general, and the difficulty is heightened by the large number of generic drug markets. Second, the government would need to not just contract for a sufficient national quantity of each drug, but also distribute the right quantity of drugs to different providers throughout the country. Finally, government procurement also raises significant risks of mismanagement and wasteful spending (e.g., Liebman and Mahoney 2017).

We think the Covid-19 vaccine experience illustrates the promise of large-scale federal pharmaceutical procurement, even though the slow initial rollout shows how the government can be an imperfect drug distributor (Jha 2020). The federal government also successfully procured large quantities of drugs for treating Covid-19, but some medical providers and patients have been frustrated by the uneven distribution of remdesivir (Kolata 2020), monoclonal antibodies (Aleccia 2021), and—most recently—the antivirals Paxlovid and molnupiravir (Park 2022). These problems serve as a reminder that we should not compare the flawed competition-based system with an idealized vision of the public sector. The key policy question is whether the flawed public sector can still get closer to the optimal point within the generic drug trilemma.
## Conclusion

This chapter introduces the generic drug trilemma as a conceptual framework for evaluating policy reforms in the pharmaceutical patent afterlife. By highlighting the tradeoffs among three key goals of generic drug policy—low prices, adequate quantities, and high quality—we think the trilemma can help policymakers and commentators assess the relative importance of each goal, including the extent to which quantity and quality are worth paying for.

The goal of low prices has attracted the most recent policy attention for both on-patent and off-patent pharmaceuticals, but these discussions often conflate the price paid to the pharmaceutical manufacturer (the “price” corner of our trilemma) with the price paid by a patient. Out-of-pocket patient costs are important because they can impede access to medicines; in one recent survey, three in ten U.S. adults reported not taking medicines as prescribed at some point in the past year “due to cost” (Kearney et al. 2021). Reducing these cost barriers is a critical policy goal. Most other high-income countries have healthcare systems with more comprehensive coverage at lower patient cost than the United States (Emanuel 2020), and policies such as expanding Medicaid, reducing Medicare coinsurance rates, and capping out-of-pocket spending can move the United States in this direction.

But the price paid by a patient is a separate and independent policy choice from the price paid to a pharmaceutical manufacturer. Patients can pay zero cost through a universal insurance system even if manufacturers receive a high return. And separating the access question from the question of the optimal award to a manufacturer helps focus policy decisions on what manufacturers are being paid for. During the patent life, this policy choice is focused on innovation. The key challenge for pharmaceutical innovation policy is setting a reward aligned with a drug’s comparative effectiveness, which existing institutions often over- or under-value (Hemel and Ouellette 2023). For generic drugs in the patent afterlife, the trilemma highlights the way price reforms inevitably entail some tradeoff with quality and quantity.

How much should society pay for a sufficient quantity of high quality generic drugs? As we have explained, the optimal number of drug shortages and safety concerns is not zero: at some point, the price that would have to be paid to reduce the risk of any supply shocks and to inspect every pill would outweigh any benefit from that risk reduction. Rather, society ideally should pay for quantity and quality up to the point that the marginal benefit of those expenditures equals their marginal cost. The current U.S. approach to managing these goals is largely focused on maintaining high regulatory barriers related to safety and efficacy while encouraging generic entry and hoping that competition will both drive down price and provide adequate quantities. And in some cases, competitive markets—aided by robust antitrust enforcement—indeed may be the best way to balance these tradeoffs. But competition is just one of multiple tools to accomplish the underlying goals of cheap, plentiful, and high-quality drugs. Framing the patent afterlife in terms of the price-
quantity-quality trilemma helps illustrate ways in which the tensions among these goals may be better balanced by more active government procurement.

Finally, our analysis has focused on the U.S. context, but the price-quantity-quality trilemma is global in scope. One reason to begin with the domestic context is that the problems of price, quantity, and quality are more tractable in the United States, where resource constraints are less binding and a strong quality regulation regime already exists. Successful navigation of the policy trilemma in the U.S. context is likely to yield lessons for abroad. In the global context, access-to-medicines advocates have primarily emphasized one element of the trilemma—price—and have focused on removing barriers to generic entry. Yet serious quantity and quality concerns apply in less developed countries as well (WHO 2016; Chokshi, Mongia, and Wattal 2015). Our analysis of the U.S. context serves as a reminder that price is only one element of the global access-to-medicine challenge and that a holistic approach to access-to-medicine problems will need to be three-dimensional.

References


———. 2021b. *Drug Shortages for Calendar Year 2020.* [https://www.fda.gov/media/150409/download](https://www.fda.gov/media/150409/download)


