Center on Health Policy at BROOKINGS

FEBRUARY 2026

BOUND BY QUOTA:
DRUG SHORTAGE VULNERABILITY FOR SCHEDULE II MEDICINES

Marta E. Wosińska and Ben Graham

AUTHOR NOTES AND ACKNOWLEDGEMENTS

Marta E. Wosińska, Ph.D., is a senior fellow and Ben Graham is a senior research assistant at the Center on Health Policy at Brookings.

This project is sponsored in part by the Uniformed Services University of the Health Sciences (USU); however, the information or content and conclusions do not necessarily represent the official position or policy of, nor should any official endorsement be inferred on the part of, USU, the Department of Defense, or the U.S. government.

We would like to thank numerous individuals for helpful comments and discussions: Bill Grubb, Carlo de Notaristefani, John Gilbert, Richard Frank, and Amy Goldstein. We would also like to thank Erin Hu for fact-checking assistance and Rasa Siniakovas for editorial and publication assistance.

DISCLOSURES

Introduction

Drug shortages have been a persistent feature of the U.S. market for more than two decades, with headline-grabbing examples including amoxicillin, attention-deficit/hyperactivity disorder (ADHD) medicines, cancer therapies, and saline. Shortages lead to treatment delays, substitution to less effective alternatives, and in some cases, rationing. They carry clinical, financial, and psychological consequences for patients, families, and health care providers.

Drug shortages arise when supply chains cannot adjust quickly enough to demand or supply shocks, such as changes in disease incidence, manufacturing disruptions, or natural disasters. The likelihood that a given drug will experience a shortage depends on economic incentives and production structures, including how many manufacturers produce that drug, how profitable it is, and how much spare capacity or diversification firms are willing to maintain. Once a disruption occurs, the speed of recovery depends on how quickly existing manufacturers can ramp up output and how easily additional capacity can be brought online, which, in turn, depends on whether facilities have appropriate Food and Drug Administration (FDA) approvals.

This paper focuses on a lesser-known driver of drug supply chain vulnerability: the restrictive regulatory framework that governs certain controlled substances known as Schedule II drugs. These drugs have accepted medical uses in the United States, can be highly beneficial for appropriately selected patients, but also carry a high potential for abuse and dependence. They include ADHD medications, oral-dose painkillers (oxycodone and hydrocodone), and three intravenous (IV) opioids (morphine, fentanyl, and hydromorphone) that appear on the FDA Essential Medicines List and are critical hospital care drugs for pain management and sedation. At the time of this report's publication, FDA listed morphine, fentanyl, and hydromorphone in shortage, along with key stimulants for treating ADHD.

In addition to the economic pressures and FDA regulations that affect all prescription drugs, Schedule II medicines are subject to an additional layer of supply-side controls administered by the Drug Enforcement Administration (DEA). DEA's core mandate is a law enforcement function: to prevent the diversion of legally manufactured drugs into illicit or nonmedical channels. DEA has multiple enforcement and regulatory tools, but the quota system—which caps how much controlled substance can be manufactured—is its most far reaching supply-side control.

But quotas are a blunt instrument, and they distort access differently across Schedule II drugs. For ADHD stimulants, where a sizable share of use is inappropriate, tight national ceilings force clinically appropriate patients and non-medical users to compete for a fixed volume, with appropriate patients at risk of being crowded out. For hospital IV opioids, where inappropriate use is far less common, the same restrictive and inflexible framework leaves little slack to absorb manufacturing or supply disruptions, so shocks quickly deepen and prolong shortages.

This paper proceeds as follows. First, we lay out a high-level framework for assessing drug supply chain vulnerability. We then provide background on key manufacturing dynamics for Schedule II medicines, apart from DEA regulations, and describe how DEA's quota and inventory systems operate, grounding the analysis in statute, regulation, and agency practice. Then we apply the vulnerability framework to Schedule II drugs, discuss what it implies for the current oversight structure, and conclude with recommendations for reform within the existing statutory framework.

Conceptual framework for drug shortages

Drug shortages arise when supply chains cannot adjust quickly enough to demand or supply shocks. In turn, the severity of a shortage depends on how those shocks interact with the market's structure and the regulatory framework around it.

Building on prior work, we describe supply chain vulnerability along four dimensions and then examine the dynamics that shape each:

- Probability of a shock
- Size of the shock
- Ability to absorb the shock
- Ability to recover from the shock

The **probability of a shock** in prescription drug markets reflects both demand-side volatility and supply-side risks. Demand can change because of epidemics, new clinical uses, or shifts in treatment guidelines, while supply can be disrupted by manufacturing quality problems, plant closures, natural disasters, geopolitical events, or firms eliminating less profitable products from their portfolios. Some of these drivers are largely controllable—such as investments in quality systems, maintenance, and diversification of suppliers—while others, like hurricanes or earthquakes, cannot be prevented but can be managed, for example, through choices about where facilities are located.

The **size of a supply chain shock** depends not just on the disruption itself (for example, the strength of a hurricane hitting a manufacturing facility) but on the size of the disruption relative to the overall market. When a single facility that accounts for a substantial share of production goes offline, it can remove enough capacity to trigger a system-wide shortage. Panic buying can then further enlarge the effective shock: When buyers rush to stockpile affected products at the first sign of trouble, they pull demand forward and exhaust remaining inventories before manufacturers have time to adjust production, turning a localized disruption into a broader shortage.

A system's **ability to absorb shocks** depends on operational capacity and inventory buffers. Backup manufacturing arrangements (such as pre-validated secondary sites or lines), and inventories or stockpiles of finished product and key inputs can all bridge temporary disruptions, allowing shocks to be absorbed without creating visible shortages for patients or providers. In generic drug markets, low prices and thin margins discourage both investment in redundant capacity and holding of large inventories. Health system or government stockpiles can supplement the system's overall buffer, but those remain limited.

The system's **ability to recover from shocks** in prescription drug markets depends on how quickly manufacturers can either restore existing production or bring new capacity online and move product through the supply chain. Once buffers are exhausted and shortages emerge, the speed of recovery determines how long patients and providers experience constrained access and how long the clinical and operational harms from the shortage persist. The nature of the underlying shock also matters: A fixable equipment failure or quality deviation on an existing line is far easier to recover from than a permanent loss of capacity, or to recover from a situation that requires new firms to enter a market. The system's ability to recover is also shaped by technological constraints, regulatory requirements, and contracting arrangements.

We return to this framework after providing background on how Schedule II medicines are manufactured and how DEA quota and inventory systems operate.

Manufacturing process

Before discussing how the DEA quota and inventory system affects supply reliability for Schedule II drugs, this section provides background on how these medicines are manufactured, focusing on production timelines and the role of batch sizes in shaping them.

Schedule II painkillers and stimulants in the United States are small-molecule drugs, typically produced through chemical synthesis rather than biologic processes, even when derived from plant-based materials like opium or poppy straw concentrates. They are mainly formulated as oral tablets, capsules, or sterile injectables, with other forms such as sprays and patches used less often. Some products are available as extended-release formulations, and a subset incorporates abuse-deterrent technologies that slow or limit rapid release of the active ingredient or make manipulation for nonoral use more difficult, moderating both onset and psychotropic effects.

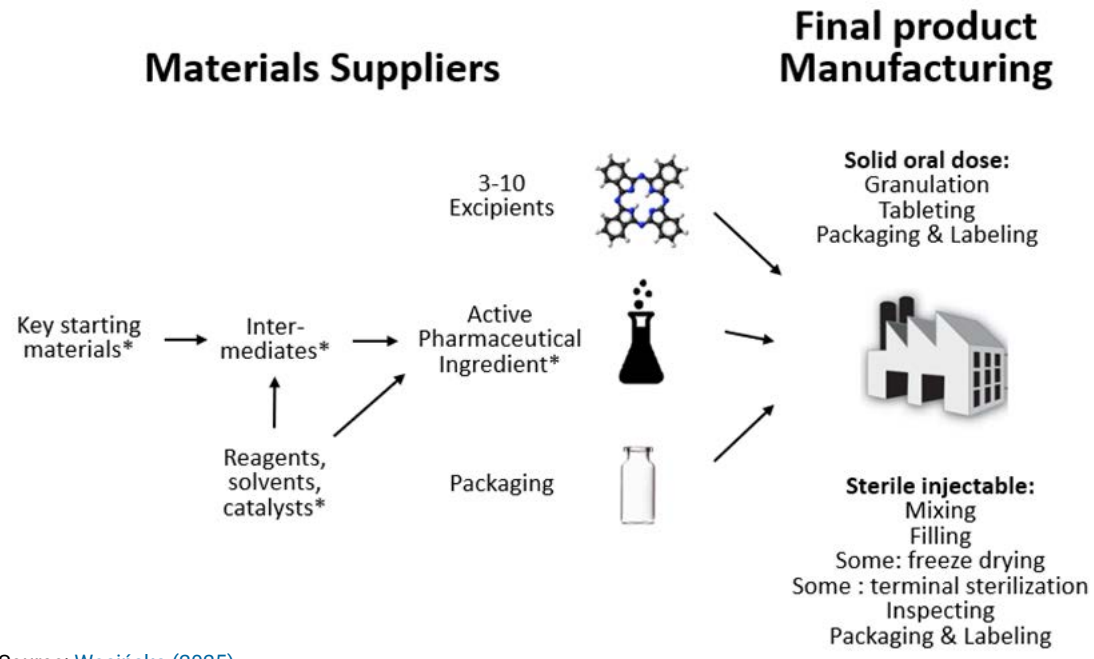
The manufacturing process for Schedule II controlled substances is similar to that for other small-molecule drugs and follows the sequence shown in Figure 1.

In this sequence, key starting materials (KSMs) are first converted, often with reagents, solvents, and catalysts, into one or more intermediates, which we refer to here as precursors. These precursors are then transformed through additional synthetic steps into the final active pharmaceutical ingredient (API). In the case of hydromorphone, for example, concentrated poppy straw enriched for oripavine serves as the KSM; it is first processed to isolate oripavine, and that oripavine intermediate is then converted through additional synthetic steps into hydromorphone, the final API.

At the formulation stage, manufacturers combine the active ingredient with excipients and process the mixture into finished dosage forms—tablets, capsules, or vials. Excipients are inactive ingredients that help control properties such as tablet hardness, dissolution rate, and stability, and it is at this stage that products are formulated as immediate-release versus extended-release tablets or capsules. Typical solid oral prod-

FIGURE 1

Inputs and manufacturing steps for small-molecule drugs



Source: Wosińska (2025)

ucts contain three to ten excipients, with abuse-deterrent formulations often relying on more complex excipient systems.

In some cases, Schedule II products are then handled by separate repackaging or relabeling firms that place the finished drug into new packages or under different labels. These firms are regulated much like manufacturers rather than simple distributors, because changing how a drug is packaged or labeled can affect its quality, traceability, and how it is used, even though the underlying formulation stays the same. Because repackaging and relabeling apply only to a subset of products and do not alter the underlying sequence from key starting materials to finished dosage forms, these activities are not shown in Figure 1.

Active ingredient manufacturing is typically carried out in batches, in which a series of processing steps—chemical reactions, separations, purifications, drying, and milling—are performed in a sequence on the same material. A batch has its own record and release testing. In practice, batches of the same product are typically run back-to-back as a campaign on the same equipment, with only limited cleaning between batches. For commercially supplied small-molecule active ingredients, the typical production time for a batch is about three to four months.

Manufacturing the finished dosage form introduces its own steps and timelines. Converting active ingredients into tablets, capsules, or sterile injectables packaged in vials can take several weeks to a few months for each commercial batch, depending on dosage form and testing requirements. Extended-release tablets typically require more complex formulations and additional testing than immediate-release products, while sterile injectables require multiple complex manufacturing steps and stringent microbiological testing, so injectable batches usually take longer to release, with total production time often exceeding six months.

Because production is organized into discrete campaigns on lines with limited capacity, batch- and campaign-size decisions sit at the center of how manufacturing realities translate into calendar time. Manufacturers usually select batch sizes close to the largest that fit available equipment and downstream packaging. Campaign decisions—how many batches of a product to run in sequence—must balance efficiency gains from longer runs against the need to limit inventories and preserve capacity for other products on shared lines, constraining how flexibly manufacturers can ramp production up or down over the year. Campaign planning also has important quality management implications, because cleaning and line preparation between campaigns introduce additional opportunities for quality issues.

DEA quota and inventory controls

To understand DEA's role in the reliability of Schedule II medicine supply chains, it is useful to first lay out how the current quota and inventory systems are designed. We begin by outlining the legal and institutional foundations of the quota system and explaining how aggregate and facility-level quotas are set, allocated, and adjusted. We then turn to regulations and practices around inventory management.

BRIEF HISTORY OF THE CONTROLLED SUBSTANCES ACT

Medications with the potential to be abused pose significant public health risks, including addiction, overdose, and other social harms. Because of these risks, governments across the globe have intervened for more than a century either on the demand side, to reduce inappropriate prescribing and use, or on the supply side, to prevent diversion of legally produced medicines into illicit or nonmedical channels.

The modern scheme of legal supply-side controls began in 1961 when the U.S. and other United Nations member countries signed the Single Convention on Narcotic Drugs. This comprehensive treaty consolidated and effectively replaced a patchwork of earlier opium- and coca-focused conventions, which had proven complex and poorly enforced. The Single Convention required United Nations member countries to establish a basic set of controls over the supply and trade of narcotic drugs that included the concept of quota, while still allowing considerable flexibility in

how countries design and administer their domestic control systems.

Following the Single Convention, the U.S. Congress enacted the Comprehensive Drug Abuse and Control Act of 1970, also known as the Controlled Substances Act (CSA), both to implement the treaty obligations and to replace its own patchwork of federal narcotic and psychotropic drug laws. The CSA vested regulatory and enforcement authority with the Attorney General and established five "schedules," with Schedules I and II subject to the strictest supply controls through quotas. In the original scheduling actions under the CSA, methylphenidate and certain amphetamine combination products used to treat ADHD were initially placed in less restrictive schedules and then, within months, were reclassified into Schedule II, bringing them under production quotas, inventory limits, and stringent recordkeeping.

To implement the CSA, President Nixon used an executive branch reorganization to create the Drug Enforcement Administration (DEA) within the Department of Justice. This 1973 change consolidated federal drug law enforcement responsibilities into a single agency, which has since served as the federal entity charged with enforcing the CSA and related federal controlled substances laws.

Because the CSA's quota provisions were sparse when enacted, barely over a page of text, the controlled substances framework has, in practice, evolved mainly through regulation rather than new legislation. Over time, successive DEA rules and practices have refined quota administration for Schedule II substances, including how DEA calculates aggregate production quotas, allocates individual quotas, and uses inventory data in those decisions. Later legislative changes, such as measures adopted in 2018, further expanded DEA's tools, contributing to a more complex quota framework for these drugs.

THE QUOTA SYSTEM

The quota framework is the main way DEA shapes the legal supply of Schedule II medicines, so assessing DEA's role in supply chain resilience requires a clear picture of how that framework works in law and in practice. This section first describes the main types of DEA quotas and how they fit together. It then explains how aggregate production quotas are set, and how individual manufacturers' quotas are allocated and adjusted over time.

Types of DEA quotas

DEA's quota system has three main components, each defined in terms of the weight of a controlled substance (in kilograms) rather than in numbers of dosage units:

- Aggregate production quota
- Manufacturing quota
- Procurement quota

The CSA establishes aggregate and manufacturing quotas in statute and requires that the sum of all manufacturing quotas for a substance not exceed its aggregate quota. Procurement quotas are created and governed solely by DEA regulation.

The aggregate production quota (or **aggregate quota** for brevity) operates at the national level: it is an annual national ceiling, in kilograms, on how much of each Schedule II substance may be manufactured for legitimate use in the United States and for export. DEA sets a separate aggregate quota for every covered substance, so, for example, hydromorphone and its precursor oripavine, each have their own kilogram ceilings. Because aggregate quotas can apply to precursors, quota decisions can affect the supply of some medicines that are not themselves scheduled—for example, intranasal naloxone is not a scheduled drug, but its synthesis relies on controlled precursors, making part of its supply chain governed by manufacturing quota.

Manufacturing quotas authorize, for each individual facility, how much of a controlled substance it may produce within the limit set by the aggregate ceiling, including both active ingredients and certain controlled precursors. Because the term "manufacturing quota" applies to whoever manufactures the controlled substance active pharmaceutical ingredient (API), not

to who uses it to make the final product, we will use a more descriptive term: **API manufacturing quotas.**

In turn, **procurement quotas** authorize facilities to acquire controlled substances as inputs for production or repackaging. Unlike API manufacturing quotas, which govern how much of a controlled substance a facility may produce, procurement quotas govern how much of a controlled substance a facility may receive for use in further manufacturing. This means that a facility that turns one controlled substance into another (e.g., a facility that makes hydromorphone from oripavine), therefore operates subject to API manufacturing quota for what it produces (hydromorphone) and a procurement quota for the controlled substance it uses as an input.

In 2023 and 2024, DEA took steps to further subdivide each facility's manufacturing and procurement quotas. First, the agency divided the facility quota into commercial manufacturing, product development, packaging, transfer, and replacement. Then they announced additional requirements that further subdivided several of these sub-quota categories into domestic use and exports. Now, a given facility will contend with several sub-quotas, depending on facility type. For example, an active-ingredient manufacturer may have commercial manufacturing procurement quota for domestic use, commercial manufacturing procurement quota for export, API manufacturing quota for commercial manufacturing for domestic use, and API manufacturing quota for commercial manufacturing for export, with corresponding product development for domestic or export and replacement quotas, if needed.

With several dozen controlled substances subject to quota and multiple active-ingredient and finished-dose manufacturers for each (we estimate there are over 350 actively marketed manufacturer-drug pairs for Schedule II medicines according to our assessment of Drugs@FDA), this structure translated into 4,000 individual facility quota decisions each year, even before sub-quotas were introduced.

Figure 2 illustrates how the various quota types map onto a simplified hydromorphone supply chain, from raw material importers through precursor, active-ingredient, and finished dosage form manufacturing facilities. For simplicity, we ignore the possibility of separate facilities labeling or repackaging the finished product, each with their own set of procurement quotas. We also ignore the need for product development or replacement sub-quotas. The facility-specific quota quantities shown are purely illustrative rather than what DEA actually assigns, although the aggregate quotas are similar to DEA's 2026 decision. Likewise, the manufacturing yield rates (e.g., how much hydromorphone is acquired from oripavine) are based on current industry methods, derived from UN reports and recent patents.

The diagram shows a Schedule II precursor facility converting concentrated poppy straw into oripavine under a precursor-level manufacturing quota (bound by oripavine aggregate production quota), with the resulting oripavine then flowing to two active-ingredient manufacturers of hydromorphone that each hold API manufacturing quotas whose sum is below the hydromorphone aggregate production quota. The remaining oripavine is procured by API manufacturers producing other substances that require oripavine (e.g., naloxone). The number of active-ingredient manufacturers is also simplified here; FDA data list as many as five possible active-ingredient manufacturers.

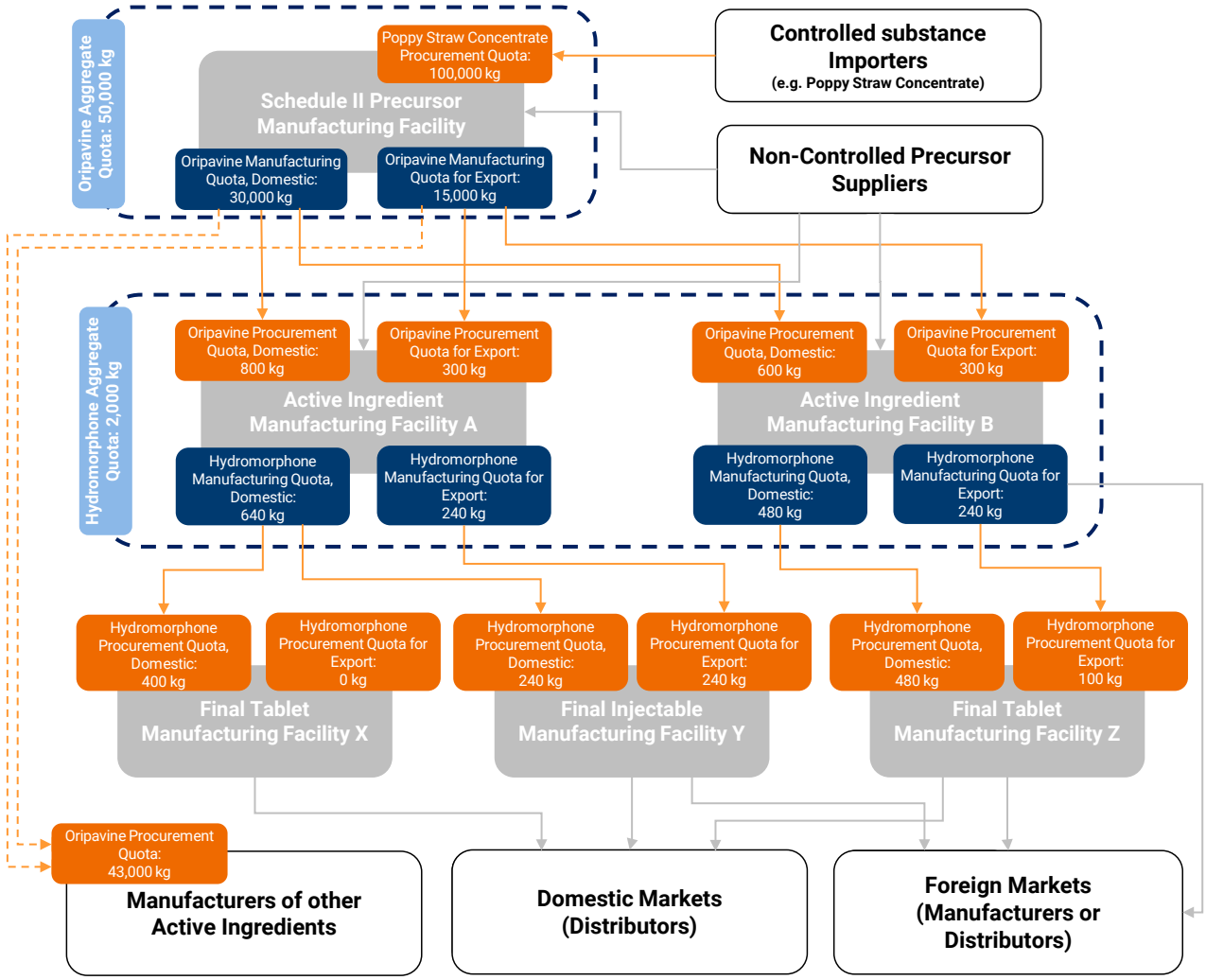
In our example, those API manufacturers, in turn, supply three finished dosage form manufacturing facilities, one making vials and two making tablets. Manufacturer B also directly exports some API to foreign dosage facilities. Each of the domestic facilities operates under its own procurement quotas for hydromorphone destined for domestic and export markets. Here again, the picture is simplified; FDA data list as many as 11 firms with active (not discontinued) authorizations to market hydromorphone in either oral or injectable form.

Setting of aggregate quota

To determine the national aggregate production quota (or aggregate quota for short), the CSA directs the Attorney General to calculate an amount sufficient to "provide for the estimated medical, scientific, research, and industrial needs of the United States, for lawful

FIGURE 2

Illustration of DEA quota system



export requirements, and for the establishment and maintenance of reserve stocks."

In implementing the aggregate quota mandate, DEA considers multiple factors, including historical consumption data, trends in legitimate medical use, projected export requirements, and input from federal health agencies. FDA, in particular, compiles annual estimates of medical and scientific need for selected controlled substances using a combination of sales data and other utilization inputs, and transmits those estimates to DEA as part of setting aggregate quota. The public record does not make clear how FDA's estimates are constructed or how DEA weights them relative to its own projections, but DEA Federal Register notices proposing and establishing aggregate production quotas indicate that the agency relies heavily on historical sales data to forecast domestic demand.

For five Schedule II opioids—fentanyl, oxycodone, hydrocodone, oxymorphone, and hydromorphone— the long-standing aggregate quota calculation requirement was changed through the SUPPORT Act of 2018. The CSA now directs DEA to estimate diversion based on "rates of overdose deaths and abuse and overall public health impact" and to reduce the aggregate quota accordingly.

Setting the aggregate quota follows a relatively regular procedural and timing pattern, with DEA publishing proposed quotas, reviewing public comments, and then issuing a final order. Although the CSA does not

specify a deadline, regulation and agency practice call for the process to begin by September 1 of the preceding year, and in recent years, proposed quotas have been published in the fall and finalized in December or very early in the quota year.

Although DEA establishes aggregate production quotas on an annual basis, the agency can adjust these ceilings midyear when new information emerges about medical need, diversion risk, or supply disruptions. Any increase requires notice-and-comment rulemaking, followed by recalculation and reallocation of individual API manufacturing quotas, a cumbersome process that can take several months.

Year-start quota allocation for facilities

DEA issues individual procurement and API manufacturing quotas only to domestic facilities because only DEA registrants are eligible, and registration is limited to manufacturing sites on U.S. territory (though they may be foreign-owned). Schedule II controlled substances can technically be imported, but imports are rare. In contrast, exports are more common: available United Nations and DEA data suggest the United States typically exports about 10% of annual production for many Schedule II substances, with much higher shares for certain products such as methylphenidate and fentanyl.

For each substance, facilities must apply for procurement quota by April 1 before the year the quota will govern, and for API manufacturing quota by May 1. In both cases, they must provide recent and projected sales, inventory levels, production cycles, and intended uses of the substance. DEA must, by statute and regulation, consider this information, along with any other factors the DEA Administrator deems relevant to ensure an adequate and uninterrupted supply for legitimate medical and scientific needs.

DEA's approach to issuing individual quotas affects not only how much Schedule II product can be made in a year, but also when those quotas are available for manufacturers to use. Even after the annual aggregate quota for a substance has been established, DEA does not typically allocate the full volume to manufacturers at the beginning of the quota year: stakeholder comments describe firms receiving only a portion of the quota they historically use and then having to seek multiple increases as the year progresses. In practice, this means manufacturers often operate under a series of mid-year quota increases, with their timing and size determining how much they can actually produce at different points in the year.

Mid-year quota allocation for facilities

Once DEA has assigned individual API manufacturing quotas to manufacturers, its ability to increase, decrease, or reallocate those quotas is constrained by statute and by DEA's own regulations. This limited flexibility may help explain why, in practice, firms often receive only a portion of the quota they request at the start of the quota year and must return for mid-year adjustments to obtain additional volume.

The CSA permits DEA to increase API manufacturer quotas if a registrant applies and demonstrates that its original quota is insufficient to meet legitimate demand, including maintaining appropriate inventory levels. However, if DEA has already effectively exhausted the aggregate quota in its original allocation of API manufacturing quotas, it must first increase the aggregate quota through notice-and-comment rulemaking before granting additional API manufacturing quota, a process that can take several months from the manufacturer's request, hindering timely shortage responses.

Procurement quotas are a regulatory DEA creation not mentioned in the CSA, so on paper, they appear more flexible than API manufacturing quotas. In practice, however, they are constrained by how much active ingredient is available under the aggregate and API manufacturing quotas, so mid-year changes mainly redirect a fixed supply across manufacturers and supply chain stages rather than increasing the total amount of product that can be made. As a result, firms may obtain additional procurement quota midyear yet still be unable to increase output if there is no corresponding increase in API manufacturing quota.

DEA's tools for reducing or reallocating preassigned quotas are limited, which tends to make quotas

"sticky" once granted. The CSA explicitly contemplates only two situations: when a registrant loses its DEA registration, in which case its quota is suspended or revoked, and when DEA lowers the aggregate production quota, in which case it must apply proportional reductions across all manufacturers. The statute is, however, silent on the more common situation where a compliant facility cannot use its quota because of manufacturing problems. In those cases, DEA has generally taken the view that it cannot unilaterally reassign quota, instead relying on manufacturers to voluntarily relinquish unused quota so that DEA can reapportion it to other registrants.

In part to address the lack of flexibility with reallocating quota, DEA briefly shifted in 2024 to a quarterly procurement quota allocation model, under which manufacturers received roughly 25% of their annual quota each quarter, with subsequent releases contingent on DEA's assessment of demand, inventories, and diversion risk. But DEA abandoned the quarterly approach within a few months after many quota assignments were delayed and operational problems threatened morphine supply. DEA then moved to a current "hybrid" quota system that differs by product and quota type.

Under the current quota system, aggregate quota remains annual and manufacturers still submit a single yearly application, but DEA is now supposed to release procurement quota for most noninjectable Schedule II products in two installments—one at the start of the year and another around midyear—while API manufacturer quota and procurement quota for sterile injectables are supposed to be granted on an annual basis. Recent public comments to DEA do not praise the hybrid as the needed fix; instead, they continue to describe missed deadlines, inadequate initial allocations, and the need for repeated quota applications.

INVENTORY CONTROLS

For regulatory purposes, a manufacturing facility's inventory includes basic drug weight of the controlled substance on hand at a site, meaning the weight of the controlled substance itself in kilograms. This inventory includes the controlled substance in raw drug substance form, in materials that are partway through the manufacturing process, in finished dosage forms undergoing quality control testing, and in warehouse stock ready for distribution.

Unlike manufacturing quotas and procurement quotas, which reset on January 1st each year, inventory can carry over. But DEA limits a facility's current-year inventory to a fixed percentage of the amount of controlled substance in finished-product shipments, product losses, and transfers (collectively, "product outflows") that left that facility over the previous year (with DEA's formal calculation averaging last year's outflows with the current year's expected outflows). This effectively ties how much controlled substance can be in process, awaiting processing, or awaiting release, to how much was moved through the facility last year. DEA refers to these limits as maximum allowable inventories.

Under rules established in 2023, recipients of API manufacturing quota are expected to keep inventories no higher than 40% of product outflows averaged over the current and prior year, while recipients of procurement quota are subject to a 35% cap, with higher allowances preserved for active ingredients intended for sterile injectable formulations, which often require larger, less frequent production runs. These caps represent a tightening from the long-standing practice of allowing maximum allowable inventories of 50%.

By regulation, exceeding the inventory caps by more than 15 percentage points triggers an automatic suspension of additional quota until inventories fall below specified restart thresholds. DEA may waive these restrictions in certain circumstances, such as a declared shortage. Manufacturers are expected to track their inventories against the caps and avoid placing or relying on orders that would require additional quota once they are over the limit. DEA primarily monitors inventory levels through required year-end inventory and transaction reports.

DEA's role in drug supply chain reliability

DEA's quota and inventory framework affects each of the four shortage elements for Schedule II drugs: it shapes the probability and size of shocks, and it constrains both the system's ability to absorb them and to recover. We discuss each element, in turn.

PROBABILITY OF A SHOCK

DEA's quota framework shapes the probability that Schedule II markets will run into trouble when either demand rises or supply is disrupted.

The system is particularly vulnerable to **increases in demand**, even if those increases are measured and steady. When quotas are not limiting production, an increase in demand can be met by making and selling more product, so the higher demand shows up directly in the sales data DEA uses for future quotas. But when quotas are binding, and manufacturers are already producing up to the ceiling, they cannot supply the extra units even if providers and patients want them; in that case, the sales data reflect only what the quota allowed to be sold, not the full level of clinical need. Because DEA relies on historical sales data to forecast demand, this feedback loop can understate true need precisely when the system is already constrained, increasing the risk that even steady demand growth will trigger or prolong shortages.

Delays in issuing quotas and manufacturing disruptions can further distort these demand estimates. When firms cannot produce up to the levels they were given because quotas arrive late or plants are temporarily offline, the effect is similar to further lowering a fully binding quota, as if the effective ceiling were set lower than on paper. When these artificially depressed sales figures are then fed back into DEA's forecasting process, subsequent aggregate quotas may be set even lower than previously set binding quotas, creating an even greater wedge between actual medical need and DEA's demand estimates.

Legislation in response to the opioid epidemic has introduced additional tension into DEA's demand estimation. Under the CSA, quotas are to be set at a level necessary to meet estimated legitimate medical and other lawful needs, along with exports and reserves. But for five opioids (fentanyl, oxycodone, hydrocodone, oxymorphone, and hydromorphone), the 2018 SUPPORT Act directs DEA to further reduce these quota estimates to account for estimated inappropriate use.

Spillover demand adds another pathway by which shocks in one product's supply can create shortages in others. When one drug becomes scarce, prescribing doctors shift to clinically similar substitutes. But aggregate quotas for those substitutes are set under assumptions of stable demand, with no explicit upward adjustment for potential spillovers, which can then lead to shortages of those substitutes. This dynamic was evident when amphetamine mixed salt products (Adderall) first went into shortage. Prescribing then shifted toward methylphenidate (Concerta) and lisdexamfetamine (Vyvanse), contributing to both products being added to the drug shortage list, where they remain as of the publication date for this report.

In principle, domestic active ingredient and finished-dose manufacturing could reduce exposure to **geopolitical risk**—that is, the risk that supply is disrupted by trade restrictions, conflicts, or other country-level shocks. In practice, the effect is more ambiguous: key inputs remain globally sourced, so domestic production of downstream manufacturing steps does not fully insulate the system from international disruptions. At best, a predominantly domestic footprint of later production stages could buy time in the face of external shocks, but that buffer would depend on having larger inventories, a topic we return to in a later section.

For other supply disruptions such as manufacturing delays, quality failures, and local natural disasters, the direction of DEA's overall effect on risk is uncertain. Quota allocation decisions can concentrate production in a small number of facilities, so placing large shares of quotas at sites with higher underlying risk makes the system more exposed to quality failures

or local disasters. Yet there is no evidence that quota decisions systematically account for plant-level risk factors such as FDA inspection history, geographic disaster exposure, or reliance on single-source inputs. As a result, it is unclear whether, on balance, current quota allocations reduce or increase this dimension of supply chain risk.

SIZE OF SHOCK

Beyond how often shocks occur, shortages also depend on the size of each disruption relative to total market supply.

In principle, DEA's framework could reduce the size of supply shocks by explicitly accounting for facility-specific risks when allocating quota—for example, shifting quota away from plants with poor quality histories or high exposure to natural disasters, so that a disruption at any single site would take out a smaller share of total capacity. The CSA explicitly calls for consideration of such risks. But as mentioned above, there is little public evidence that plant-level risk indicators such as FDA inspection findings or geographic disaster exposure influence quota allocations.

Brand-to-generic transitions show how the quota system can precipitate the shock even when the shock is predictable. For high-volume, small-molecule drugs, brand volumes typically fall by about 75% to 85% within six months of generic entry, so a large shift in demand toward generics is both routine and foreseeable. In the case of lisdexamfetamine (Vyvanse), which lost market exclusivity in August 2023 and for which multiple generic manufacturers sought procurement quotas, brand sales declined only modestly after the loss of exclusivity. This suggests that DEA must have given quotas to the brand and not the generics.

Separately, Schedule II medicines are less likely to experience a dynamic that often makes shortages worse for other hospital-administered medicines: stockpiling at the first sign of trouble. When hospitals rapidly expand orders to build local inventories, this behavior can precipitate and intensify a shortage. For Schedule II medicines, large wholesalers, responding to opioid-related enforcement actions, have implemented monitoring systems that flag and sometimes block Schedule II drug orders that deviate from a customer's historical purchasing patterns. These "red flag" mechanisms make such stockpiling behavior difficult. But while this dampens one important amplifier of shortages, it also means the system can struggle to accommodate legitimate increases in clinical demand, as discussed above.

ABILITY TO ABSORB A SHOCK

For Schedule II controlled substances, even when operational capacity buffers exist, the quota framework can prevent them from being used. DEA's quota system, by definition, limits how much controlled substance can be at a facility at any given time, so manufacturers with spare line time or the ability to ramp up production are unable to increase output without an appropriate quota. As a result, unused capacity at some facilities can coexist with unused quota at others, and the system has little practical ability to ramp up production in response to changes in supply chain pressures.

Technically, the inventory buffers that DEA allows for Schedule II controlled substances should help compensate for limits on operational capacity. However, many readers may intuitively equate "inventory" with finished product on the shelf, making a 40% to 50% allowance seem ample or generous. In reality, DEA's inventory definition includes unprocessed controlled substance, substance in processing, and processed substance awaiting release, much of which cannot be made available to the market on short notice.

As a result, shortages of Schedule II products occur and persist, suggesting that the allowed inventory caps, while nontrivial, do not always provide enough cushion. Inventories can help other manufacturers move their own production forward, but when an af-

fected facility cannot restart and continue processing its own in-process inventory, or work through enough of its quota before it expires, some stock and quota will be unused, and the system still ends up with a production shortfall.

Inventory limits can be particularly tight for manufacturers with a growing market share. Because allowable inventories are set as a fixed percentage of each firm's recent controlled substance outflows (sales, transfers, and losses), a company whose market share is rising—whether due to spillovers from other shortages or because it has just been given more quota—will see its permitted inventory levels lag behind its expanding volume. The Vyvanse (lisdexamfetamine) generic launches illustrate this problem: new entrants had little or no prior-year sales, which kept their allowable inventories low even as they were expected to help address spillover demand from amphetamine salts.

ABILITY TO RECOVER FROM A SHOCK

Recovery from disruptions in Schedule II drug markets follows the same basic logic as in other prescription drug markets. If the affected manufacturer can resolve the issue on its existing production lines—for example, by correcting an equipment failure or addressing a quality deviation—recovery mainly depends on how quickly it can restore output and work through backlogs. By contrast, when the underlying problem is a lasting loss of capacity, recovery typically requires other manufacturers to bring new or dormant capacity online, much as they would after a sustained demand increase.

When the affected manufacturer attempts to recover from a disruption, it will typically still have some combination of remaining quota and inventory. But if production is halted for part of the year, some quota may expire before production can resume. If inventory or in-process batches are damaged or must be discarded—for example, because of contamination or flooding—the firm loses both product and the quota used to make it. In both cases, fully restoring supply may require mid-year quota increases.

When recovery depends on shifting production to other manufacturers, the quota system can impede that shift in two ways. The affected manufacturer has strong incentives to hold on to its quota—even if it cannot fully use it—because future allocations are informed by historical production and sales. At the same time, DEA may be reluctant to grant substantial additional quota to other manufacturers while unused quota technically remains with the original firm, even when that quota is unlikely to be converted to product before it expires. This combination makes it harder for new or expanding manufacturers to obtain the quota they would need to step in quickly after a major disruption, prolonging the period of constrained supply.

Mid-year increases in aggregate quota increases, which could help in these situations, are legally possible, but in practice, relatively rare and often slow. If the existing aggregate quota has already been fully allocated, DEA must first increase the aggregate quota through notice-and-comment rulemaking and then process individual manufacturing and procurement quota applications, a sequence that can add months before firms can begin to translate additional quota into finished product.

Our analysis of DEA publications revealed that, of the 383 aggregate quotas issued from 2021 through 2025, only 20 received mid-year adjustments. Six of them were decreases and include four for FDA essential medicines which have been persistently in shortage since 2017.

Discussion

The U.S. framework for Schedule II controlled substances embodies a deep structural tension. These medicines are approved for medically accepted uses and are central to managing pain, ADHD, and anesthesia, yet the system that governs their availability is designed and led as an enforcement apparatus rather than a public health initiative. The core levers that determine whether patients can receive these drugs—quotas, registration, and inventory rules—are regulatory tools, but they sit in a justice-sector agency whose primary mandate is to prevent channeling of legally manufactured controlled substances into illegal or nonmedical use, not to sustain reliable access for medically appropriate use.

Because Congress placed oversight of Schedule II medicines inside the Department of Justice, DEA approaches these regulatory levers through a law enforcement lens. The agency's practices focus on detecting and punishing oversupply, rather than as instruments for ensuring that medically appropriate demand can be met.

This was not always how the quota system was framed. The CSA directs DEA to set aggregate quotas at levels sufficient to meet legitimate medical and other lawful needs, along with exports and reserve stocks, putting access on equal footing with diversion control.

In the years leading up to and during the opioid epidemic, DEA routinely approved large increases in aggregate production quotas for Schedule II opioids, contributing to a sharp expansion in supply. In response, Congress publicly and repeatedly faulted DEA for having approved large opioid quotas, and pressed the agency to restrict quota far more aggressively. The SUPPORT Act of 2018 formalized that by requiring DEA to estimate "inappropriate use" and, for some Schedule II medicines, subtract those amounts from the aggregate production quota, reinforcing the expectation that cutting quotas is a primary response to misuse rather than a tool of last resort.

Once quota cuts became a central tool for preventing the diversion of drugs for illicit purposes, this change created risks extending well beyond the opioid tablets at the center of the addiction crisis. Because aggregate quotas are set largely from historical sales, adjusted only slowly midyear, and complicated by unused quotas that are difficult to reallocate, they cannot respond quickly when demand grows or shifts across products.

The prolonged shortages of ADHD treatments illustrate how this framework plays out in practice. Following a steady increase in prescribing of amphetamine mixed salts (Adderall) and manufacturing problems at one important facility, multiple ADHD medications went into shortage as demand spilled over from one ADHD drug to another. The DEA and FDA responded by emphasizing that manufacturers had used only about 70% of their allotted amphetamine quota in 2022. As discussed earlier, however, unused quota often reflects structural features of the quota system—rather than a genuine lack of willingness or capacity to produce.

These dynamics have been reinforced by implementation choices during the Biden administration. DEA reduced allowable inventories, shrinking the buffers that might otherwise help manufacturers maintain supply when shocks occur or quota decisions are delayed. It introduced new sub-quota categories that divide each Schedule II substance into separate ceilings for different uses, creating additional ways that usable quota can be stranded. DEA also briefly shifted procurement quotas to a quarterly schedule, a change manufacturers had warned would be unworkable and risky for injectable opioids, and then reversed course within months amid operational problems and emerging concerns about morphine supply, underscoring how little weight it placed on preserving appropriate medical use when redesigning quota mechanics.

This fragmentation, coupled with delays in issuing quotas, leaves markets with little resilience when shocks occur. When manufacturers do not receive quotas in time or cannot shift them across products and facilities, they cannot use available line time or inventories to stabilize supply, so even modest disruptions can trigger or prolong shortages. As the

prolonged shortages of sterile injectable opioids on the FDA Essential Medicines List suggest, these constraints make it harder for already constrained markets to recover fully from shocks, so shortages tend to persist and oscillate between more severe and less severe rather than ever fully resolving.

Against this backdrop, the U.S. approach to medically used controlled substances stands out internationally. Other countries, including Canada and the United Kingdom, place primary responsibility for medically used controlled substances in health ministries or health-sector regulators. And they emphasize tools that govern prescribing, dispensing, and treatment—such as prescription monitoring programs, clinical governance requirements, and audit trails for controlled drug use in care settings—rather than tight, top-down limits on aggregate national production or binding national quotas. In turn, these prescribing- and dispensing-focused tools give regulators a more precise view of clinical need and of where use is appropriate or inappropriate.

More than a half a century ago, the United States briefly moved in a similar direction. In the early 1960s, as countries were adapting their systems to the United Nations Single Convention that compelled countries to restrict controlled substances production, President Kennedy's Advisory Commission on Narcotic and Drug Abuse recommended placing U.S. authority over medically useful controlled substances under what was then the Department of Health, Education, and Welfare (now the Department of Health and Human Services).

Instead, when Congress enacted the CSA, it vested primary authority in the Attorney General and built a justice sector–led framework that manages medically used controlled substances through registration, monitoring, and aggregate production quotas. Since the opioid epidemic, debates over how to manage Schedule II drugs have increasingly centered on using quotas and inventory limits as core control tools. But quotas are a blunt tool: for drugs with substantial inappropriate use, they work by forcing appropriate patients and nonmedical users to compete for a fixed supply, and for hospital IV opioids, where inappropriate use is far less common, they leave so little slack in the system that even routine disruptions can push supply chains into repeated, prolonged shortages.

It is worth underscoring that tighter national production quotas were not what ultimately drove progress against the opioid crisis. The prescription medicines most closely associated with the first wave of that crisis—high-dose oxycodone and hydrocodone—were brought under control primarily through demand-side measures, including state prescription drug monitoring program mandates, the 2016 CDC opioid prescribing guideline, insurer and health-system prescribing controls, and aggressive enforcement actions targeting "pill mills," fraudulent prescribing, and other opioid diversion schemes.

Conclusion

The current system for Schedule II controlled substances poses a central policy question: Can a top-down, quota-based framework by a justice-sector agency support access to medically important controlled substances and simultaneously prevent diversion?

In practice, the experience documented in this paper suggests that the answer is no. The DEA's laser focus on limiting diversion of Schedule II medications into illicit channels sits uneasily with the role of ensuring an adequate supply of drugs with approved medical uses. The same institution that must respond to illicit fentanyl trafficking and pill-mill prescribing effectively runs key supply chains for all Schedule II drugs, acting as a de facto capacity planner that not only assigns market shares but also constrains how much of that capacity manufacturers can bring to patients.

These findings make clear that improving supply chain resilience for Schedule II drugs will require rebalancing both DEA's institutional objectives and its tools. At a minimum, that means that the quota program should be used as a public health instrument that supports adequate, reliable access for legitimate medical use, and that diversion control must be pursued subject to a binding constraint of maintaining that access. Coordination with federal public health agencies should be strengthened and formalized, with DEA expected not only to draw on FDA's shortage intelligence and forecasting work but also to share timely quota and diversion data so FDA can help identify and prevent emerging gaps in medical access.

Beyond clarifying mandates and improving coordination, DEA's quota tools and their implementation also need reform, much of which can be accomplished administratively under current law. The Administration can begin by setting aggregate production quotas high enough to include an explicit reserve that can be released quickly when shortages emerge, consistent with the statute's direction to provide for "the establishment and maintenance of reserve stocks." It can also allow unused quotas to roll over into the next year, so that temporary disruptions do not permanently ratchet supply downward.

Beyond adjusting headline quotas, the framework for facility-level quotas should be greatly simplified. DEA should eliminate procurement quotas that function mainly as monitoring tools, collapse unnecessary subcategories within the manufacturing quota structure, and streamline the quota request and adjustment process. Finally, DEA should revise allowable inventory levels—especially for drugs on Essential Medicines lists—so manufacturers can hold inventories robust enough to cushion disruptions rather than operating with levels that are too thin to absorb shocks.

In turn, rather than micromanaging manufacturing-facility allocations through complex quota categories, DEA should adopt a more targeted, risk-based approach to preventing diversion. This direction aligns with the Department of Justice Office of Inspector General's 2019 review, which urged DEA to improve data integration and analytics-driven enforcement. In practice, that means using integrated data to detect suspicious or anomalous prescribing, dispensing, and distribution patterns. When such patterns emerge, DEA should apply targeted tools—such as investigations, registration enforcement, corrective action plans, civil and criminal penalties, and enhanced monitoring—while leaving routine production and distribution subject to fewer front-end constraints.

Without such rebalancing, the United States will remain in the position of managing standard-of-care medicines through a system built primarily to reduce oversupply, and with too little regard for access and outcomes for appropriate patients. A more sensible approach would replace blunt, backward-looking production caps with oversight that targets problematic behaviors, using timely, integrated data on prescribing, dispensing, and purchasing to show where the product is in the system. With that visibility, DEA should then intervene surgically through investigations, registration actions, and other enforcement tools rather than constraining production across the board.

Such a shift would better align diversion control with what the Controlled Substances Act explicitly requires—the obligation to sustain an adequate, reliable supply of Schedule II medicines for legitimate medical use.

BROOKINGS

1775 Massachusetts Ave NW
Washington, DC 20036
(202) 797-6000
www.brookings.edu