Regulated Prices, Global Capacity, and Drug Shortages

Pierre Dubois[*] Gosia Majewska[†] Valentina Reig[‡]

July 2026[§]

Abstract

Drug shortages affect most pharmaceutical markets. We develop a model of capacity choice for a firm supplying multiple countries under regulated prices, in which supply shocks generate shortages: the highest-price countries remain fully served, middle-price countries are rationed, and the lowest-price countries are excluded, so that domestic and foreign prices affect both the frequency and the magnitude of shortages. We then develop a method to detect shortages and measure their magnitude using only national sales data. Applying it to monthly pharmacy sales in France over 2019-2024, we find an average shortage probability of 26%. Higher French prices reduce both the probability and the magnitude of shortages in France. Higher UK prices reduce the probability of French shortages but increase their magnitude, while higher Polish prices reduce both, consistent with the capacity and allocation channels of the model. Counterfactuals on antipsychotics show that targeted price increases can substantially reduce shortages at limited expenditure cost.

Keywords: Drug Shortages, Pharmaceuticals, Price Regulation, Capacity Investment, International Spillovers.

JEL codes: I11, I18, L51, L65.

[*]Toulouse School of Economics, pierre.dubois@tse-fr.eu

[†]ESSEC Business School and THEMA, majewska@essec.edu

[‡]Toulouse School of Economics, valentina.reig@tse-fr.eu

[§]We thank GERS for accessing the Sell Out Data of pharmaceutical sales and the Toulouse School of Economics Health Center sponsors for their support. We thank participants of the TSE Workshop on Innovation, regulation and organization in healthcare in Paris, the 30th European Workshop On Econometrics and Health Economics in Madrid, ESSEC internal seminar, the Mapping the Pathways Workshop in Verona, the EARIE conference in Amsterdam, the Paris IO Day and the Finnish Health Economics Seminar Series. Financial support from the European Research Council and the European Union under ERC grant agreement No. 101141890-EHCI, as well as the ANR under grant No. ANR17-EURE-0010 (Investissements d'Avenir program) and ANR-16-CE36-0004.

1 Introduction

Drug shortages are a problem widely documented around the world. The main implications of shortages include treatment disruption (specifically, switching to inferior alternative therapies (Becker et al., 2013)), additional costs (e.g., increased hospital pharmacy workloads (Kaakeh et al., 2011) and price increases for affected drugs (Dave et al., 2018)), and the emergence of illicit drug markets (Fittler et al., 2018). While eliminating supply chain uncertainty may be impossible, maintaining oversized production capacities to cover demand during factory disruptions is often prohibitively expensive. However, it is important to measure the right trade-off between spending on drug manufacturing and shortage risk to reduce shortage events.

In this paper, we analyze how prices affect the probability and magnitude of drug shortages. Studies in the US context suggest that downward pressure on prices is correlated with more frequent shortages (e.g., Yurukoglu et al., 2017; Ridley et al., 2016; Conti and Wosińska, 2025), a literature we discuss in more detail below. In Europe, this mechanism warrants close investigation, as prices are often set by regulators (e.g., in France, the UK, and other countries).

Drugs are globally traded goods, and the prices in all countries can affect shortages. There are several mechanisms that can lead to different effects of prices on shortages. First, through capacity decisions: the supply side adapts its capacity of production to the anticipated drug prices. All else equal, higher prices lead to higher profits, incentivizing firms to build adequate production capacities. This will reduce the likelihood of shortages, which might happen with supply disruptions, for example, if factories need maintenance. Thus, higher prices are likely to reduce the likelihood of shortages. At the same time, higher prices also reduce demand, limiting the need to increase capacity.

Second, through the distribution of the limited supply across countries. Once the global capacity is fixed, the total production is distributed across countries. If the limited quantity for one country is lower than the demand at the regulated price, shortages occur because prices cannot adjust upward. Ceteris paribus, the magnitude of the shortage decreases as the local price increases. Additionally, foreign prices can exert a negative externality on the local quantity supplied. When geographical markets compete for limited supply, higher prices attract more supply, and reduce the supply of the drug in other countries. Thus, external prices can have either a positive or negative effect on local shortages. This mechanism of scarce capacity allocated across sources according to margins rather than costs is the force that Asker et al. (2019) show generates misallocation in global oil extraction. In our setting, the margins steering the allocation are administered by national price regulators rather than shaped by market power.

We formalize these intuitions in a simple model of firm capacity choice. In the model, the firm chooses its global capacity and the quantity supplied to each country, and shortages are driven by supply shocks. Introducing a monetary penalty for shortages, we demonstrate that countries with the highest price levels remain unaffected by shortages, those in the middle of the price distribution experience partial shortages, and those with the lowest prices receive no supply. Our model allows also for differential effects of prices on shortages. Higher prices in high price countries have a negative effect on capacity building, while higher prices in rationed countries generally increase capacity, provided that demand is not too elastic. Once capacity is decided, and a shortage occurs, the higher the price in the country, the more supply the firm will direct to it, but the effect of prices in other countries is ambiguous.

To test our model's predictions, we develop a methodology to detect shortage events and quantify their intensity using national-level sales data. Shortages occur when the quantities supplied do not meet demand. This gap can be significant for prescription drugs in Europe because their prices are regulated in most countries. Detecting shortage events in sales data is not trivial because sales do not reflect demand during shortage periods. Consequently, estimating demand using data containing shortage periods yields biased estimates.

Our method relies on the assumption that shortages are driven by supply shocks, and that the size of the market studied is small in comparison to the global market, making shortage events exogenous to local demand. Indeed, these shocks correspond to production capacity disruption events that alter significantly the production level of a drug during enough time so that stocks are insufficient to smooth the shock. This implies that a domestic demand surge cannot trigger a global shortage, as local demand remains small relative to world demand. Under these circumstances, when we observe large sales in a small country, it means there is no shortage in that country. Thus, we can define a threshold level of sales above which shortages do not happen. In the limit case where demand is constant, this method simply uses the fact that if sales are smaller than the regular demand, it means that there is a shortage and thus the supply is lower than the country's demand. Then, the question is how to detect the frequency of shortages that can be identified using the sales variations and demand estimates of the drug in the country. Of course, the demand for a drug in France is not constant but varies with prices, drug diffusion, innovation, thus our method needs to account for varying demand.

We consider a threshold depending on the quantile of the observed sales distribution, fixed at the drug-class level. We choose the threshold level by comparing the resulting predicted demand (estimated using time periods when sales are above the threshold) to the realized sales. We define shortages as periods when the observed sales are below the threshold and the predicted demand. The magnitude of the shortage is the sum of the differences between the predicted demand and the realized sales in these periods.

We apply the methodology to monthly data covering sales of retail pharmacies in France in the period 2019-2024 provided by the GERS association. We find that the average probability of a drug shortage, across product-month observations, is 26%, though with substantial heterogeneity.

Using these estimates, we analyze the impact of domestic and foreign prices (specifically, in the UK and Poland) on shortages. The results align with the predictions of our theoretical model. We find that higher domestic prices in France lead to a lower probability of a shortage and a smaller shortage magnitude, that is, a smaller share of demand left unserved, in case of a shortage. On the other hand, while higher prices abroad are also associated with a lower probability of a shortage in France, conditional on a shortage, they lead to different effects on the shortage magnitude in France. Prices in the UK, a country with generally more expensive drugs than France, are associated with larger shortage magnitudes in France. In contrast, higher prices in Poland, where price levels are typically lower than in France, are associated with lower shortage magnitudes in France. These results underscore the trade-off between maintaining low prices and mitigating the probability and magnitude of shortages. The effects of the UK prices suggest that, in the event of a shortage, firms shift the remaining supply toward countries with higher prices. But they also suggest positive spillover effects of prices on the shortage probability, probably through the production capacity that higher global prices allow firms to build.

To investigate further the relationship between prices and shortages, we estimate the effects of counterfactual price increases in France, using the example of the market for antipsychotics. Since observed sales understate demand during shortage episodes, we take as the quantity demanded the demand predicted by our detection method and estimate a discrete choice model on the full sample. The model captures price elasticities and intra-class substitution patterns, which are critical when evaluating price counterfactuals. We estimate a supply model under shortage periods to understand how higher prices can reduce shortages by increasing supply and decreasing demand. We find individuals are price elastic in this market, as they can substitute across multiple products within the class. Our counterfactuals include changing prices for all drugs, and increasing prices for one drug per class. Results show that minor price increases can substantially reduce shortages, but their welfare implications are not obvious if we account for the decrease in demand they cause in non-shortage periods.

Our paper relates to several strands of the literature on drug shortages. In the existing literature, the problem is most often analyzed by medical professionals, discussing clinical implications of the shortages, case studies of particular drugs, and surveys of the personnel directly affected by them. Some policy-oriented studies feature informal discussions of the economic forces in place (Acosta et al., 2019; Woodcock and Wosinska, 2013; Parsons et al., 2016; Dave et al., 2018). Relative to these descriptive contributions, we provide a formal model of capacity choice and supply allocation across countries, together with a systematic method to measure shortage events and their magnitude.

A first strand of economic studies documents the relationship between prices or reimbursement levels and shortages, mostly in the US institutional context. Conti and Wosińska (2025) synthesize this literature, arguing that generic drug shortages persist because prices cannot easily rise in response to scarcity, owing to contracts and procurement institutions, and because buyers undervalue supply reliability. The price rigidity at the core of their account is precisely what our model formalizes, in a setting where prices are administered directly by regulators and where the firm allocates scarce supply across countries with different administered prices. Yurukoglu et al. (2017) use the sterile injectable products market to study the relationship between reimbursement and shortages. In 2003 and 2005 Medicare substantially reduced the amounts reimbursed to practitioners administering the drugs. The paper hypothesizes that drugs with lower fixed costs and serving more Medicare patients are affected more severely by the decline in reimbursement. The empirical strategy, a quasi-difference-in-differences design using the drug's Medicare market share as a measure of treatment intensity, suggests that drugs more exposed to this policy change experienced longer durations of shortages. A theoretical model in Kim and Scott-Morton (2015) lends support to this conclusion, while Yang (2020) finds spillovers of the policy on the Californian Medicaid system.

Within this strand, low prices are often linked to generic competition or to buyer bargaining power. Parsons et al. (2016) document a positive correlation between the probability of shortage and the generic status of the drug. Low margins in the generic market mean that the manufacturing sites are used at their capacity. The lack of spare capacity limits the ability to perform regular maintenance, increasing the risk of disruptions. At the same time, without backup production lines, any disruption can translate into a shortage. The weak failure-to-supply clauses existing in this market aggravate this problem, as the manufacturer faces a low cost of undersupplying. Finally, limited capacity makes the producers shift to the most profitable products - Haninger et al. (2011) show that drugs experiencing shortages in 2008-2011, faced, on average, decreasing sales volumes and prices in 2006-2008. A study of generics dispensed in outpatient pharmacies confirms the relationship between price and risk of shortage, finding that lower-priced generics are more often in shortage (Dave et al., 2018). Apart from generic competition, prices can be driven down if the buyer has substantial bargaining power. Half of the childhood vaccines in the US are paid for by the government. The CDC acquires them directly from the manufacturers and uses strict price caps. Ridley et al. (2016) estimate that a 1% increase in price is associated with a 0.086 pp decrease in the probability of shortage in the following year. In the period studied, the shortages reach a peak in 2007, with 7 vaccines unavailable. Relatedly, Pauly (2005) suggests that vaccine shortages stem mostly from firms' underspending on the quantity and quality of labor needed to ensure the lack of errors and disruptions. Closest to our setting, Liebman et al. (2023) study the rationing of a pediatric vaccine during a shortage, showing how scarce supply is allocated when prices cannot adjust. We complement their single-product analysis by modeling the allocation across countries and estimating price effects on shortages at scale.

Evidence from Europe, where prices are set by regulators, is scarcer. Unlike in the US, branded drugs are relatively more often affected by shortages in Europe, with reference pricing policies quoted as the cause (Pauwels et al., 2014). In Switzerland, Blankart and Felder (2022) show that prices of medicines in which there is never a shortage are higher than for medicines that presented a shortage. They also estimate the economic cost of shortages, calculated as the counterfactual expenditure had the prescribed medicine been substituted with alternative options, reaching 0.3% of expenditures in 2018. We contribute to this strand by providing evidence from a large European market with regulated prices and, guided by our model of the international allocation of supply, by estimating not only the effect of domestic prices but also the cross-country spillovers of foreign prices on domestic shortages.

A second strand focuses on the supply side, namely manufacturing quality and the organization of production. Stomberg (2016) focuses on non-injectables drugs and how shortages react to the changes in the intensity of FDA's quality control. When the probability of inspection increases, manufacturers not meeting the standards might experience disruptions. Galdin (2025) studies the international market for injectable drugs and the incentives of manufacturers to offshore and price generic drugs. She finds that drugs offshored to Asia are 56% more likely to experience shortages compared to their domestically produced counterparts, with shortage spells lasting on average 130 days longer. This provides causal evidence for her structural model, in which suppliers trade off supply reliability against lower production costs. She evaluates counterfactual policies and shows that failure-to-supply clauses in contracts can reduce shortage incidence, yielding net consumer welfare gains that are not offset by higher drug prices. Our model shares with this work the emphasis on supply disruptions as the proximate cause of shortages, but focuses on how regulated prices shape both the firm's capacity investment and the allocation of scarce supply across countries.

A third strand assesses how regulatory dispositions across countries affect shortages. Lee et al. (2021) analyze mandates on reporting shortages from regulatory agencies in the US and Canada. Using a difference-in-differences design, in which Canada is the treatment group, they find that compulsory reporting of shortages reduces the duration of a drug's shortage (its time-to-recovery) by 41% relative to voluntary reporting. They explain this is due to positive pressure to end the shortage in a short time because competitors are now aware of this shortage. Kortelainen et al. (2023) study the effects of different regulatory regimes across the Nordic countries on generic drug expenditure and pharmaceutical availability. While the most stringent price regulation regime decreases drug expenditure by 40% compared to the less stringent regime, they find no effects on the availability of drugs. Our counterfactual analysis speaks directly to this policy debate by quantifying how targeted price increases would affect the frequency and magnitude of shortages. Finally, while prior empirical work on shortage incidence relies heavily on officially reported shortage events, our detection method requires only sales data, allowing us to measure both the occurrence and the magnitude of shortages even where reporting is incomplete or definitions differ across countries.

The paper proceeds as follows: Section 2 presents the model of the international market for drugs and the methodology to identify drug shortages from sales data (Section 2.2); in Section 3 we bring our analysis to the French context and study the effect of prices on shortages and the implications of possible counterfactual price changes; Section 4 concludes.

2 Capacity Choice, Shortages and the Detection Method

Pharmaceutical manufacturers typically scale their production facilities to meet global demand. However, the high fixed costs of building and maintaining these plants, including infrastructure, equipment, and regulatory compliance, make it impractical to invest in excess capacity that could cover demand under all possible scenarios. For instance, ensuring uninterrupted supply during temporary facility shutdowns for maintenance or safety inspections would require prohibitively expensive spare capacity. As documented in Stomberg (2016), regulatory oversight alone imposes substantial costs. Given these constraints, firms instead invest in capacity up to a point where the marginal cost of expansion becomes increasingly steep, leading to a convex relationship between fixed costs and expected production capacity. We begin by developing a simple model of capacity choice for an international pharmaceutical firm to examine how price regulation influences production decisions and, consequently, drug shortages.

2.1 Capacity Choice and Allocation of Production

A pharmaceutical firm sells drug j in N countries $c \in \{1, \ldots, N\}$ in each period t.[1] Prices p_j^c differ across countries because regulators in each market set reimbursement levels independently.

Countries are indexed in decreasing order of price:

$$p_j^N \leq \cdots \leq p_j^c \leq \cdots \leq p_j^1. \tag{1}$$

Demand in country c is $q_{jc}^d(p_j^c)$, a standard downward-sloping demand function. Before the selling period begins, the firm faces a marginal production cost c_j, common across markets.

The firm chooses a global production capacity $\overline{K}_j$, incurring a fixed installation cost $C\left(\overline{K}_j\right)$ that is convex and increasing: each additional unit of capacity becomes progressively more expensive, reflecting the escalating costs of pharmaceutical plant infrastructure, equipment quali-

[1] We omit the time subscript for ease of notation.

fication, and regulatory compliance. Therefore, it does not invest in unlimited backup capacity; instead, it chooses $\overline{K}_j$ to balance the marginal cost of installation against the expected benefit of avoiding shortages. After capacity is installed, a supply shock κ reduces effective production to $\overline{K}_j - \kappa$ with probability π. This shock represents real-world disruptions, such as factory maintenance shutdowns, contamination events, or forced regulatory closures. With probability $1-\pi$, no shock occurs and capacity $\overline{K}_j$ is fully available; under this scenario, the firm supplies fully all demand in every country. When shortages occur the firm incurs a penalty $\gamma(q^d_{jc} - q^s_{jc})$, increasing in the magnitude of the gap between demand and supply. This penalty can represent explicit regulatory fines, contractual failure-to-supply clauses, or implicit reputational damage. We parameterize $\gamma(\cdot)$ as a smooth, convex function with $\gamma(0) = \gamma'(0) = 0$.

The firm chooses its capacity $\overline{K}_j$ before the shock is realized. If the shock occurs, it then chooses the quantities $\{q^s_{jc}\}_{c=1}^N$ to supply to each country, subject to the binding capacity constraint. It takes into account price-cost margins $(p^c_j - c_j)$, the demand in each country $q^d_{jc}\left(p^c_j\right)$ and the shortage penalty $\gamma()$.

In periods without shock, the firm fully satisfies demand in all markets. The firm capacity choice problem, knowing that shocks may happen, is as follows:

$$\begin{aligned}&\max_{\overline{K}_j}\Bigg\{(1-\pi)\sum_{c=1}^{N}(p^c_j - c_j)q^d_{jc}(p^c_j)+\\&\pi \max_{\left\{0\le q^s_{jn}\le q^d_{jn}(p^n_j)\right\}_{n=1,\ldots,N},\sum_{c=1}^N q^s_{jc}\le \overline{K}_j-\kappa_j}\left\{\sum_{c=1}^{N}\left((p^c_j - c_j)q^s_{jc} - \gamma\left(q^d_{jc}(p^c_j) - q^s_{jc}\right)\right)\right\} - C(\overline{K}_j)\Bigg\}\end{aligned}$$

which is equivalent to:

$$\max_{\overline{K}_j}\left\{\pi \max_{\left\{0\le q^s_{jn}\le q^d_{jn}(p^n_j)\right\}_{n=1,\ldots,N},\sum_{c=1}^N q^s_{jc}\le \overline{K}_j-\kappa_j}\left\{\sum_{c=1}^{N}\left((p^c_j - c_j)q^s_{jc} - \gamma\left(q^d_{jc}\left(p^c_j\right) - q^s_{jc}\right)\right)\right\} - C\left(\overline{K}_j\right)\right\}$$

The outer maximization determines how much capacity to install ex ante. The inner maximization determines how the firm allocates the limited capacity $\overline{K}_j - \kappa$ across countries in a shock state. When capacity is scarce, the firm optimally allocates each unit to the country where it generates the highest marginal payoff: the price-cost margin $p_j^c - c_j$ plus the marginal shortage penalty avoided $\gamma'(q_{jc}^d - q_{jc}^s)$.

We impose the following mild ordering condition, which guarantees a well-behaved threshold structure for the supply allocation prioritizing high-price countries as shown in proposition 1.

Assumption 1 (Monotone ordering of effective priorities). *We assume the following: If*

$$p_j^c + \gamma'\left(q_{jc}^d(p_j^c)\right) < p_j^{c+1} + \gamma'\left(q_{jc+1}^d(p_j^{c+1})\right),$$

then

$$p_j^{c+1} + \gamma'\left(q_{jc+1}^d(p_j^{c+1})\right) < p_j^{c+2} + \gamma'\left(q_{jc+2}^d(p_j^{c+2})\right).$$

For instance, this holds if $p_j^c + \gamma'(q_{jc}^d(p_j^c))$ is convex in the country index c.

The indices (m^*, n^*) are determined in equilibrium, jointly with the shadow value ρ of the capacity constraint in the shock state (characterized in Proposition 1), by the complementarity conditions

$$\begin{array}{ll} p_j^c - c_j \geq \rho & \text{for } c = 1, \ldots, m^* - 1 \text{ (full service)}, \\ p_j^c - c_j < \rho < p_j^c - c_j + \gamma'\left(q_{jc}^d\left(p_j^c\right)\right) & \text{for } c = m^*, \ldots, m^* + n^* - 1 \text{ (rationing)}, \\ p_j^c - c_j + \gamma'\left(q_{jc}^d\left(p_j^c\right)\right) \leq \rho & \text{for } c = m^* + n^*, \ldots, N \text{ (exclusion):} \end{array}$$

a country is fully served when its margin weakly exceeds the shadow value of capacity, rationed when its margin falls below ρ while the marginal penalty of leaving its entire demand unserved would exceed it, and excluded otherwise. These conditions involve only the primitives $\left(p_j^c, q_{jc}^d\right)$ and the scalar ρ, which is pinned down by the binding capacity constraint as stated in Proposition 1; under Assumption 1 the three sets are consecutive intervals of the price ranking, so that (m^*, n^*) is well defined. We can then establish the following supply decisions (see proof in Appendix A.1):

Proposition 1 (Threshold allocation under a supply shock). *Fix $\overline{K}_j$ and consider the shock state in which the available capacity is $\overline{K}_j - \kappa_j$. Using Assumption 1 and indices (m^*, n^*), we* *have the following supply regimes:*

1. ***Full-service regime:*** *for* $c = 1, \ldots, m^* - 1, \quad q_{jc}^s = q_{jc}^d(p_j^c)$*, the highest-price countries are fully served,*

2. ***Rationed regime:*** *for* $c = m^*, \ldots, m^* + n^* - 1$*,*

$$(p_j^c - c_j) + \gamma'\Big(q_{jc}^d(p_j^c) - q_{jc}^s\Big) = \rho. \tag{2}$$

   where ρ *is the shadow value of capacity in the shock state, determined by the binding capacity constraint as the solution to:*

$$\sum\nolimits_{c=m^*}^{m^*+n^*-1} \gamma'^{-1}\left(\rho - \left(p_j^c - c_j\right)\right) = \sum\nolimits_{c=1}^{m^*+n^*-1} q_{jc}^d\left(p_j^c\right) - \overline{K}_j + \kappa_j$$

3. ***Excluded regime:*** *for* $c \geq m^* + n^*, \quad q_{jc}^s = 0.$

For middle-price countries receiving partial supply, the firm equalizes the marginal value of allocating an additional unit across all rationed countries. Consequently, each faces a positive shortage $q_{jc}^d(p_j^c) - q_{jc}^s > 0$ that increases as the country index c rises (as the price falls).

Using the threshold structure that partitions countries into distinct supply regimes, we derive how the supply allocated to a rationed country responds to price changes in countries in other regimes. To obtain closed-form expressions and make the comparative statics transparent, we impose a quadratic form for the shortage penalty:

Assumption 2 (Quadratic shortage cost)**.** *Shortage cost is* $\gamma(x) = \gamma\dfrac{x^2}{2}$ *with* $\gamma > 0$*.*

Assumption 2 should be read as a normalization for expositional purposes rather than a substantive restriction. As we show in Appendix A.2, all the qualitative results that follow–the signs and elasticity conditions in Propositions 2 and 3 and the total effects combining them–extend to any strictly convex penalty $\gamma(\cdot)$ with $\gamma(0) = \gamma'(0) = 0$: the constant $1/\gamma$ is replaced by the local inverse curvature $1/\gamma''(\cdot)$, evaluated at the equilibrium shortage of the relevant country, and the equal sharing weights $1/n^*$ by weights proportional to these inverse curvatures. Only the closed-form expressions rely on the quadratic specification.

Under this assumption, Proposition 2 derives the supply in rationed countries and characterizes both the own-price response and the cross-price effects from other countries:

Proposition 2 (Supply in rationed regions)**.** *Using the quadratic shortage cost assumption, the quantity supplied in rationed countries $c \in \{m^*, \ldots, m^* + n^* - 1\}$ is:*

$$q^s_{jc} = q^d_{jc}\left(p^c_j\right) - \frac{1}{n^*}\sum_{c=1}^{m^*+n^*-1} q^d_{jc}\left(p^c_j\right) + \frac{1}{\gamma}\left[\left(p^c_j - c_j\right) - \frac{1}{n^*}\sum_{c=m^*}^{m^*+n^*-1}\left(p^c_j - c_j\right)\right] + \frac{1}{n^*}\left(\overline{K}_j - \kappa_j\right)$$

and we obtain the following comparative statics for $c \in \{m^*, \ldots, m^* + n^* - 1\}$:

$$\left.\frac{\partial q^s_{jc}}{\partial p^c_j}\right|_{m^*,n^*,\overline{K}_j} = \left(\frac{\partial q^d_{jc}\left(p^c_j\right)}{\partial p^c_j} + \frac{1}{\gamma}\right)\left(1 - \frac{1}{n^*}\right) > 0 \textit{ if } \frac{\partial q^d_{jc}\left(p^c_j\right)}{\partial p^c_j} > -\frac{1}{\gamma}$$

and the cross derivatives are given by:

$$\left.\frac{\partial q^s_{jc}}{\partial p^{c'}_j}\right|_{m^*,n^*,\overline{K}_j} = \begin{cases} -\dfrac{1}{n^*}\dfrac{\partial q^d_{jc'}\left(p^{c'}_j\right)}{\partial p^{c'}_j} > 0 \textit{ if } c' \in \{1, \ldots, m^* - 1\} \\ < 0 \textit{ if } c' \in \{m^*, \ldots, m^* + n^* - 1\} \textit{ and } \dfrac{\partial q^d_{jc'}\left(p^{c'}_j\right)}{\partial p^{c'}_j} > -\dfrac{1}{\gamma} \\ = 0 \textit{ if } c' \in \{m^* + n^*, \ldots, N\} \end{cases}$$

and

$$\left.\frac{\partial q^s_{jc}}{\partial \overline{K}_j}\right|_{m^*,n^*} = \frac{1}{n^*}$$

These comparative statics show how the supplied quantity of a rationed country responds to different price changes, conditional on a fixed capacity. First, consider an own-price change for the analyzed rationed country. The supply allocated to a rationed country increases with its local price, provided its demand is not too elastic. In this model, the condition bounding this elasticity depends on the inverse of the shortage penalty parameter $\frac{\partial q_{jc}^{d}\left(p_j^c\right)}{\partial p_j^c} > -\frac{1}{\gamma}$. During a supply shock, a price increase in the rationed country, with other countries' prices fixed, incentivizes the firm to shift supply away from other rationed countries toward this one. This is optimal because the demand is still high. However, if the price increase is too big or if the demand elasticity is very high, increasing the price would only reduce the total demand, counteracting the incentive to redirect supply. Second, consider cross-price effects from a rival country. If the rival country is in the full-service regime, a price increase reduces its quantity demanded, and the freed capacity is distributed equally among the rationed countries. Conversely, if the price of a rival rationed country increases—and its demand is not excessively elastic—the supply allocated to the focal country decreases. In this scenario, it becomes optimal for the firm to redirect capacity toward the rival country experiencing the price increase. The allocation rule in Proposition 2 is thus a regulated-price analogue of the margin-based (rather than cost-based) ordering of production across heterogeneous sources studied by Asker et al. (2019).

With an additional assumption on the fixed costs of capacity, we analyze the choice of capacity in Proposition 3:

Proposition 3 (Choice of capacity)**.** *The optimal capacity choice satisfies*

$$\begin{aligned}\overline{K}_j &= C'^{-1}\left(\pi\rho\right)\\ &= C'^{-1}\left(\pi\left(\frac{\gamma}{n^*}\sum_{c=1}^{m^*+n^*-1} q_{jc}^{d}\left(p_j^c\right)+\frac{1}{n^*}\sum_{c=m^*}^{m^*+n^*-1}\left(p_j^c-c_j\right)-\frac{\gamma}{n^*}\left(\overline{K}_j-\kappa_j\right)\right)\right)\end{aligned}$$

and, with quadratic fixed cost $C(x)=\frac{x^2}{2}$, *is*

$$\overline{K}_j=\pi\frac{\gamma\sum_{c=1}^{m^*+n^*-1} q_{jc}^{d}\left(p_j^c\right)+\sum_{c=m^*}^{m^*+n^*-1}\left(p_j^c-c_j\right)+\gamma\kappa_j}{n^*+\pi\gamma}$$

with

$$\left.\frac{\partial \overline{K}_j}{\partial p_j^c}\right|_{m^*,n^*} = \begin{cases} \dfrac{\pi\gamma}{n^*+\pi\gamma}\dfrac{\partial q_{jc}^d\left(p_j^c\right)}{\partial p_j^c} < 0 \;\; \textit{if } c \in \{1,\ldots,m^*-1\} \\ \dfrac{\pi\gamma}{n^*+\pi\gamma}\left(\dfrac{\partial q_{jc}^d\left(p_j^c\right)}{\partial p_j^c} + \dfrac{1}{\gamma}\right) \;\; \textit{if } c \in \{m^*,\ldots,m^*+n^*-1\} \;\; \textit{and } > 0 \;\; \textit{if } \dfrac{\partial q_{jc}^d\left(p_j^c\right)}{\partial p_j^c} > -\dfrac{1}{\gamma} \\ 0 \;\; \textit{if } c \in \{m^*+n^*,\ldots,N\} \end{cases}$$

The allocation rule in Proposition 2 and the capacity choice in Proposition 3 imply that prices in different countries can affect the global capacity in different ways. In general, capacity is increasing in the prices of countries that face rationing under supply shocks, but decreasing in the prices of countries that are always fully served, as higher prices depress quantity demanded. Moreover, capacity can also be decreasing in prices of rationed countries when they have very elastic demand or large shortage penalties.

The indices (m^*, n^*) themselves respond to prices only through the complementarity conditions defining the supply regimes. Generically, when none of these conditions holds with equality, a small price change leaves (m^*, n^*) unchanged, which is what justifies deriving the comparative statics of Propositions 2 and 3 holding these indices fixed. When a boundary condition binds, the direction in which the rationed set adjusts is governed by the induced movement of the shadow value ρ (and of capacity once $\overline{K}_j$ adjusts) rather than by demand monotonicity alone: under Assumption 2 and at fixed capacity, $\partial\rho/\partial p_j^c = \left(1+\gamma\dfrac{\partial q_{jc}^d}{\partial p_j^c}\right)/n^* < 1$ for a rationed country c, so an increase in a rationed country's own price raises its margin by more than ρ and weakly moves that country toward full service, while whether adjacent countries enter the rationed set from full service or from exclusion depends on the sign of the change in ρ, and hence on demand elasticities.

Finally, consider the total own-price effect on supply, combining the fixed-capacity reallocation of Proposition 2 with the capacity adjustment of Proposition 3. For countries that remain fully served in the shock state ($c \in \{1,\ldots,m^*-1\}$), supply simply tracks demand, so that

$\left.\frac{\partial q_{jc}^s}{\partial p_j^c}\right|_{m^*,n^*} = \frac{\partial q_{jc}^d\left(p_j^c\right)}{\partial p_j^c} < 0$, while for excluded countries ($c \in \{m^*+n^*, \ldots, N\}$) it is zero. For a rationed country $c \in \{m^*, \ldots, m^*+n^*-1\}$:

$$\begin{aligned}
\left.\frac{\partial q_{jc}^s}{\partial p_j^c}\right|_{m^*,n^*} &= \left.\frac{\partial q_{jc}^s}{\partial p_j^c}\right|_{m^*,n^*,\overline{K}_j} + \frac{1}{n^*}\left.\frac{\partial \overline{K}_j}{\partial p_j^c}\right|_{m^*,n^*} \\
&= \left(\frac{\partial q_{jc}^d\left(p_j^c\right)}{\partial p_j^c} + \frac{1}{\gamma}\right)\left(1 - \frac{1}{n^*} + \frac{1}{n^*}\frac{\pi\gamma}{n^*+\pi\gamma}\right)
\end{aligned}$$

This shows that $\frac{\partial q_{jc}^d\left(p_j^c\right)}{\partial p_j^c} > -\frac{1}{\gamma}$ is a sufficient condition for $\left.\frac{\partial q_{jc}^s}{\partial p_j^c}\right|_{m^*,n^*} > 0$ in the rationed countries $c \in \{m^*, \ldots, m^*+n^*-1\}$: even after accounting for the adjustment of the installed capacity, a higher domestic price attracts supply to a rationed country provided its demand is not too elastic.

Similarly, the total effect of a price change in another country $c' \neq c$ on the supply allocated to a rationed country c combines the fixed-capacity reallocation effect of Proposition 2 with the capacity adjustment of Proposition 3:

$$\begin{aligned}
\left.\frac{\partial q_{jc}^s}{\partial p_j^{c'}}\right|_{m^*,n^*} &= \left.\frac{\partial q_{jc}^s}{\partial p_j^{c'}}\right|_{m^*,n^*,\overline{K}_j} + \frac{1}{n^*}\left.\frac{\partial \overline{K}_j}{\partial p_j^{c'}}\right|_{m^*,n^*} \\
&= \begin{cases}
-\frac{1}{n^*+\pi\gamma}\frac{\partial q_{jc'}^d\left(p_j^{c'}\right)}{\partial p_j^{c'}} > 0 \text{ if } c' \in \{1, \ldots, m^*-1\} \\
-\frac{1}{n^*+\pi\gamma}\left(\frac{\partial q_{jc'}^d\left(p_j^{c'}\right)}{\partial p_j^{c'}} + \frac{1}{\gamma}\right) \text{ if } c' \in \{m^*, \ldots, m^*+n^*-1\} \\
0 \text{ if } c' \in \{m^*+n^*, \ldots, N\}
\end{cases}
\end{aligned}$$

Two features of this total effect are worth noting. First, under the quadratic fixed cost, the capacity adjustment dampens but never reverses the fixed-capacity reallocation effect. A higher price in a fully served country c' therefore unambiguously increases the supply allocated to rationed countries: it lowers the quantity demanded in c', and the capacity thereby freed in shock states more than compensates for the smaller installed capacity. Second, for a rationed rival c', the sign depends on its demand elasticity: if demand in c' is not too elastic $(\frac{\partial q_{jc'}^d}{\partial p_j^{c'}} > -\frac{1}{\gamma})$, the higher margin attracts scarce capacity toward c' and the supply allocated to the other rationed countries falls; if instead demand in c' is sufficiently elastic, the contraction of its quantity demanded dominates and the supply allocated to the other rationed countries increases. Finally, since the prices of excluded countries affect neither the allocation of capacity nor the capacity choice, a foreign price can matter for domestic shortages only if the foreign country is itself served in shock states.

We consider also the alternative case where there is no shortage penalty ($\gamma \equiv 0$). In this case, as shown in Appendix A.1, at most one country is rationed: there exists a marginal country m^* such that countries $1, \ldots, m^* - 1$ are fully served, country m^* is rationed while countries $m^* + 1, \ldots, N$ are excluded (in the notation of the previous propositions, $n^* = 1$). The supply in the rationed country is then increasing in the price in that country. The response of the optimal capacity depends on whether country m^* is strictly rationed. When it is, capacity is pinned down by the marginal country's margin, $\overline{K}_j = C'^{-1}\left(\pi\left(p_j^{m^*} - c_j\right)\right)$, so that it increases with $p_j^{m^*}$ and is unaffected by the other prices; when instead capacity exactly covers the demand of the served countries, it decreases with their prices through the demand effect, as for fully served countries in the case $\gamma > 0$. In either regime, the shortage in the marginal country is weakly decreasing in the price of every served country.

In summary, the effect of prices on the probability and magnitude of shortages depends on demand elasticity, shortage penalties, and the country's position within the global price distribution. Determining the impact of a domestic price increase remains an empirical question, necessitating the precise measurement of shortage events, a methodology we develop in the following section.

2.2 Detecting Shortages in Sales Data

To take our model's predictions to the data, we develop a methodology to detect drug shortages and quantify their magnitude using only product-level sales data, without relying on auxiliary datasets of regulatory shortage notifications. Indeed, while the main policy response to the shortage problem has been to increase transparency, allowing the practitioners as well as competing manufacturers to anticipate disruptions to their operations (Acosta et al., 2019), data on the reported shortages are not always available and inconsistent definitions complicate international comparisons[2]. Another advantage of using sales data directly is that we can also estimate the magnitude of the shortages, i.e., the size of the gap between the quantity demanded and the quantity supplied. We will show explicitly in which cases this method is valid and when it cannot be used.

As a shortage corresponds to a situation in which the quantities supplied (quantities observed in the sales data) are less than the quantities demanded, identifying shortages requires having an estimate of the demand function. In a market where we know shortages occur, obtaining the correct demand function is complicated: including shortage periods, when sales do not reflect quantity demanded, biases the demand estimates. Therefore, we need a way of isolating time periods without shortages for demand estimation.

Our method relies on defining a sales threshold, denoted τ, and assuming that shortages occur only when sales fall below this threshold. In other words, we focus exclusively on shortages driven by supply shocks that bring sales to abnormally low levels. The importance of supply shocks as drivers of shortages is confirmed in the literature, for example, the review in Acosta et al. (2019) identifies four main, interconnected, categories of causes of shortages: the market, supply chain management, manufacturing process, and political and ethical issues. Indeed, even if there are demand variations at the world level, the supply side production and storage capacities are typically sufficient to smooth demand shocks and provide drugs where needed. However, supply-side shocks do happen (e.g. accidents in the supply chain, factory closures), and lead to shortages even with constant demand.

[2]Moreover, the timing of regulatory declarations is only loosely related to observed disruptions in sales. Using GERS sell-in data for France, Baudet and Dherbécourt (2025) show that sales to pharmacies do not decline until about two months after a declaration to the ANSM (falling to 85% of the pre-declaration volume for risk-of-shortage declarations and to 76% for declarations of actual shortages), consistent with declarations largely concerning anticipated shortages. Declarations of actual shortages are even accompanied by a contemporaneous 17% increase in sales to pharmacies, likely reflecting stockpiling by pharmacies and wholesalers. Declaration dates therefore provide a poor benchmark for the monthly timing and magnitude of actual supply disruptions, which our sales-based method measures directly.

2.2.1 Set-up

Considering a national market (for example France), we established in Section 2.1 that a drug's global production capacity depends on demands and prices across all countries, and thus could be very large compared to the demand from France. Indeed, France constitutes a small share of aggregate world demand in general: in 2020, French pharmaceutical spending was approximately €30 billion compared to global sales of roughly €1,250 billion, representing less than 2.5% of the total market. The stylized model in Section 2.1 does not introduce demand uncertainty and assumes that the maximum possible demand from France is small compared to the world production capacity and also that the supply shock has a distribution, specified below, implying that demand shocks cannot cause shortages.

Denoting q_{jt}^d and Q_{jt}^d the French and rest of the world demands for drug j in period t, respectively, we assume that the world production capacity is $\overline{K}_{jt}$ and that, as in Section 2.1, a supply shock reduces effective production to $\overline{K}_{jt} - \kappa_{jt}$, where the shock size $\kappa_{jt} \geq 0$ is now a random variable with support $[0, \bar{\kappa}]$, the realization $\kappa_{jt} = 0$ meaning that no shock occurs. Capacity is scaled to meet world demand in the absence of a shock: $\overline{K}_{jt} \geq q_{jt}^d + Q_{jt}^d$.

A world level shortage happens if the world production is insufficient that is if:

$$q_{jt}^d + Q_{jt}^d > \overline{K}_{jt} - \kappa_{jt}$$

Assuming that κ_{jt} is independent of the distribution of demands q_{jt}^d and Q_{jt}^d, a shortage will happen with probability $P(q_{jt}^d + Q_{jt}^d > \overline{K}_{jt} - \kappa_{jt}) = P(\kappa_{jt} > q_{jt}^d + Q_{jt}^d > \overline{K}_{jt}) = 1 - F_\kappa(\overline{K}_{jt} - q_{jt}^d - Q_{jt}^d)$ where F_κ is the cumulative distribution function of κ_{jt}.

Denoting $\underline{Q}^d$ the minimum rest-of-the-world demand, we then assume that κ_{jt} has an atom at zero, $P(\kappa_{jt} = 0) = 1 - \pi$ with $0 < \pi < 1$, and that F_κ is flat on $(0, \overline{K}_{jt} - \underline{Q}^d]$. Therefore, conditional on a shock occurring, $\kappa_{jt} \geq \overline{K}_{jt} - \underline{Q}^d$. This implies that the shortage probability is $1 - F_\kappa(\overline{K}_{jt} - q_{jt}^d - Q_{jt}^d) = \pi$ for any demand realization: it is independent of the French demand, and depends only on the occurrence of supply shocks that are large relative to the rest-of-the-world excess capacity of production. κ_{jt} is thus such that $P(\kappa_{jt} = 0) = 1 - \pi$ and $P(\kappa_{jt} \geq \overline{K}_{jt} - \underline{Q}^d) = \pi$.

This assumption means that with probability $1-\pi$ no shock occurs, the full capacity $\overline{K}_{jt}$ is available and all demands are served. With probability π, a shock occurs and destroys enough production capacity that the effective production $\overline{K}_{jt} - \kappa_{jt}$ falls below even the minimum rest-of-the-world demand $\underline{Q}^d$. This means that even if the French demand were zero, the shock on production capacity is such that there is a world shortage. The probability of a shortage at the world level $P(q^d_{jt} + Q^d_{jt} > \overline{K}_{jt} - \kappa_{jt}) = \pi$ is then exogenous to the French demand, consistently with the exogenous shock probability π of Section 2.1. However, when there is a world shortage, part of the shortage likely affects the French market.

2.2.2 Detecting Drug Shortages

To detect shortages and their magnitude, we first specify a demand prediction process for drugs in France. The demand for a given drug in French pharmacies depends on product identity (reflecting its quality and therapeutic necessity), the time period (capturing demographic and epidemiological variations), and other unobservable factors. Moreover, the demand for a drug may depend also at the monthly level on its dynamics because treatments sometimes last longer and the diffusion of the drug may depend on its use. As prices may also affect demand, we include them in the prediction. We thus specify a process for drug j in month t as follows:

$$\log(q^d_{jt}) = \sum\nolimits_{k=1}^{K} \alpha_k \log(q^d_{jt-k}) - \beta \log(p_{jt}) + X'_{jt}\theta + \varepsilon^d_{jt} \tag{3}$$

where q^d_{jt} is the quantity of demand in France of drug j at period t, which depends on lagged demands up to K periods, its price p_{jt}, a vector of observable characteristics X_{jt} and some unobservable demand shock ε^d_{jt}.

However, the quantity sold in period t may be lower than the quantity theoretically demanded q^d_{jt} if there is a shortage. We thus denote by q_{jt} the quantity sold in month t for each product j, such that $q_{jt} = q^d_{jt}$ in the absence of a shortage, and $q_{jt} < q^d_{jt}$ during a shortage.

We then define a threshold τ_j for drug j as being the threshold such that we can consider there is no shortage if sales are larger than τ_j. This threshold is not necessarily equal to the theoretical supply capacity as it could be that there is no shortage even if sales are smaller. Rather, it represents the maximum possible supply of a drug j where shortages are still possible.

Indeed, with the assumptions made earlier on the variability of the demand in France, shortages cannot be caused by unexpected high demand shocks in France but result from global shortages caused by negative supply shocks or large rest-of-the-world demand shocks. Thus, if French sales volumes exceed this threshold, the market cannot be experiencing a shortage.

In particular, suppose that the French and rest-of-the-world demands are proportional with a factor M (not to be confused with the number of countries N of Section 2.1), such that $Q^d_{jt} = Mq^d_{jt}$, and that in a shortage state the available supply is allocated across markets in proportion to demands, so that France receives the share $\dfrac{1}{M+1}$ of the effective production. Conditional on a shock occurring, we assumed that $\kappa_{jt} \geq \overline{K}_{jt} - \underline{Q}^d$, so that the effective production is at most $\underline{Q}^d$ and French sales in any shortage state satisfy $q_{jt} = \dfrac{\overline{K}_{jt} - \kappa_{jt}}{M+1} \leq \dfrac{\underline{Q}^d}{M+1}$. Defining the threshold as $\tau_j \equiv \dfrac{\underline{Q}^d}{M+1}$, observed sales above the threshold are therefore incompatible with a shortage state: if $q_{jt} > \tau_j$, there is no shortage and $q_{jt} = q^d_{jt}$. Note that τ_j is precisely the maximum possible French supply in a shortage state, as in the verbal definition above, and that it is constant over time since it depends only on the minimum rest-of-the-world demand. The converse inference does not hold: sales below τ_j may simply reflect low demand rather than a shortage, which is why the threshold serves only to select periods that are free of shortages.

We now make the following assumption:

$$E[\varepsilon^d_{jt} | \alpha_k, \beta, \theta, q_{jt-k}, \ldots, q_{jt-1}, q_{jt} > \tau_j] = 0, \tag{4}$$

meaning that unobservable demand shocks are mean-independent of observable demand determinants when sales exceed the maximum threshold compatible with shortages. Specifically, if sales are sufficiently high to rule out a shortage during a given period, then demand equals observed sales. This holds even in the presence of large unobservable demand shocks, as we assume supply fully accommodates demand in these instances. Alternatively, this implies that unobserved, large positive demand deviations cannot cause shortages when sales exceed the threshold. We believe this assumption is reasonable because large demand increases are in general either permanent or predictable by lagged demand increases.[3]

Then, using (4), we can estimate the demand equation conditional on the observation of τ_j and predict the demand $\widehat{q}_{jt}^{d}\left(\tau_j\right)^4$ using:

$$\log(\widehat{q}_{jt}^{d}(\tau_j)) = \sum\nolimits_{k=1}^{K} \widehat{\alpha}_k \log(q_{jt-k}^{d}) - \widehat{\beta}\log(p_{jt}) + X_{jt}'\widehat{\theta} \tag{5}$$

Given the realized quantity sold q_{jt}, we define the shortage quantity $s_{jt}(\tau_j)$ as a function of τ_j, equal to the difference between demand and realized sales (if demand is larger than supply).

To account for estimation precision, we also compute the estimated variance of the predicted demand of Equation (5), denoted $var(\widehat{q}_{jt}^{d}(\tau_j))$. We infer that a shortage occurs if the observed sales quantity falls below the lower bound of the 95% confidence interval for the predicted demand, and define the detected shortage indicator accordingly:

$$D_{jt}(\tau_j) \equiv 1_{\left\{q_{jt} \leq \widehat{q}_{jt}^{d}(\tau_j) - 1.96\sqrt{var(\widehat{q}_{jt}^{d}(\tau_j))} \wedge q_{jt} \leq \tau_j\right\}} \tag{6}$$

The shortage quantity is then:

$$s_{jt}\left(\tau_j\right) = \max\left\{0, \widehat{q}_{jt}^{d}(\tau_j) - q_{jt}\right\} \times D_{jt}\left(\tau_j\right) \tag{7}$$

Given our assumptions regarding supply shocks as the source of shortages, we would ideally

[3] To illustrate the rationale of the assumption of Equation (4), simplifying to the case where $q_{jt}^{d} = \alpha_j + \varepsilon_{jt}^{d}$ and when $\tau_j \equiv \dfrac{Q^d}{M+1}$, the assumption 4 is satisfied if ε_{jt}^{d} is bounded below by $\dfrac{Q^d}{M+1} - \alpha_j$ because then $E[\varepsilon_{jt}^{d}|\alpha_j, q_{jt} > \tau_j] = E[\varepsilon_{jt}^{d}|\alpha_j, \varepsilon_{jt}^{d} > \dfrac{Q^d}{M+1} - \alpha_j] = E[\varepsilon_{jt}^{d}|\alpha_j] = 0$.

[4] For shortage periods, or periods within three months of a shortage we use an extrapolation of the predictions, following the logic outlined in the following Section 2.2.3

define τ_j as the λ quantile of the demand of product j

$$\tau_j(\lambda) = \arg\max_{\tau}\{\ P_j[q^d_{jt} \geq \tau] \geq 1 - \lambda\}$$

However, we do not observe the demand q^d_{jt} but only the realized sales. In practice, we therefore implement $\tau_j(\lambda)$ as the λ quantile of the observed sales q_{jt} of product j. Since sales never exceed demand, each quantile of the sales distribution lies weakly below the corresponding quantile of the demand distribution, and the data-driven choice of λ compensates for this gap: we search for the value of λ such that the demand process estimated on the subsample $q_{jt} > \tau_j(\lambda)$ yields the best fit, using the Akaike information criterion as described in the estimation details below.

2.2.3 Estimation details

We calibrate the λ parameters separately for each ATC2 class. For each product j, λ defines the threshold τ_j, the λ percentile of log sales, $\log(q_{jt})$. We regress $\log(q_{jt})$ on the first three lags, log price, indicators for genericized markets, generics, and combination products, product strength and the number of doses in the package. We also include molecule fixed effects, and ATC4-specific time fixed effects. We restrict the sample to observations where both $\log(q_{jt})$ and its first three lags exceed τ_j. Our selection criterion for λ is the AIC.

After selecting λ, we re-estimate Equation (3) separately for each ATC2 class, restricting the sample to observations where $q_{jt} > \tau_j(\lambda)$ holds for the current sales and all three lags. We use these estimates to predict ATC2-specific demand using Equation (5). To avoid using sales values from shortage periods in the prediction, we first focus on periods where both current sales and all three lags exceed τ_j. Next, we iterate over a series of equations to extrapolate the fitted values, giving priority to valid observed sales values (exceeding τ_j). For example, we start with observations where observed sales in both $t-2$ and $t-3$ exceed τ_j, and use the fitted value (if available) for the first lag. We consider every possible combination and repeat this process until no further changes can be made. Using τ_j and these demand predictions, we apply the rule in Equation (6) to detect shortage periods.

3 Empirical Analysis of Drug Shortages in France

3.1 The French Context

In France, the French National Agency of Medicines and Health Product Safety (ANSM) defines a shortage as the inability of a pharmacy to deliver a drug within 72 hours. Shortages are classified by context: a supply problem resulting from supply chain issues (*rupture d'approvisionnement*), or a stock problem where the medicine cannot be manufactured (*rupture de stock*). Drug shortages in France have been well documented since 2012, when marketing authorization holders of pharmaceuticals were legally required to report actual or predicted shortages to the ANSM.

Quantitative evidence on drug shortages in France is very limited. Benhabib et al. (2020) use ANSM shortage reports to analyze their occurrence and evolution, finding that 3,530 pharmaceutical products were reported on shortage during 2012-2018, including 1,833 different active substances. The number of reported shortages increased fourfold between 2012 and 2018, with a particularly sharp spike in 2017 and 2018. This trend was accompanied by an increase in the number of active substances in shortage. Manufacturing and material supply issues were the main reported causes of shortages each year (30%).

3.2 Data and Descriptive Statistics

We analyze drug shortages using monthly retail pharmacy sales data in France from January 2019 to December 2024, provided by GERS. The data include monthly sales values and quantities for each product, defined at the delivery-unit level by molecule, producer (brand), form, strength, mode of administration and packaging.

The dataset covers 1,544 molecules (ATC level 5 class in the WHO classification) representing 17,173 unique products and 982,726 total observations. Table 1 presents an overview of the data by ATC level 1 class, the most general level of the WHO drug classification. Three classes: L (Antineoplastic and immunomodulating agents), N (Nervous system), and A (Alimentary tract and metabolism) generate substantially more total revenue than the rest.

Products in ATC level 1 classes J (Antiinfectives), B (Blood and blood forming organs) and S (Sensory organs) generate the highest mean revenues per product. Class L exhibits the highest average product prices, followed by classes H (Systemic hormonal preparations) and R (Respiratory system). The lowest product prices are in class C (Cardiovascular system) and D (Dermatologicals).

Finally, class L averages fewer brands per ATC level 5 class (molecule), a proxy for generic penetration, though the variation across classes is much smaller than for prices or revenues.

Table 1: Mean sales, prices and number of firms per molecule (ATC5) by ATC1 class

	Mean				
ATC	Class Revenue	Product Revenue	Product Sales	Product Price	Firms in ATC5
A Alimentary tract and metabolism	2,690.4	1,388.8	252.6	45.9	3.4
B Blood and blood forming organs	2,010.9	3,759.9	244.0	107.5	2.4
C Cardiovascular system	1,923.5	554.4	73.5	15.6	5.2
D Dermatologicals	691.8	1,146.3	202.5	19.2	3.0
G Genito urinary system	737.8	779.9	78.2	23.0	3.3
H Systemic hormonal preparations	569.4	1,778.5	189.1	233.7	2.7
J Antiinfectives for systemic use	2,178.8	1,912.6	129.4	142.5	3.4
L Antineoplastic and immunomodulating	5,900.7	7,369.8	24.0	786.2	2.5
M Musculo skeletal system	541.3	999.7	183.0	40.5	4.2
N Nervous system	2,910.6	1,046.3	358.6	29.8	4.5
R Respiratory system	2,195.1	2,649.5	276.6	157.5	3.7
S Sensory organs	1,182.5	3,478.1	263.6	37.4	2.3

Notes: Class revenue is the mean annual revenue of all drugs in the ATC1 class in millions of euros. Product is defined as the combination of molecule, brand, firm, formulation and pack size. Product revenue is the mean annual product revenue in thousands of euros. Product sales is the mean product sales quantity in thousands of units.

3.3 Drug shortages estimation results

We apply the methodology described in Section 2.2 to identify shortages in our dataset. Strictly speaking, our method identifies shortages with some statistical error. We use the detected shortage indicator D_{jt} defined in Section 2.2, which equals one whenever the probability that demand exceeded observed sales is at least 95%, even if the shortage magnitude is small.

After selecting λ and predicting demand for each product, we identify the periods in which each drug is in shortage following the definition in Equation (7). Because predicted log demand is normally distributed with an estimated mean and variance, a product is considered to be in shortage in a given month if observed sales fall below the lower bound of the 95% confidence interval. This implies a 95% probability that the demand exceeded observed sales. This means there is a 5% chance that we falsely infer a shortage. Table 2 presents the shortage estimates for France predicted by our method. The first column reports the share of product-month observations in shortage, maintaining a maximum Type I error rate of 5% for each observation. The second and third columns report the total number of products in the class and the share of molecules ever affected by a shortage, respectively. Averaging across the 17,173 products in the sample (weighting the class means by the number of products), the average probability of a drug shortage across product-month observations is 26%.

Shortages affect different ATC 1 classes to varying degrees. Products in class D face an average shortage probability of 19%, while drugs in classes G and J exhibit probabilities of 36% and 41%, respectively. This pattern is consistent with Benhabib et al. (2020), who also found the highest rate of shortages among anti-infectives.

Table 2: Shortage Estimates

ATC	Mean Shortage	Total N products	Share molecules with shortage
A Alimentary tract and metabolism	0.28	2,360	0.88
B Blood and blood forming organs	0.25	676	0.73
C Cardiovascular system	0.21	4,163	0.89
D Dermatologicals	0.19	694	0.77
G Genito urinary system	0.36	1,124	0.93
H Systemic hormonal preparations	0.25	395	0.94
J Antiinfectives for systemic use	0.41	1,348	0.80
L Antineoplastic and immunomodulating	0.23	1,059	0.79
M Musculo skeletal system	0.24	641	0.92
N Nervous system	0.28	3,311	0.94
R Respiratory system	0.25	1,003	0.85
S Sensory organs	0.22	399	0.93

Notes: The mean shortage column corresponds to the mean of the inferred shortage variable (probability of inferring a shortage while there is none being 5%) which is when $q_{jt} \leq \widehat{q}^d_{jt}(\tau_j) - 1.96\sqrt{var(\widehat{q}^d_{jt}(\tau_j))}$.

Figure 1 shows the distribution, across all 17,173 products in our sample, of the average probability that the product is on shortage over the five-year sample period. While many products are rarely in shortage, the bulk of those that experience shortages have a shortage probability that seldom exceeds 40%. Figure 2 shows that almost half of the detected shortages last one month or less, and over 80% are resolved within six months.

Figure 1: Frequency of Shortage by Product

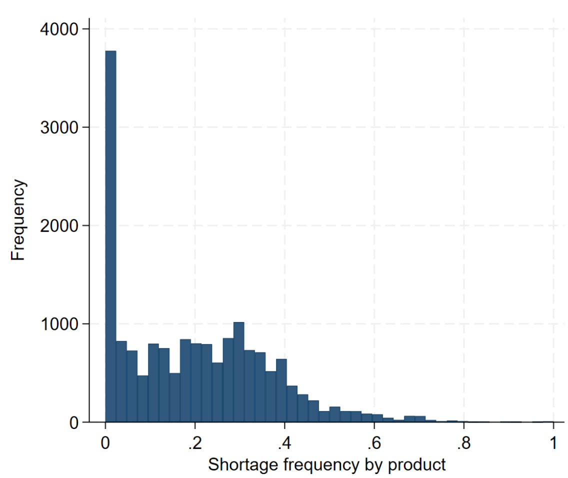


Figure 2: Shortage Spell Duration

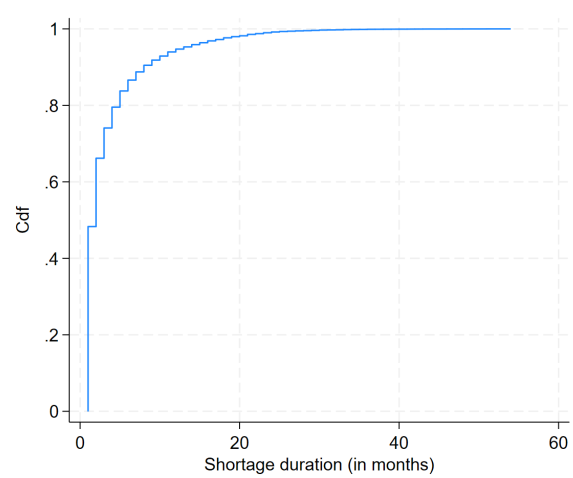


Table 3 shows the mean shortage probability across different product categories. Branded products exhibit a higher shortage probability (28%) than unbranded products (24%). Branded products are defined as those whose molecule is still on patent or that have not experienced entry of generics of the same molecule.[5] They represent around 50% of all products and 50.4% of observations. Additionally, the shortage probability is higher for injectables (29%) than for oral drugs (27%) or other modes of administration (21%).

Table 3: Shortage Estimates by Type of Product

Groups of drugs	Shortage Probability
By mode of administration	
Injectable	0.29
Oral	0.27
Other	0.21
By number of brands	
Single brand	0.28
Multiple brands	0.24

Conditional on identifying a shortage, we examine its extent as a share of predicted demand. Table 4 presents summary statistics on the distribution of these shortage shares by ATC1 class, including the mean along with the 25th, 50th, and 75th percentiles. The data suggests that France frequently operates in a regime of product rationing. Across classes, the mean shortage share ranges between 40% and 59%, but reaches as high as 93% at the 75th percentile. The quantiles vary across classes, specifically, classes with higher mean shortages exhibit a right-shifted distribution across all quantiles. For nearly all classes the median is below the mean, meaning the distribution of shortage shares is right-skewed.

[5] Here, a branded product corresponds to molecules marketed under a single brand. If we observe multiple brands per molecule, we infer the product is unbranded.

Table 4: Shortage Share Conditional on Shortages

	Shortage Share			
ATC	Mean	P25	P50	P75
A Alimentary tract and metabolism	0.46	0.15	0.35	0.81
B Blood and blood forming organs	0.51	0.20	0.46	0.84
C Cardiovascular system	0.43	0.11	0.30	0.80
D Dermatologicals	0.55	0.22	0.52	0.93
G Genito urinary system	0.45	0.16	0.35	0.74
H Systemic hormonal preparations	0.59	0.30	0.59	0.92
J Antiinfectives for systemic use	0.49	0.21	0.42	0.78
L Antineoplastic and immunomodulating agents	0.44	0.17	0.35	0.69
M Musculo skeletal system	0.49	0.16	0.44	0.86
N Nervous system	0.43	0.13	0.33	0.76
R Respiratory system	0.52	0.23	0.48	0.82
S Sensory organs	0.40	0.12	0.29	0.66

Notes: The shortage share conditional on having a shortage is the estimated share of demand that is not satisfied by supply conditionally on having a shortage. P25, P50 and P75 represent respectively the 25, 50 and 75 the percentiles of the distribution.

3.4 Effect of Prices on Shortages

Having identified the shortage events, we examine how prices affect shortage probability. Our theoretical model in Section 2.1 predicts that, provided demand is not too elastic relative to the shortage penalty (i.e., $\frac{\partial q^d_{jc}}{\partial p^c_j} > -\frac{1}{\gamma}$), higher domestic prices decrease both the probability and magnitude of shortages, whereas foreign prices can impact local shortages in either direction. Because these predictions depend on demand elasticities, shortage penalties, and the country's position in the global price distribution, which effects matter and dominate in practice is an empirical question that we investigate below.

We complement our dataset with external information on drug prices from other countries. Using public data on drug prices from the National Health Service (NHS) website, we incorporate prices from the UK, a country with a substantially higher overall level of pharmaceutical prices than France (TLV, 2023). We also extract price data from the public lists of reimbursed drugs in Poland, published by the Polish Health Ministry, that include negotiated manufacturer prices for over 6000 products. Pharmaceutical prices in Poland are on average lower than in France, and patients usually face copays of at least 30% for reimbursed drugs.

We model the detected shortage indicator D_{jt} as a function of the product prices, including the domestic price in France, and the international prices (UK and Poland), and we control for characteristics of the product X_{jt}, in a linear probability model:

$$D_{jt} = \beta_{FR} \log(p_{jt}^{FR}) + \beta_I \log(p_{jt}^{I}) + X_{jt}'\theta + \varepsilon_{jt} \tag{8}$$

Table 5 presents the estimation results of Equation (8) on the full sample (Columns (1) and (2)) and the restricted samples matched with UK price data (Column (3)) or Polish price data (Column (4)). All regressions include time fixed effects and ATC5 fixed effects to control for unobserved time-invariant heterogeneity at the molecule level. We also control for product characteristics that may influence shortage risk: including indicators for whether the drug is a generic or belongs to a genericized market, its administration route (oral or intravenous), and firm size categories based on portfolio count (fewer than 10 or more than 50 products). Column (1) presents a baseline OLS regression evaluating the effect of domestic prices on shortage probability across the full sample. Our estimate suggests that a 10% price increase is associated with a 0.12 pp decrease in shortage probability. Recognizing that prices of drugs may change over time with unobserved factors that may explain the shortage probability and be correlated with prices, Column (2) instruments for domestic price. We use the number of different formats available in the market within ATC5 class, for example different strengths and modes of administration, and dummy variables that represent bins on the number of products within an ATC5 class.

The validity of our instruments (the number of formats and the number of products within an ATC5 class) relies on the exclusion restriction that, conditional on our extensive controls, these market-structure variables affect a product's shortage probability primarily by shifting its regulated price. In pharmaceutical markets, the number of available presentations and competing products are key determinants of therapeutic substitutability and competitive pressure, which directly influence price negotiations with regulators. While one might be concerned that the number of products or formats also directly captures physical supply redundancy or manufacturing complexity, our empirical specification isolates the pricing channel by absorbing these physical supply mechanisms. Specifically, ATC5 fixed effects absorb persistent molecule-level differences in baseline supplier structure, production technology, and chronic vulnerability to shortages. Furthermore, our controls for generic status and genericized markets capture the primary physical channel of supplier redundancy, while administration-route and firm-size indicators absorb differences in operational scale and packaging complexity. We argue that the residual within-class variation in the competitive environment primarily shifts the firm's relative bargaining power and pricing incentives, rather than reflecting high-frequency physical supply disruptions.

The IV effect of price on the shortage probability is larger than in the OLS regression: a 10% increase in price reduces the shortage probability by 2.2 percentage points. Through the lens of the model, this domestic price effect operates mainly through the allocation margin rather than through global capacity: since France accounts for a small share of world demand, the French price has a negligible effect on the capacity the firm installs, but a higher French price raises France's priority in the allocation of scarce capacity during supply shocks, making it less likely that a global shortage translates into a French one. Column (3) includes the UK price, and while the French price effect is stronger compared to the previous columns, a 10% increase in the UK price reduces the shortage probability by 0.6 percentage points. Foreign prices can lower the French shortage probability through two distinct margins of the model. The first is the capacity margin: in the framework of Section 2.2, the frequency of shortage events is governed by the distribution of supply shocks relative to the world excess capacity, so that by raising the capacity the firm installs (Proposition 3), higher prices in rationed countries reduce the frequency with which supply shocks translate into world shortages, and thus into French ones. This positive externality is the natural reading of the UK effect, given the shortage-magnitude evidence below indicating that, conditional on a shock, the UK attracts scarce capacity when its price rises. In Column (4) we include the Polish price. Although not directly comparable, the Polish price seems to have a larger effect on the French shortage probability than the UK price. This effect can be explained by the second margin affecting the probability of a French shortage: the allocation margin; as shown in Section 2.1, when a rationed country with relatively elastic demand raises its price, its quantity demanded contracts and capacity is freed in shock states, so that French sales are less likely to fall below French demand even when a world shortage occurs. For the international prices in Columns (3) and (4), we adopt an instrumental variables approach analogous to that used for domestic prices, constructing instruments from the number of product formats and the total product count per ATC5 class.

Table 5: Shortage probability

	(1)	(2)	(3)	(4)
$\log(p_{jt}^{FR})$	-0.012***	-0.223***	-0.324**	-0.147*
	(0.001)	(0.037)	(0.099)	(0.074)
$\log(p_{jt}^{UK})$			-0.061***	
			(0.017)	
$\log(p_{jt}^{PL})$				-0.283***
				(0.033)
genericized market	-0.032***	-0.016***	0.013	-0.039***
	(0.002)	(0.004)	(0.017)	(0.007)
Generic	0.013***	-0.043***	-0.074**	-0.009
	(0.002)	(0.010)	(0.025)	(0.018)
Drug administered orally	0.017***	0.012*	-0.184***	-0.192***
	(0.005)	(0.005)	(0.045)	(0.045)
Drug administered intravenously	0.061***	0.099***	0.230***	0.389***
	(0.007)	(0.010)	(0.034)	(0.066)
Firm with <10 products on the market	-0.014**	-0.023***	0.012	-0.070***
	(0.004)	(0.005)	(0.013)	(0.012)
Firm with >50 products on the market	0.003	-0.016***	-0.028**	-0.036***
	(0.002)	(0.004)	(0.010)	(0.009)
N	768285	768285	346700	310255
Instrumental variables		✓	✓	✓
FE	✓	✓	✓	✓

Notes: Standard errors in parenthesis. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$. Column (3) includes only drugs for which the UK price can be observed, and Column (4) includes only drugs for which the Polish price can be observed. Columns (2)-(4) use instrumental variables and all regressions include ATC5 and time fixed effects.

Next, we examine the effect of prices on the magnitude of shortages, conditional on a shortage being detected ($D_{jt} = 1$). Because observed sales rarely go to zero even during shortage periods, measuring this magnitude is vital for evaluating welfare effects. We use the log quantity missing, $\log(\hat{q}^d_{jt} - q^s_{jt})$, as our dependent variable and estimate the following equation, controlling for drug characteristics X:

$$\log(\hat{q}^d_{jt} - q^s_{jt}) = \beta_{FR} \log(p^{FR}_{jt}) + \beta_I \log(p^I_{jt}) + X'_{jt}\theta + \varepsilon_{jt} \tag{9}$$

for all j, t such that $D_{jt} = 1$.

Table 6: Shortage magnitude

	(1)	(2)	(3)	(4)
$\log(p^{FR}_{jt})$	-0.483***	-0.169*	-1.144***	-1.225***
	(0.005)	(0.078)	(0.026)	(0.071)
$\log(p^{UK}_{jt})$			0.203***	
			(0.015)	
$\log(p^{PL}_{jt})$				-0.705***
				(0.064)
genericized market	0.593***	0.556***	0.227***	0.980***
	(0.017)	(0.020)	(0.028)	(0.032)
Generic	0.000	0.089**	-0.077***	-0.042
	(0.016)	(0.027)	(0.023)	(0.032)
Drug administered orally	-0.180***	-0.107**	-0.786***	-1.964***
	(0.029)	(0.035)	(0.069)	(0.109)
Drug administered intravenously	-0.237***	-0.310***	-0.297**	-0.335**
	(0.039)	(0.044)	(0.094)	(0.130)
Firm with <10 products on the market	-0.576***	-0.520***	-0.185*	-1.220***
	(0.035)	(0.038)	(0.085)	(0.116)
Firm with >50 products on the market	0.090***	0.153***	0.221***	-0.123***
	(0.015)	(0.022)	(0.030)	(0.036)
N	199137	199137	98306	77596
Instrumental variables		✓	✓	✓
FE	✓	✓	✓	✓

Notes: Standard errors in parenthesis. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$. Column (3) includes only the drugs for which the UK price can be observed. Column (4) includes only the drugs for which the Polish price can be observed. All regressions include ATC3 and time fixed effects.

Table 6 shows how domestic and international prices affect the shortage magnitude in case of shortage in France. Column (1) and (2) display results for the full sample of shortage observations. Columns (3) and (4) restrict the sample to shortage observations matched with UK and Polish price data, respectively. We include time and drug class at ATC3 level fixed effects[6] and instrument for French and international prices in the last 3 columns, as in Table 5.

For drugs in shortage, a higher domestic price reduces the shortage magnitude. Conversely, a higher UK price increases the shortage magnitude in France, whereas the effect of the Polish price is negative. Our model in Section 2 explains the economic mechanisms driving these opposite effects. Note first that, in the model, the prices of countries that are excluded from supply in shock states affect neither the allocation of capacity nor the capacity choice, so the significant UK and Polish price coefficients indicate that both countries are themselves served, and plausibly rationed, during global supply shocks. For the UK, the joint pattern of a lower shortage probability (Table 5) and a larger shortage magnitude in France is consistent with the UK belonging to the rationed set with a priority above France: a higher UK price raises the capacity the firm installs, lowering the frequency with which global shortages reach France, but, conditional on a supply shock, the higher UK margin attracts the scarce capacity toward the UK and away from France. Being a high-price country per se would not deliver this pattern: if the UK were always fully served, the model predicts that a higher UK price would reduce UK demand and free capacity for France, decreasing rather than increasing the French shortage magnitude. For Poland, the negative effect on the French shortage magnitude corresponds to the opposite, elastic-demand case of the total cross-price effect derived in Section 2.1: when the demand of a rationed country is sufficiently elastic –plausible for Poland, where patients face copays of at least 30%– a higher price contracts its quantity demanded enough that, in shock states, capacity is freed for the other rationed countries, including France. Although a negative coefficient is also consistent with Poland being fully-served, Poland's lower relative prices place it below France in the priority ordering, meaning the effect most plausibly operates through

[6] As we condition on shortages taking place, adding fixed effects at a more disaggregated level does not allow us to obtain meaningful estimates.

Poland's highly elastic copay structure as a rationed country.

3.5 Counterfactuals

To evaluate how larger price mitigate shortages, we design counterfactual exercises that simulate price changes for specific drugs and measure the subsequent impact on shortages. For this, we first estimate a nested logit demand model and specify a stylized supply model. This framework accounts for substitution effects within ATC5 classes and captures how mitigating shortages for a given product spills over onto competing products.

Following a price increase, shortages can decrease for two reasons. First, a price increase can reduce the quantity demanded. Second, manufacturers can respond to the higher price by increasing the quantity supplied or production capacity, reducing shortage probabilities. The structural demand model also captures cross-product substitutions: if the price of a single drug rises, demand redirects toward close substitutes.

We conduct these counterfactual exercises using a sample of antipsychotic drugs, which are mainly used to treat schizophrenia and other conditions as bipolar disorder and Parkinson disease. Details on sample selection are in Appendix A.3. The sample includes five ATC4 classes and eight ATC5 classes, with 15,612 observations. We assume that the true quantity demanded is the predicted demand from Section 2.2: market shares are computed from these corrected quantities rather than from observed sales. By using the correct demand quantity of Section 2.2 we can safely include shortage periods such that our estimates are not distorted by the censored periods. This relies on the generated variables from the first stage predictions, and therefore the results in this section are sensitive to those assumptions. We then simulate how consumption patterns shift under alternative pricing scenarios. Since some products are likely to be highly substitutable within a class, we use a nested logit demand model where nests are defined by ATC5 class and the outside option.

3.5.1 Demand

We specify a discrete choice model to define the demand for pharmaceuticals. A market is defined at the monthly level. Patients, in consultation with their physicians, choose a drug j to treat their condition, by first selecting an ATC5 class g, which represents the nest, and then a product j within the class. The indirect utility of individual i consuming drug j at period t is given by:

$$u_{ijt} = \beta \log(p_{jt}^{FR}) + \theta x_{jt} + \phi_t + \xi_{jt} + \zeta_{igt} + (1 - \omega)\, \varepsilon_{ijt} \tag{10}$$

Utility depends on the product price and on characteristics x_j: the strength of the pill, if it is a generic drug, if it belongs to a genericized market, if the mode of administration is oral, and if it is an oral solution. We also include ATC4-level and month-year fixed effects.

Potential market size is calculated based on the number of active patients diagnosed with some psychotic condition by year, obtained from French National Health Insurance data (*Assurance Maladie*)[7]. We assume that each patient consumes 30 daily doses per month. Because official data are unavailable for 2024, we project the 2024 market size by applying a 6% growth rate to the 2023 baseline. This approach yields outside good market shares consistent with Kohn et al. (2004), who find that the proportion of untreated schizophrenia cases is around 30%.

The estimation details for the nested logit framework are provided in Appendix A.3. The final estimation equation takes the following linear form:

$$\log(\sigma_{jt}) - \log(\sigma_{0t}) = \beta\, \log(p_{jt}^{FR}) + \theta\, x_{jt} + \omega \log(\sigma_{j|g,t}) + \xi_{jt} \tag{11}$$

To estimate Equation (11) we need to instrument for both $\log(p_{jt}^{FR})$ and $\log(\sigma_{j|g,t})$, as they are likely correlated with the unobserved demand shocks contained in ξ_{jt}. Even under a regulated pricing regime, prices may respond to demand shocks, and a product's within-nest share

[7] `https://data.ameli.fr/explore/dataset/effectifs/table/?refine.patho_niv1=Maladies+psychiatriques&refine.patho_niv2=Troubles+psychotiques&sort=-prev.`

fluctuates directly with idiosyncratic demand shocks. Consequently, we use BLP instruments that capture the competitive environment of the market, these variables shift prices and within-nest shares without co-moving with unobserved demand shocks. Specifically, our instrument set includes the number of products offered by each firm in the market, the number of competitors within the ATC5 class, and the distance in dosage strength between a given product and its rivals within the same ATC5 group.

Table 7 reports the demand estimation results.[8] We find a negative price coefficient which translates to own-price elasticities in Figure 3. The nest parameter ω is close to 1, implying that products within the nest are highly substitutable, and that we are far from a standard logit specification. These results highlight two margins relevant to shortage dynamics: a price increase reduces a product's own demand, while this contraction simultaneously shifts demand toward rival products through cross-product substitution. Whether the total number of shortages falls depends on the balance between these two effects and the elasticity of supply with respect to higher margins.

3.5.2 Supply

While we estimate demand using the full sample of antipsychotic medications, including periods with observed shortages (which is possible because the dependent quantity is the imputed demand from Section 2.2 rather than observed sales), we estimate supply exclusively during shortage periods. To simulate counterfactual supply, we estimate a stylized supply relationship exclusively during shortage periods. Because observed quantities during detected shortages represent constrained supply rather than unconstrained demand, we regress log observed sales, $\log(q_{jt})$, on the log French price, $\log(p_{jt}^{FR})$, alongside controls for generic status, and fixed effects for time, ATC5 class, and mode of administration.

Price regulation in France prevents the price to be adjusted to immediate supply shocks, limiting the endogeneity problem of prices. In addition, the market under analysis contains only

[8] We have explored other specifications with other sets of instruments, including supply disruptions and re-centered instruments.

antipsychotic drugs and we can assume allocation priorities and the penalty cost of shortages is similar across classes. Nevertheless, unobserved product-level factors or selection into the shortage sample could still be a concern for quantities sold under shortage. To address this endogeneity of prices, we instrument for the French price using USD-Euro and Pound-Euro exchange-rate interactions with firm type. Conditional on time fixed effects (which absorb common macroeconomic shocks) and ATC5 fixed effects (which control for persistent molecule-level unobservables), we assume that differential exposure to exchange-rate fluctuations shifts pricing incentives without directly determining French physical allocations, restricting its effect on supplied quantities to operate through the price channel.

Table 8 reports the supply estimates. The coefficient on $\log(p_{jt}^{FR})$ is positive, confirming that higher prices are associated with higher supplied quantities during shortage periods. This is the supply response that we use in the counterfactual exercises. Generic products and those in genericized markets are likewise associated with larger quantities supplied during shortages. In our baseline scenario, we define a shortage magnitude as the gap between quantity demanded and quantity supplied for each drug-month observation. We then compare these baseline shortages with those predicted under alternative counterfactual pricing scenarios.

Figure 3: Own-price elasticities

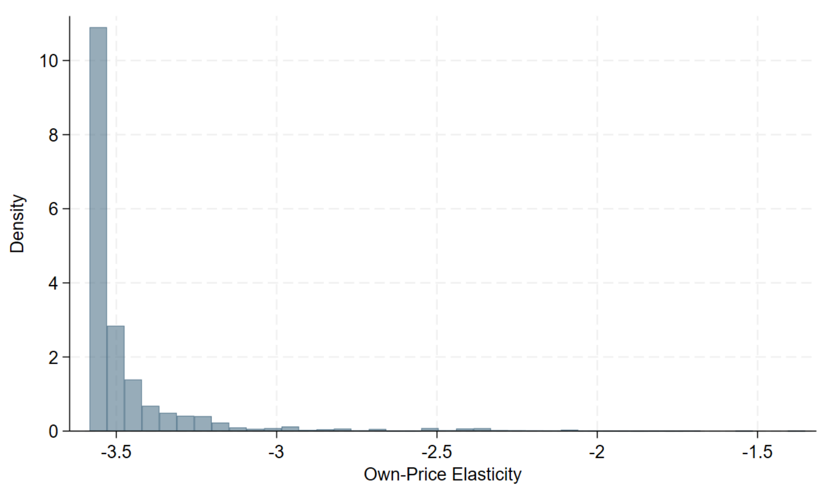


Table 7: Demand

	First stage		Demand
	$\log(p_{jt}^{FR})$	$\log(\sigma_{j\|g,t})$	$\log(\sigma_{jt}) - \log(\sigma_{0t})$
$\log(p_{jt}^{FR})$			-0.484***
			(0.015)
$\log(\sigma_{j\|g,t})$			0.865***
			(0.004)
Generic product	-0.376***	-1.380***	-0.192***
	(0.012)	(0.053)	(0.014)
Genericized market	-0.015	2.629***	0.188***
	(0.014)	(0.061)	(0.013)
Strength	0.001***	0.002***	-0.000
	(0.000)	(0.000)	(0.000)
Oral solution	0.549***	0.782***	0.320***
	(0.020)	(0.076)	(0.019)
Oral mode of administration	-1.136***	-2.176***	-0.597***
	(0.029)	(0.094)	(0.032)
Firm size	-0.000***	0.003***	
	(0.000)	(0.000)	
Mean strength rivals	0.004***	-0.003***	
	(0.000)	(0.000)	
Number of firms in ATC5	-0.020***	-0.057***	
	(0.001)	(0.005)	
N	15612	15612	15612
FE	✓	✓	✓

Notes: Robust standard errors in parenthesis. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$. Firm size measured in the number of products in the market. All regressions include month-year and ATC4 fixed effects.

3.5.3 Results

How do demand and supply change when prices increase, and how does this affect shortages? To answer this question, we use the estimated demand and supply models to study how shortages change under different price increases. For each scenario, we recompute equilibrium demand and supply under the modified price vector and compare the resulting shortages against the baseline. On the demand side, the model captures both the direct contractionary effect of the price increase on the shocked product and the subsequent reallocation of demand toward substitutes within the same therapeutic nest. On the supply side, we allow quantities to respond to the price increase

Table 8: Supply

	$\log(q_{jt})$
$\log(p_{jt}^{FR})$	2.813***
	(0.800)
Generic product	1.595**
	(0.500)
Genericized market	2.165***
	(0.258)
N	3983
FE	✓

Notes: Supply model is regression of log observed quantities on log of price in periods of shortage. Price is instrumented with USD-Euro exchange rate interacted with number of firms and generic dummy, exchange rate UK Pound-Euro interacted with a dummy for big firms (more than 50 products on the market). Fixed effects: ATC5, time, pharmaceutical mode of administration. Kleibergen-Paap F-Stat for Weak Instruments is 14.13

only for the shocked products, using the estimated supply equation. For the other products, we impose a simple capacity rule: additional demand generated by substitution away from the shocked product is fully absorbed by products that were not in shortage in the baseline, but it is not absorbed by products that were already under shortage.

Figure 4 shows the effect of increasing prices for all drugs on the number of shortages. The outcome is the percentage change in the number of shortages, computed by ATC5 and period and then averaged over time. The response is large: eliminating 50% of shortages requires an 8% increase in the price of all drugs. Figure 5 shows the same counterfactual separately by ATC5 class, illustrating significant heterogeneity across molecules. We then consider a more targeted counterfactual, where the price increase is applied only to the product with the largest number of shortages in the baseline within each ATC5. In this case, the direct effect should reduce shortages for the shocked product, but the aggregate effect also depends on the baseline shortages of other products that do not receive a price increase. This is because part of the demand moves toward those products, and some of them may already be constrained. On average across time and ATC5 classes, a 50% price increase for the products with the largest baseline shortages reduces the number of shortages by 80%, as shown in Figure 8 in Appendix

A.3.

Figure 4: Decrease in shortages due to a price increase for all drugs

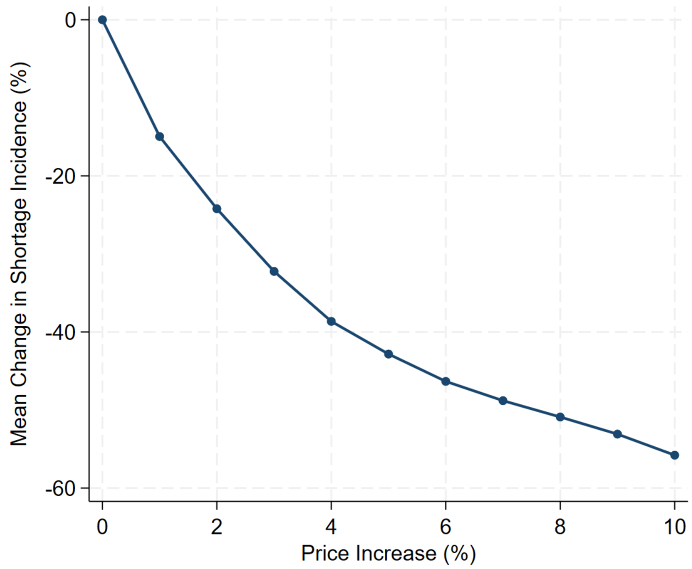
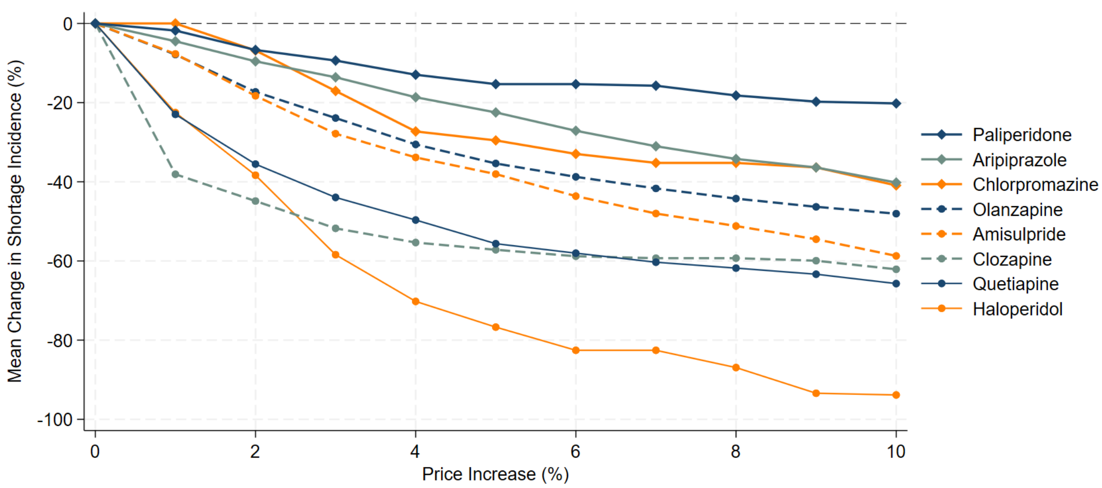


Figure 5: Decrease in number of shortages due to a price increase for all drugs for each ATC5



Notes: Percentage change in number of shortages when price of all drugs increases compared to baseline. The percentage change is computed by ATC5 and time period, then averaged over time.

To understand the mechanisms behind these results, we focus on one product and report the effects on shortage magnitude and shortage probability, together with the effects on rival products. The price increase is constant across all periods in the data. The results correspond to a 10% price increase for one product in the Amisulpride class, Amisulpride Teva 200mg.

Table 9: Demand: Counterfactual Effect of a 10% Price Increase on Amisulpride Teva 200mg on Shortage Magnitude and Probability

	Δ shortage magnitude			Δ shortage
	Absolute (DDD)	% of shortage	% of demand	probability
Shocked product	-2,056.4	-96.39	-8.19	-92.86
Same ATC5 rivals	984.6	3.65	0.09	0.00
Other ATC5s	143.2	0.03	0.00	0.00

Notes: All statistics are means across periods. Columns: change in shortage magnitude in absolute DDD units; Change in shortage magnitude as percentage of shortage: shortage change as % of baseline shortage q; shortage periods only; Change in shortage magnitude as percentage of demand: shortage change as % of baseline demand quantity of shocked product; Change in shortage probability: change in number of shortages as % of periods.

Table 9 shows that this price increase reduces the shortage magnitude of the shocked product by 96%. This large effect comes from both a reduction in demand and an increase in supply for the shocked product. However, the same price increase also raises demand for rival products. For products in the same ATC5, shortage magnitude increases by 984.6 DDD on average, although this represents only 3.65% of their baseline shortages and does not change their shortage probability. The effect on products in other ATC5 classes is much smaller.

Figure 6 shows the time series for the shocked product, separating demand and sales in the baseline from their counterfactual values after the price increase. The figure shows that the counterfactual reduces the gap between demand and sales in most shortage periods. Figure 7 shows the corresponding change in shortage magnitude for rival products. Under the model-based supply assumption, rivals that are already in shortage do not expand supply to meet the additional demand generated by substitution. As a result, some of the shortage reduction for the shocked product is partly offset by higher shortage pressure on rival products.

Table 10 reports the effect of the same 10% price increase on expenditures. For the shocked product, expenditures decrease by 14.2%, even though its price increases. This happens because the reduction in quantity is large enough to offset the higher price. For rival products, expenditures increase, but the effect is small under the model-based supply assumption: 0.48% for rivals in the same ATC5 and almost zero for products in other ATC5 classes. Under the full absorption assumption, rival products can serve all the additional demand generated by substitution, so their expenditures increase slightly more and there is no new quantity left unserved among rivals. These results show that the price increase reduces shortages for the shocked product without generating a large increase in total expenditures for other products.

Table 10: Counterfactual Effect of a 10% Price Increase of Amisulpride Teva 200mg on Expenditures

		Model-Based Supply		Full Absorption	
	Expenditure (baseline)	Δ Exp. (%)	q not served	Δ Exp. (%)	q not served
Shocked product	22,297.5	-14.183	261.3	-14.183	261.3
Same ATC5 rivals	739,794.4	0.484	984.6	0.596	0.0
Other ATC5s	20,183,949.0	0.003	143.2	0.004	0.0

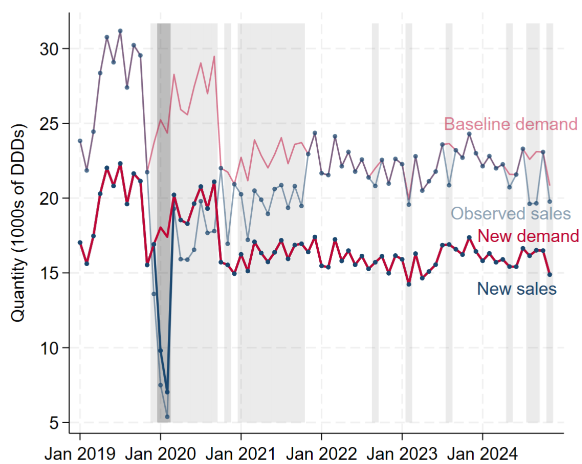
Notes: All statistics are means across periods. Expenditure = price × equilibrium q. Δ Expenditure = (counterfactual − baseline) / baseline × 100. Shocked product uses new price (×1.10). **Model-based:** rivals use model-implied q^s. **Full absorption:** rivals absorb only extra quantities from counterfactual, that come from shocked product. Rivals' q not served = 0 by construction (only for new extra demand). q not served captures the counterfactual shortage magnitude; for rivals already in shortage under the model, it captures new demand not absorbed by supply.

We then repeat this targeted counterfactual for all ATC5 classes, increasing prices for the products with the largest number of shortages in the baseline within each class. Table 11 reports the effects separately for the shocked products and for their rivals in the same ATC5, only for the model-based supply.

The reduction in shortages is large for most shocked products, but there is important heterogeneity across classes. In Haloperidol and Aripiprazole, shortages for the shocked products are eliminated. Quetiapine and Amisulpride also show very large reductions, above 94%. In contrast, the reduction is smaller for Paliperidone, where shortage magnitude falls by 41.5% and shortage probability falls by only 11%.

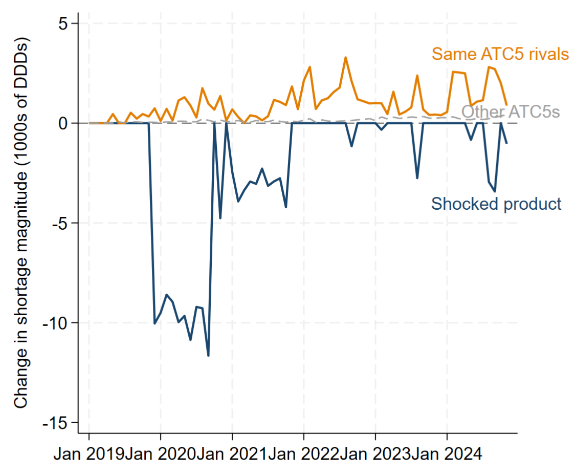
The table also shows that the gains for the shocked products can come with higher shortage pressure on rivals in the same ATC5. This is especially visible for Haloperidol and Chlorpromazine, where rival shortage magnitude increases by 179.9% and 54.5%, respectively. For other classes, such as Quetiapine, Aripiprazole, and Clozapine, the spillovers to rivals are much smaller. Overall, the results suggest that targeting products with the largest baseline shortages can substantially reduce shortages, but the effectiveness of the policy depends on how concentrated shortages are in the shocked products and on whether close substitutes are already constrained.

Figure 6: Demand and Sales



Notes: Amisulpride Teva 200mg. New demand refers to the counterfactual demand under a 10% price increase. New sales plots the minimum between the counterfactual demand and the counterfactual supply. Light shade region mark shortage periods in the baseline, dark shaded region marks shortage under the counterfactual.

Figure 7: Change in shortage magnitude



Notes: Shocked product is Amisulpride Teva 200mg, Same ATC5 rivals refers to other products in the class.

The magnitude of these counterfactual effects may appear large when compared with the reduced-form estimates of Section 3.4, where a 10% increase in the French price reduces the unmet quantity by around 2% on average (and by 11–12% on the subsamples matched to foreign prices, Table 6). Several differences between the two exercises explain this gap. First, the estimands differ: the reduced-form coefficient measures the average effect of prices on the intensity of shortages, conditional on a shortage occurring, across all drugs in our data, whereas the counterfactual computes the total equilibrium effect of a permanent, targeted price increase, including the extensive margin by which shortage episodes are eliminated entirely once supply reaches demand. Second, the shortage gap happens to be small, but it is actually the difference between two much larger quantities: for the shocked product in Table 9, the baseline shortage represents about 8% of demand, so that a moderate contraction of demand combined with a moderate expansion of supply suffices to close almost all of the gap –the 96% reduction in the shortage corresponds to a change of only 8.2% of baseline demand. Third, the antipsychotics market is particularly substitution-rich: the nest parameter close to one implies product-level demand elasticities that are much larger than the average price responsiveness of drug-level demand in the rest of our analysis, and the supply elasticity of 2.8 is likewise estimated on this market only. For these reasons, the counterfactual results should be interpreted as specific to markets with close therapeutic substitutes rather than extrapolated directly to the average drug in our sample.

Table 11: Counterfactual: 10% Price Increase on Highest-Shortage Products per ATC5

	Shocked products			Same ATC5 rivals	
ATC5	Δ short q (%)	Shortage prob. (%)	Δ Exp. (%)	Δ short q (%)	Δ Exp. (%)
Chlorpromazine	-80.47	-68.42	-4.078	54.52	10.264
Haloperidol	-100.00	-100.00	-7.633	179.90	8.715
Clozapine	-69.30	-60.53	-12.436	2.25	0.550
Olanzapine	-87.14	-88.24	-11.622	2.95	0.556
Quetiapine	-98.66	-97.37	-12.416	0.23	0.035
Amisulpride	-94.29	-96.43	-14.337	3.51	0.881
Aripiprazole	-100.00	-100.00	-20.555	0.63	0.423
Paliperidone	-41.51	-11.00	-0.227	27.67	16.434

Notes: Many products can be shocked at the same time in the same ATC5. Δ short q is as % of shortage: the change in shortage magnitude as a percentage of baseline shortage magnitude, computed as a ratio of totals over baseline-shortage periods only. Shortage probability is the relative change in the share of periods in shortage. Δ Exp. (%) is $(CF - baseline)/baseline$, aggregated over all periods. Rivals are products in the same ATC5 excluding the shocked product(s). If rivals have no baseline shortage, Δ short q is set to 0.

4 Conclusion

Shortages affect most major pharmaceutical markets. Their high prevalence among critical therapeutic areas for vulnerable patient populations, such as emergency medicine and oncology, has fostered a sizable medical and policy literature examining their incidence and clinical implications. While the structural peculiarities of the pharmaceutical market have been posited as underlying causes of these supply deficits, formal empirical analyses remain scarce and have rarely extended beyond the United States.

In the case of France, we demonstrate that shortages are directly dependent on pricing structures and escalate as prices fall. This adds a crucial short-term trade-off between fiscal expenditures and product availability to the standard dynamic, long-term efficiency trade-offs between current cost containment and future innovation. The conventional justification for elevated pharmaceutical pricing is that industry rents generate greater incentives for innovation (Acemoglu and Linn, 2004; Dubois, de Mouzon, Scott-Morton, and Seabright, Dubois et al.); however, this argument falls short for generic markets or small nations whose contribution to aggregate global revenue is marginal. Nonetheless, our findings show that because shortage risks escalate at lower price points, a short-run trade-off between expenditure and market availability persists even for non-innovative products, where long-run innovation rents are irrelevant, and applies a fortiori to innovative therapies.

Parallel trade within Europe provides another prominent example of international pricing spillovers linked to shortages. Prior work indicates that cross-border arbitrage alters price negotiations and industry returns (Kyle, 2011; Dubois and Sæthre, 2020). Crucially, cross-country price spillovers may arise not only from strategic manufacturer behavior, such as prioritizing higher-priced markets during global supply shocks, but also via the arbitrage activities of parallel traders, operating independently of manufacturer intent.

Empirically, we demonstrate that UK prices reduce the probability of drug shortages in France but increase their intensity. One interpretation is that higher regional prices bolster global industry profitability and capacity investment, thereby reducing the baseline frequency of shortage events. However, conditional on a shortage occurring, drugs with higher relative prices in the UK see significantly lower quantities supplied to France. This is consistent with a supply-diversion mechanism: when global demand outstrips available capacity, firms prioritize higher-margin jurisdictions. Through the lens of our model, this joint pattern indicates that the UK itself faces rationing during global supply shocks, with a priority above France in the firm's allocation of scarce capacity. This nuanced analysis is made possible by our structural identification strategy, which infers shortage events using assumptions grounded in institutional features of the pharmaceutical market.

While current policy frameworks center on increasing transparency and mitigating downstream impacts, structural proposals targeting root causes include implementing price floors for generic chemotherapy drugs, accelerating regulatory reviews for generic entries (Kantarjian, 2014), and increasing the frequency of quality audits to enforce compliance (Scott Morton and Boller, 2017).

Our findings suggest that while higher domestic prices could mitigate the occurrence of shortages and satisfy a larger share of demand, achieving zero shortages may be economically prohibitive. Consequently, if regulatory authorities aim to minimize overall clinical harm under budget constraints, our results suggest a natural prioritization: they might concentrate price increases on drugs with few or no therapeutic alternatives, where shortages are most critical. Conversely, they could tolerate supply deficits for products with readily available substitutes, where patient demand can be absorbed by rivals. Finally, the cooperative international price effect we identify on shortage frequencies calls for explicit coordination among European regulators to internalize the positive externalities of price-setting on capacity investment.

Bibliography

Acemoglu, D. and J. Linn (2004). Market Size In Innovation: Theory And Evidence From The Pharmaceutical Industry. *The Quarterly Journal of Economics 119*(3), 1049–90.

Acosta, A., E. P. Vanegas, J. Rovira, B. Godman, and T. Bochenek (2019). Medicine Shortages: Gaps Between Countries and Global Perspectives. *Frontiers in Pharmacology 10*(July), 1–21.

Asker, J., A. Collard-Wexler, and J. De Loecker (2019). (Mis)Allocation, market power, and global oil extraction. *American Economic Review 109*(4), 1568–1615.

Baudet, G. and C. Dherbécourt (2025). Tensions et ruptures de stock de médicaments déclarées par les industriels : quelle ampleur, quelles conséquences sur les ventes aux officines ? Études et Résultats 1335, DREES.

Becker, D. J., S. Talwar, B. P. Levy, M. Thorn, J. Roitman, R. H. Blum, L. B. Harrison, and M. L. Grossbard (2013). Impact of oncology drug shortages on patient therapy: Unplanned treatment changes. *Journal of Oncology Practice 9*(4).

Benhabib, A., S. Ioughlissen, C. Ratignier-Carbonneil, and P. Maison (2020). The French reporting system for drug shortages: Description and trends from 2012 to 2018: An observational retrospective study. *BMJ Open 10*(3), 1–7.

Blankart, K. E. and S. Felder (2022). Do medicine shortages reduce access and increase pharmaceutical expenditure? a retrospective analysis of switzerland 2015-2020. *Value in Health 25*(7), 1124–1132.

Conti, R. M. and M. E. Wosińska (2025). The economics of generic drug shortages: The limits of competition. *Journal of Economic Perspectives 39*(2), 79–102.

Dave, C. V., A. Pawar, E. R. Fox, G. Brill, and A. S. Kesselheim (2018). Predictors of Drug Shortages and Association with Generic Drug Prices: A Retrospective Cohort Study. *Value in Health 21*(11), 1286–1290.

Dubois, P., O. de Mouzon, F. Scott-Morton, and P. Seabright. *RAND Journal of Economics 46(4)*(October).

Dubois, P. and M. Sæthre (2020). On the Effect of Parallel Trade on Manufacturers' and Retailers' Profits in the Pharmaceutical Sector. *Econometrica 88*(6), 2503–2545.

Fittler, A., R. G. Vida, V. Rádics, and L. Botz (2018). A challenge for healthcare but just another opportunity for illegitimate online sellers: Dubious market of shortage oncology drugs. *PLoS ONE 13*(8), 1–17.

Galdin, A. (2025). Resilience of global supply chains and generic drug shortages. *Dartmouth College, Tuck Business School Working Paper*.

Haninger, K., A. Jessup, and K. Koehler (2011). Economics Analysis of the Causes of Drug Shortages. *ASPE Issue Brief*.

Kaakeh, R., B. V. Sweet, C. Reilly, C. Bush, S. DeLoach, B. Higgins, A. M. Clark, and J. Stevenson (2011). Impact of drug shortages on U.S. health systems. *American Journal of Health-System Pharmacy 68*(19), 1811–1819.

Kantarjian, H. M. (2014). Chemotherapy drug shortages in the United States revisited. *Journal of Oncology Practice 10*(5), 329–331.

Kim, S.-H. and F. Scott-Morton (2015). A Model of Generic Drug Shortages: Supply Disruptions, Demand Substitution and Price Control. (FDA 2011).

Kohn, R., S. Saxena, I. Levav, and B. Saraceno (2004). The treatment gap in mental health care. *Bulletin of the World Health Organization 82*(11), 858–866.

Kortelainen, M., J. Markkanen, M. Siikanen, and O. Toivanen (2023). The Effects of Price Regulation on Pharmaceutical Expenditure and Availability. *Working paper*.

Kyle, M. (2011). Strategic responses to parallel trade. *The B.E. Journal of Economic Analysis and Policy 11*(2).

Lee, J., H. S. Lee, H. Shin, and V. Krishnan (2021). Alleviating Drug Shortages: The Role of Mandated Reporting-Induced Operational Transparency. *Management Science 67*(4), 2326–2339.

Liebman, E., E. C. Lawler, A. Dunn, and D. B. Ridley (2023). Consequences of a shortage and rationing: Evidence from a pediatric vaccine. *Journal of Health Economics 92*, 102819.

Parsons, H. M., S. Schmidt, A. B. Karnad, Y. Liang, M. J. Pugh, and E. R. Fox (2016). Association between the number of suppliers for critical antineoplastics and drug shortages: Implications for future drug shortages and treatment. *Journal of Oncology Practice 12*(3), e289–e298.

Pauly, M. V. (2005). Improving vaccine supply and development: Who needs what? *Health Affairs 24*(3), 680–689.

Pauwels, K., I. Huys, M. Casteels, and S. Simoens (2014). Drug shortages in European countries: A trade-off between market attractiveness and cost containment? *BMC Health Services Research 14*(1).

Ridley, D. B., X. Bei, and E. B. Liebman (2016). No shot: US Vaccine Prices And Shortages. *Health Affairs 35*(2), 235–241.

Scott Morton, F. M. and L. T. Boller (2017). Enabling competition in pharmaceutical markets. *Hutchins Center Working Paper* (May).

Stomberg, C. (2016). *Drug Shortages, Pricing, and Regulatory Activity*, pp. 323–348. University of Chicago Press.

TLV (2023). International price comparison 2023: An analysis of swedish pharmaceutical prices in relation to 19 other european countries. Technical Report 03523/2023, Dental and Pharmaceutical Benefits Agency (TLV), Stockholm, Sweden. Director General: Agneta Karlsson. Working group: Christoffer Karlsson, Martin Löwing Jensen, Jonas Nilsson, and Marie Orre.

Woodcock, J. and M. Wosinska (2013). Economic and technological drivers of generic sterile injectable drug shortages. *Clinical Pharmacology and Therapeutics 93*(2), 170–176.

Yang, Y. H. (2020). National Drug Shortage Impacts on Medi-Cal. *Journal of Supply Chain and Operations Management 18*(1), 55–68.

Yurukoglu, A., E. Liebman, and D. B. Ridley (2017). The role of government reimbursement in drug shortages. *American Economic Journal: Economic Policy 9*(2), 348–382.

A Appendix

A.1 Firms Supply and Capacity Choice with Risk of Shortage

This appendix provides the proofs of Propositions 1, 2 and 3 of Section 2.1, as well as the analysis of the case without shortage penalty ($\gamma \equiv 0$) discussed at the end of that section. Throughout, we use the notation of the main text, Assumptions 1 and 2, and the definitions of the thresholds m^* and n^* given before Proposition 1. In particular, countries are indexed in decreasing order of price as in equation (1) and, unless stated otherwise, all derivations refer to the shock state, in which the available capacity is $\overline{K}_j - \kappa_j$.

Proof of Proposition 1. In the shock state, the firm allocates the available capacity by solving the inner maximization of the capacity choice problem of Section 2.1:

$$\max_{\left\{0 \le q^s_{jn} \le q^d_{jn}(p^n_j)\right\}_{n=1,\dots,N}, \sum_{c=1}^N q^s_{jc} \le \overline{K}_j - \kappa_j} \left\{ \sum\nolimits_{c=1}^N \left((p^c_j - c_j) q^s_{jc} \ - \gamma \left(q^d_{jc}\left(p^c_j\right) - q^s_{jc} \right) \right) \right\} \tag{12}$$

To find the solution, remark that the marginal cost of not supplying any unit demanded in country c, where the demand is $q^d_{jc}\left(p^c_j\right)$, is $\left(p^c_j - c_j\right) + \gamma'\left(q^d_{jc}\left(p^c_j\right)\right)$ and the marginal cost of not supplying the last unit demanded in country c is $\left(p^c_j - c_j\right)$ (since $\gamma'(0) = 0$).

This implies the following:

If $\left(p^c_j - c_j\right) > \left(p^{c'}_j - c_j\right) + \gamma'\left(q^d_{jc'}\left(p^{c'}_j\right)\right)$ then it cannot be that we have any shortage in c even if there is full shortage in c'.

If $\left(p^c_j - c_j\right) \le \left(p^{c'}_j - c_j\right) + \gamma'\left(q^d_{jc'}\left(p^{c'}_j\right)\right)$ then it cannot be that we have complete shortage in c' and no shortage at all in c.

Given the price ordering (1), we have $\left(p^c_j - c_j\right) > \left(p^{c+1}_j - c_j\right)$ for any c but it can be that $\left(p^c_j - c_j\right) < \left(p^{c+n}_j - c_j\right) + \gamma'\left(q^d_{jc+n}\left(p^{c+n}_j\right)\right)$ in which case there will be some shortage in $c, c+1, c+n$ countries and full shortage in country $c+n+1$ where n is such that $\left(p^c_j - c_j\right) > \left(p^{c+n+1}_j - c_j\right) + \gamma'\left(q^d_{jc+n+1}\left(p^{c+n+1}_j\right)\right)$. Assumption 1 guarantees that the resulting partition of countries into supply regimes is well defined, with each regime consisting of consecutive countries.

With m^* and n^* defined as in the main text, it follows that the highest-price countries are fully served, $q^s_{jc} = q^d_{jc}\left(p^c_j\right)$ for $c = 1, \dots, m^* - 1$, and that the supplies $q^s_{jm^*}, \dots, q^s_{jm^*+n^*-1}$ of the rationed countries equalize marginal payoffs:

$$p^{m^*+n-1}_j - c_j + \gamma'\left(q^d_{jm^*+n-1}\left(p^{m^*+n-1}_j\right) - q^s_{jm^*+n-1}\right) \equiv \rho \text{ for } n = 1, \dots, n^* \tag{13}$$

where ρ is the Lagrange multiplier of the capacity constraint in (12), while the remaining countries are excluded:

$$q^s_{jm^*+n-1} = 0 \text{ for } n \geq n^* + 1 \tag{14}$$

As (13) implies that for $c = m^*, \ldots, m^* + n^* - 1$

$$q^s_{jc} = q^d_{jc}\left(p^c_j\right) - \gamma'^{-1}\left(\rho - \left(p^c_j - c_j\right)\right) \tag{15}$$

the binding capacity constraint implies

$$\begin{aligned}
\overline{K}_j - \kappa_j &= \sum\nolimits_{c=1}^{m^*-1} q^d_{jc}\left(p^c_j\right) + \sum\nolimits_{c=m^*}^{m^*+n^*-1} q^s_{jc} \\
&= \sum\nolimits_{c=1}^{m^*+n^*-1} q^d_{jc}\left(p^c_j\right) - \sum\nolimits_{c=m^*}^{m^*+n^*-1} \gamma'^{-1}\left(\rho - \left(p^c_j - c_j\right)\right)
\end{aligned}$$

thus ρ solves

$$\sum\nolimits_{c=m^*}^{m^*+n^*-1} \gamma'^{-1}\left(\rho - \left(p^c_j - c_j\right)\right) = \sum\nolimits_{c=1}^{m^*+n^*-1} q^d_{jc}\left(p^c_j\right) - \overline{K}_j + \kappa_j$$

as stated in the proposition. $\square$

Proof of Proposition 2. Under Assumption 2, $\gamma(x) = \gamma\frac{x^2}{2}$, implying $\gamma'(x) = \gamma x$ and $\gamma'^{-1}(x) = \frac{x}{\gamma}$, so that the equation defining ρ in Proposition 1 becomes

$$\frac{1}{\gamma}\sum\nolimits_{c=m^*}^{m^*+n^*-1}\left(\rho - \left(p^c_j - c_j\right)\right) = \sum\nolimits_{c=1}^{m^*+n^*-1} q^d_{jc}\left(p^c_j\right) - \overline{K}_j + \kappa_j$$

hence

$$\rho = \frac{\gamma}{n^*}\sum\nolimits_{c=1}^{m^*+n^*-1} q^d_{jc}\left(p^c_j\right) + \frac{1}{n^*}\sum\nolimits_{c=c^*+1}^{m^*+n^*-1}\left(p^c_j - c_j\right) - \frac{\gamma}{n^*}\left(\overline{K}_j - \kappa_j\right) \tag{16}$$

Substituting into (15), the supply to a rationed country $c \in \{m^*, \ldots, m^* + n^* - 1\}$ is

$$\begin{aligned}
q^s_{jc} &= q^d_{jc}\left(p^c_j\right) - \frac{\rho - \left(p^c_j - c_j\right)}{\gamma} \\
&= q^d_{jc}\left(p^c_j\right) - \frac{1}{n^*}\sum\nolimits_{c=1}^{m^*+n^*-1} q^d_{jc}\left(p^c_j\right) + \frac{1}{\gamma}\left[\left(p^c_j - c_j\right) - \frac{1}{n^*}\sum\nolimits_{c=m^*}^{m^*+n^*-1}\left(p^c_j - c_j\right)\right] + \frac{1}{n^*}\left(\overline{K}_j - \kappa_j\right)
\end{aligned}$$

which is the expression in the proposition. The comparative statics, holding m^*, n^* and $\overline{K}_j$ fixed, follow by differentiating this expression. For the own price p^c_j, the terms in $q^d_{jc}\left(p^c_j\right)$ and

$\left(p_j^c - c_j\right)$ each enter with coefficient $\left(1 - \frac{1}{n^*}\right)$, giving

$$\left.\frac{\partial q_{jc}^s}{\partial p_j^c}\right|_{m^*,n^*,\overline{K}_j} = \left(\frac{\partial q_{jc}^d\left(p_j^c\right)}{\partial p_j^c} + \frac{1}{\gamma}\right)\left(1 - \frac{1}{n^*}\right)$$

which is positive if $\frac{\partial q_{jc}^d\left(p_j^c\right)}{\partial p_j^c} > -\frac{1}{\gamma}$. For the price of another country $c' \neq c$: if $c' \in \{1, \ldots, m^* - 1\}$, only the term $-\frac{1}{n^*} q_{jc'}^d\left(p_j^{c'}\right)$ depends on $p_j^{c'}$, giving $-\frac{1}{n^*}\frac{\partial q_{jc'}^d\left(p_j^{c'}\right)}{\partial p_j^{c'}} > 0$; if $c' \in \{m^*, \ldots, m^* + n^* - 1\}$, the derivative is $-\frac{1}{n^*}\left(\frac{\partial q_{jc'}^d\left(p_j^{c'}\right)}{\partial p_j^{c'}} + \frac{1}{\gamma}\right)$, which is negative if $\frac{\partial q_{jc'}^d\left(p_j^{c'}\right)}{\partial p_j^{c'}} > -\frac{1}{\gamma}$; and if $c' \in \{m^* + n^*, \ldots, N\}$, no term depends on $p_j^{c'}$, so the derivative is zero. Finally, differentiating with respect to capacity gives

$$\left.\frac{\partial q_{jc}^s}{\partial \overline{K}_j}\right|_{m^*,n^*} = \frac{1}{n^*}$$

□

Proof of Proposition 3. Using the threshold allocation of Proposition 1, and since both the no-shock state payoff and the full-shortage penalties for excluded countries do not depend on $\overline{K}_j$ within a fixed regime (m^*, n^*) the capacity choice objective can be written, up to additive constants, as

$$\begin{aligned}
& \max_{\overline{K}_j}\left\{\pi\left\{\sum\nolimits_{c=1}^{N}\left((p_j^c - c_j)q_{jc}^s \; - \gamma\left(q_{jc}^d\left(p_j^c\right) - q_{jc}^s\right)\right)\right\} - C\left(\overline{K}_j\right)\right\} \\
= \; & \max_{\overline{K}_j}\left\{\pi\left\{\sum\nolimits_{c=1}^{m^*-1}\left(p_j^c - c_j\right)q_{jc}^d\left(p_j^c\right) \; + \sum\nolimits_{c=m^*}^{m^*+n^*-1}\left((p_j^c - c_j)q_{jc}^s - \gamma\left(q_{jc}^d\left(p_j^c\right) - q_{jc}^s\right)\right)\right\} - C\left(\overline{K}_j\right)\right\}
\end{aligned}$$

where the supplies q_{jc}^s of the rationed countries depend on $\overline{K}_j$ as derived in Proposition 2. The first-order condition with respect to $\overline{K}_j$, using $\left.\frac{\partial q_{jc}^s}{\partial \overline{K}_j}\right|_{m^*,n^*} = \frac{1}{n^*}$, is

$$\begin{aligned}
\pi\left\{\sum\nolimits_{c=m^*}^{m^*+n^*-1}\left[\left(p_j^c - c_j\right) \; + \gamma'\left(q_{jc}^d\left(p_j^c\right) - q_{jc}^s\right)\right]\frac{\partial q_{jc}^s}{\partial \overline{K}_j}\right\} &= C'\left(\overline{K}_j\right) \\
\pi\left\{\sum\nolimits_{c=m^*}^{m^*+n^*-1}\left[\left(p_j^c - c_j\right) + \gamma'\left(q_{jc}^d\left(p_j^c\right) - q_{jc}^s\right)\right]\frac{1}{n^*}\right\} &= C'\left(\overline{K}_j\right) \\
\pi\rho &= C'\left(\overline{K}_j\right)
\end{aligned}$$

where the last equality uses (13). Hence $\overline{K}_j = C'^{-1}(\pi\rho)$, which combined with the expression

(16) for ρ under Assumption 2 gives the first expression of the proposition.

With $C(x) = \frac{x^2}{2}$, C'^{-1} is the identity, so that

$$\overline{K}_j \quad = \quad \pi\left(\frac{\gamma}{n^*}\sum\nolimits_{c=1}^{m^*+n^*-1} q_{jc}^d\left(p_j^c\right) + \frac{1}{n^*}\sum\nolimits_{c=m^*}^{m^*+n^*-1}\left(p_j^c - c_j\right) - \frac{\gamma}{n^*}\left(\overline{K}_j - \kappa_j\right)\right)$$

and solving for $\overline{K}_j$ yields

$$\overline{K}_j = \pi\frac{\gamma\sum_{c=1}^{m^*+n^*-1} q_{jc}^d\left(p_j^c\right) + \sum_{c=m^*}^{m^*+n^*-1}\left(p_j^c - c_j\right) + \gamma\kappa_j}{n^* + \pi\gamma}$$

Differentiating this closed form with respect to p_j^c, holding m^* and n^* fixed, gives the comparative statics of the proposition: for $c \in \{1, \ldots, m^* - 1\}$, only $q_{jc}^d\left(p_j^c\right)$ depends on p_j^c; for $c \in \{m^*, \ldots, m^* + n^* - 1\}$, both $q_{jc}^d\left(p_j^c\right)$ and the margin $\left(p_j^c - c_j\right)$ do; and for $c \in \{m^* + n^*, \ldots, N\}$, no term does. □

The case without shortage penalty ($\gamma \equiv 0$) We finally derive the results, stated at the end of Section 2.1, for the alternative case where there is no shortage penalty ($\gamma \equiv 0$).

Define the cumulative demand of the n highest-price countries as $Q_j^n\left(p_j^1, \ldots, p_j^n\right) \equiv \sum_{c=1}^{n} q_{jc}^d\left(p_j^c\right)$. Then, $Q_j^n\left(p_j^1, \ldots, p_j^n\right)$ is increasing in n while $C'^{-1}\left(\pi\left(p_j^n - c_j\right)\right)$ is decreasing in n because p_j^n is decreasing in n.

Thus, there exists m^* such that

$$Q_j^{m^*-1}\left(p_j^1, \ldots, p_j^{m^*-1}\right) + \kappa_j < C'^{-1}\left(\pi\left(p_j^{m^*} - c_j\right)\right) \text{ and } Q_j^{m^*}\left(p_j^1, \ldots, p_j^{m^*}\right) + \kappa_j \geq C'^{-1}\left(\pi\left(p_j^{m^*+1} - c_j\right)\right)$$

(with the convention $p_j^{N+1} = c_j$, so that m^* is well defined).

The optimal choice of capacity $\overline{K}_j$ must then be such that it is larger than the demand of all countries $\{1, \ldots, m^* - 1\}$ because the marginal cost of increasing capacity is less than the marginal benefit of supplying the m^* country, but less than supplying country m^*+1 because the marginal cost of increasing capacity above $Q_j^{m^*}$ is larger than the marginal benefit of supplying the $m^* + 1$ country in case of shortage.

Thus

$$Q_j^{m^*-1}\left(p_j^1, \ldots, p_j^{m^*-1}\right) \leq \overline{K}_j - \kappa_j \leq Q_j^{m^*}\left(p_j^1, \ldots, p_j^{m^*}\right)$$

so that, in the shock state, countries $1, \ldots, m^* - 1$ are fully served, country m^* is rationed (with zero shortage in the limiting case $\overline{K}_j - \kappa_j = Q_j^{m^*}$) and countries $m^* + 1, \ldots, N$ are excluded: in the notation of Propositions 1 and 2, $n^* = 1$ and the marginal country is m^*. The capacity choice problem of Section 2.1 is then equivalent to

$$\begin{aligned}
& \max_{\overline{K}_j}\left\{\pi \max_{\substack{\left\{0\le q^s_{jn}\le q^d_{jn}(p^n_j)\right\}_{n=m^*,\dots,N},\\ \sum_{c=1}^N q^s_{jc}\le \overline{K}_j-\kappa_j}} \left\{\sum_{c=1}^{m^*-1}(p^c_j-c_j)q^d_{jc}(p^c_j)+\sum_{c=m^*}^{N}(p^c_j-c_j)q^s_{jc}\right\}-C(\overline{K}_j)\right\}\\
\Leftrightarrow\quad & \max_{\overline{K}_j}\left\{\pi\left[(p^{m^*}_j-c_j)\left(\overline{K}_j-\kappa_j-\sum_{c=1}^{m^*-1}q^d_{jc}(p^c_j)\right)\right]-C(\overline{K}_j)\right\}
\end{aligned}$$

hence

$$\pi\left(p^{m^*}_j-c_j\right)=C'\left(\overline{K}_j\right)$$

and

$$\overline{K}_j=C'^{-1}\left(\pi\left(p^{m^*}_j-c_j\right)\right)$$

This first-order condition characterizes the optimum only when its solution is interior to the bracket above, that is, when

$$Q^{m^*-1}_j\left(p^1_j,\dots,p^{m^*-1}_j\right)+\kappa_j<C'^{-1}\left(\pi\left(p^{m^*}_j-c_j\right)\right)<Q^{m^*}_j\left(p^1_j,\dots,p^{m^*}_j\right)+\kappa_j$$

(the *rationing regime*, in which country m^* faces a strictly positive shortage). If instead $C'^{-1}\left(\pi\left(p^{m^*}_j-c_j\right)\right)\ge Q^{m^*}_j+\kappa_j$, the marginal benefit of capacity drops at $\overline{K}_j=Q^{m^*}_j+\kappa_j$ from $\pi\left(p^{m^*}_j-c_j\right)$, which exceeds the marginal cost, to $\pi\left(p^{m^*+1}_j-c_j\right)$, which does not, so that the optimum is at the kink:

$$\overline{K}_j=Q^{m^*}_j\left(p^1_j,\dots,p^{m^*}_j\right)+\kappa_j,\qquad \pi\left(p^{m^*+1}_j-c_j\right)\le C'\left(\overline{K}_j\right)\le\pi\left(p^{m^*}_j-c_j\right)$$

(the *exact-coverage regime*, in which capacity exactly covers the shock-state demand of countries $1,\dots,m^*$ and no country is rationed).

In the rationing regime, the supply to the marginal country is

$$\begin{aligned}
q^s_{jm^*} &= \overline{K}_j-\kappa_j-\sum\nolimits_{c=1}^{m^*-1}q^d_{jc}(p^c_j)\\
&= C'^{-1}\left(\pi\left(p^{m^*}_j-c_j\right)\right)-\sum\nolimits_{c=1}^{m^*-1}q^d_{jc}(p^c_j)-\kappa_j
\end{aligned}$$

$$\begin{aligned}
\left.\frac{\partial q^s_{jm^*}}{\partial p^{m^*}_j}\right|_{m^*} &= \frac{\partial C'^{-1}\left(\pi\left(p^{m^*}_j-c_j\right)\right)}{\partial p^{m^*}_j}>0\\
\left.\frac{\partial q^s_{jm^*}}{\partial p^c_j}\right|_{m^*} &= -\frac{\partial q^d_{jc}(p^c_j)}{\partial p^c_j}>0 \text{ if } c<m^*
\end{aligned}$$

so that the supply in the rationed country is increasing in the price in that country, as claimed in the main text.

Thus, in the rationing regime,

$$\frac{\partial \overline{K}_j}{\partial p_j^{m^*}} > 0 \qquad \text{and} \qquad \frac{\partial \overline{K}_j}{\partial p_j^c} = 0 \text{ for } c \neq m^* :$$

capacity is pinned down by the margin of the rationed country alone and increases with its price. In the exact-coverage regime, $\overline{K}_j = Q_j^{m^*} + \kappa_j$, so that

$$\frac{\partial \overline{K}_j}{\partial p_j^c} = \frac{\partial q_{jc}^d \left(p_j^c\right)}{\partial p_j^c} < 0 \text{ for } c \leq m^* :$$

capacity tracks the demand of the served countries and decreases with their prices, exactly as it does for the fully served countries when $\gamma > 0$ (Proposition 3).

Moreover, as $\dfrac{\partial Q_j^n \left(p_j^1, \ldots, p_j^n\right)}{\partial p_j^{n'}} \leq 0$ if $n' \leq n$, it must be that $\dfrac{\partial m^*}{\partial p_j^n} \geq 0$ if $n \leq m^*$: higher prices in served countries reduce their demand and allow the firm to reach further down the price ladder. Note that the two regimes must be distinguished when comparing capacities across values of m^*: for a fixed price vector, the interior solution $C'^{-1}\left(\pi\left(p_j^m - c_j\right)\right)$ is decreasing in m, and the optimum passes from one first-order-condition bracket to the next only through the exact-coverage regime. As the price of a country $n < m^*$ increases, the optimum moves continuously from the rationing regime, where capacity is unchanged, through the exact-coverage regime, where capacity falls with the demand of country n, to the rationing regime with marginal country $m^* + 1$, where capacity is constant at the lower level $C'^{-1}\left(\pi\left(p_j^{m^*+1} - c_j\right)\right)$. Hence, without a shortage penalty, the optimal capacity is weakly decreasing in the prices of the fully served countries and increasing only in the price of the rationed country.

The shortage quantity $q_{jm^*}^d(p_j^{m^*}) - q_{jm^*}^s$, in contrast, is weakly decreasing in every price p_j^n with $n \leq m^*$: in the rationing regime, an increase in p_j^c for $c < m^*$ lowers $Q_j^{m^*-1}$ and thus raises $q_{jm^*}^s$ at unchanged capacity, an increase in $p_j^{m^*}$ both raises capacity and lowers $q_{jm^*}^d$, and in the exact-coverage regime the shortage is zero. This establishes the properties of the optimal capacity stated at the end of Section 2.1.

A.2 Generalization to a Non-Quadratic Shortage Penalty

In this appendix we show that the quadratic form of Assumption 2 serves only to deliver closed-form expressions: all the qualitative results of Section 2.1 extend to a general strictly convex penalty. We assume throughout that γ is twice continuously differentiable with $\gamma\left(0\right) =$

$\gamma'(0) = 0$ and $\gamma''(x) > 0$ on the relevant range of shortages, and that the fixed cost of capacity satisfies $C'' > 0$. Note first that Proposition 1 and the optimality condition $\overline{K}_j = C'^{-1}(\pi\rho)$ of Proposition 3 are established in Appendix A.1 without using the quadratic form. Similarly, since the shortage of a rationed country is $x_c \equiv q^d_{jc}\left(p^c_j\right) - q^s_{jc} = \gamma'^{-1}\left(\rho - \left(p^c_j - c_j\right)\right)$ with γ'^{-1} increasing, shortages increase within the rationed set as the price falls, again without the quadratic form.

For rationed countries $c \in \{m^*, \ldots, m^* + n^* - 1\}$, define the local inverse curvature of the penalty at the equilibrium shortage, and the associated weights:

$$h_c \equiv \frac{1}{\gamma''(x_c)} > 0, \qquad H \equiv \sum\nolimits_{c'=m^*}^{m^*+n^*-1} h_{c'}, \qquad w_c \equiv \frac{h_c}{H} \in (0,1], \qquad \sum\nolimits_{c=m^*}^{m^*+n^*-1} w_c = 1.$$

Under Assumption 2, $\gamma'' \equiv \gamma$, so that $h_c = 1/\gamma$, $H = n^*/\gamma$ and $w_c = 1/n^*$: the weights reduce to the equal shares of Propositions 2 and 3.

Proposition 4 (Comparative statics with a general convex penalty)**.** *Holding m^*, n^* and $\overline{K}_j$ fixed, for a rationed country $c \in \{m^*, \ldots, m^* + n^* - 1\}$:*

$$\left.\frac{\partial q^s_{jc}}{\partial p^c_j}\right|_{m^*,n^*,\overline{K}_j} = \left(\frac{\partial q^d_{jc}\left(p^c_j\right)}{\partial p^c_j} + h_c\right)(1 - w_c) > 0 \ \textit{if}\ \frac{\partial q^d_{jc}\left(p^c_j\right)}{\partial p^c_j} > -\frac{1}{\gamma''(x_c)}$$

$$\left.\frac{\partial q^s_{jc}}{\partial p^{c'}_j}\right|_{m^*,n^*,\overline{K}_j} = \begin{cases} -w_c \dfrac{\partial q^d_{jc'}\left(p^{c'}_j\right)}{\partial p^{c'}_j} \geq 0 \ \textit{if}\ c' \in \{1, \ldots, m^* - 1\} \\ -w_c\left(\dfrac{\partial q^d_{jc'}\left(p^{c'}_j\right)}{\partial p^{c'}_j} + h_{c'}\right) < 0 \ \textit{if}\ c' \in \{m^*, \ldots, m^* + n^* - 1\} \ \textit{and}\ \dfrac{\partial q^d_{jc'}\left(p^{c'}_j\right)}{\partial p^{c'}_j} > -\dfrac{1}{\gamma''(x_{c'})} \\ = 0 \ \textit{if}\ c' \in \{m^* + n^*, \ldots, N\} \end{cases}$$

and

$$\left.\frac{\partial q^s_{jc}}{\partial \overline{K}_j}\right|_{m^*,n^*} = w_c$$

Proof. Totally differentiating the rationed first-order conditions $\left(p^c_j - c_j\right) + \gamma'(x_c) = \rho$ gives $dx_c = h_c\left(d\rho - dp^c_j\right)$ (with $dp^c_j = 0$ for prices that do not change), and the binding capacity constraint $\sum\nolimits_{c=1}^{m^*+n^*-1} q^d_{jc}\left(p^c_j\right) - \sum\nolimits_{c\in\mathcal{R}} x_c = \overline{K}_j - \kappa_j$, where $\mathcal{R} = \{m^*, \ldots, m^* + n^* - 1\}$, gives $\sum\nolimits_{c\in\mathcal{R}} dx_c = \sum\nolimits_{c=1}^{m^*+n^*-1} \frac{\partial q^d_{jc}}{\partial p^c_j} dp^c_j - d\overline{K}_j$. Solving for $d\rho$:

$$d\rho = \frac{1}{H}\left[\sum\nolimits_{c=1}^{m^*+n^*-1} \frac{\partial q^d_{jc}}{\partial p^c_j} dp^c_j + \sum\nolimits_{c\in\mathcal{R}} h_c dp^c_j - d\overline{K}_j\right]$$

Substituting into $dq^s_{jc} = \frac{\partial q^d_{jc}}{\partial p^c_j} dp^c_j - dx_c$ and evaluating one perturbation at a time yields each expression: for the own price, $\frac{\partial q^d_{jc}}{\partial p^c_j} + h_c - h_c \left(\frac{\partial q^d_{jc}}{\partial p^c_j} + h_c \right) / H = \left(\frac{\partial q^d_{jc}}{\partial p^c_j} + h_c \right)(1 - w_c)$; for the price of a fully served country only the demand term of the constraint moves, giving $-w_c \frac{\partial q^d_{jc'}}{\partial p^{c'}_j}$; for a rationed rival both its demand and its first-order condition move, giving $-w_c \left(\frac{\partial q^d_{jc'}}{\partial p^{c'}_j} + h_{c'} \right)$; prices of excluded countries enter neither the constraint nor the first-order conditions; and $d\overline{K}_j$ enters only through the constraint, giving w_c. □

Proposition 5 (Capacity choice with a general convex penalty)**.** *Define* $\widetilde{H} \equiv H + \pi / C'' \left(\overline{K}_j \right)$*. At the optimal capacity* $\overline{K}_j = C'^{-1}(\pi\rho)$*:*

$$\left.\frac{\partial \overline{K}_j}{\partial p^c_j}\right|_{m^*,n^*} = \begin{cases} \frac{\pi}{C''\left(\overline{K}_j\right)\widetilde{H}} \frac{\partial q^d_{jc}\left(p^c_j\right)}{\partial p^c_j} < 0 \text{ if } c \in \{1, \dots, m^* - 1\} \\ \frac{\pi}{C''\left(\overline{K}_j\right)\widetilde{H}} \left(\frac{\partial q^d_{jc}\left(p^c_j\right)}{\partial p^c_j} + h_c \right) \text{ if } c \in \{m^*, \dots, m^* + n^* - 1\} \text{ and } > 0 \text{ if } \frac{\partial q^d_{jc}\left(p^c_j\right)}{\partial p^c_j} > -\frac{1}{\gamma''(x_c)} \\ 0 \text{ if } c \in \{m^* + n^*, \dots, N\} \end{cases}$$

and the total effects on the supply of a rationed country $c \in \{m^*, \dots, m^* + n^* - 1\}$ *are*

$$\left.\frac{\partial q^s_{jc}}{\partial p^c_j}\right|_{m^*,n^*} = \left(\frac{\partial q^d_{jc}\left(p^c_j\right)}{\partial p^c_j} + h_c \right) \left(1 - \frac{h_c}{\widetilde{H}} \right),$$

$$\left.\frac{\partial q^s_{jc}}{\partial p^{c'}_j}\right|_{m^*,n^*} = \begin{cases} -\frac{h_c}{\widetilde{H}} \frac{\partial q^d_{jc'}\left(p^{c'}_j\right)}{\partial p^{c'}_j} > 0 \text{ if } c' \in \{1, \dots, m^* - 1\} \\ -\frac{h_c}{\widetilde{H}} \left(\frac{\partial q^d_{jc'}\left(p^{c'}_j\right)}{\partial p^{c'}_j} + h_{c'} \right) \text{ if } c' \in \{m^*, \dots, m^* + n^* - 1\} \\ 0 \text{ if } c' \in \{m^* + n^*, \dots, N\} \end{cases}$$

Proof. Differentiating the first-order condition $C'\left(\overline{K}_j\right) = \pi\rho$ gives $d\overline{K}_j = \pi\, d\rho / C''\left(\overline{K}_j\right)$. Substituting into the expression for $d\rho$ in the proof of Proposition 4 and solving:

$$d\rho = \frac{1}{\widetilde{H}} \left[\sum_{c=1}^{m^*+n^*-1} \frac{\partial q^d_{jc}}{\partial p^c_j} dp^c_j + \sum_{c \in \mathcal{R}} h_c dp^c_j \right]$$

The expressions then follow exactly as in the proof of Proposition 4, with H replaced by $\widetilde{H}$

for the total effects and $d\overline{K}_j = \pi \, d\rho / C''$ for the capacity responses. Under Assumption 2 and $C(x) = \frac{x^2}{2}$, $\widetilde{H} = (n^* + \pi\gamma)/\gamma$, and all expressions reduce to those of Section 2.1. $\square$

Three observations follow. First, every sign and every economic statement of Section 2.1 is preserved: capacity freed by a fully served country (and marginal capacity, $\partial q^s_{jc} / \partial \overline{K}_j = w_c$) is allocated across rationed countries in proportion to the inverse local curvature of the penalty – capacity flows where the marginal penalty rises most slowly– with equal sharing as the quadratic special case. Second, since $\widetilde{H} > H \geq h_c$, the factor $1 - h_c/\widetilde{H}$ lies in $(0, 1)$: the capacity adjustment dampens but never reverses the fixed-capacity reallocation, and a price increase in a fully served country unambiguously raises the supply allocated to rationed countries, for any strictly convex γ and C. Third, what does require the quadratic specification are the closed-form expressions for ρ, q^s_{jc} and $\overline{K}_j$, which are only implicitly characterized in general, and the fact that the elasticity conditions involve the primitive constant γ: in general, $\gamma''(x_c)$ is evaluated at the endogenous equilibrium shortage. The latter is easily restored with a curvature bound: if $\overline{\gamma}'' \equiv \sup_x \gamma''(x) < \infty$ on the relevant range of shortages, then $\frac{\partial q^d_{jc}\left(p^c_j\right)}{\partial p^c_j} > -1/\overline{\gamma}''$ implies the local conditions above at every equilibrium, exactly parallel to $\frac{\partial q^d_{jc}\left(p^c_j\right)}{\partial p^c_j} > -\frac{1}{\gamma}$ in the main text (of which it is the special case $\overline{\gamma}'' = \gamma$). Strict convexity itself is needed for uniqueness of the rationed allocation: weakly convex penalties with linear segments would leave the allocation across rationed countries indeterminate.

A.3 Counterfactuals and estimation details

We specify a discrete choice model in which individuals, together with their physician make a choice of a drug relevant for the treatment needed. We study the market of antipsychotic drugs, mainly used to treat schizophrenia and bipolar disorder, but drugs are also used for Parkinson's disease, for Alzheimer's in case of aggression, and severe or acute agitation and psychosis. The sample includes the following antipsychotic active substances: Paliperidone, Aripiprazole, Quetiapine, Clozapine, Olanzapine, Amisulpride, Haloperidol, and Chlorpromazine. [9]

Discrete choice model specification. The market is defined by month and year, and each individual together with their physician need to choose a drug j to treat their condition. The indirect utility of an individual i that searches for drug j is given by:

[9] ATC5 codes (in order): N05AX12, N05AH04, N05AH02, N05AH03, N05AL05, N05AE04, N05AD01, N05AA01. Other ATC5 classes for these types of conditions might include Risperidone and Ziprasidone, Lurasidone, Cariprazine and Brexpiprazole (N05AX13, N05AE05, N05AX15, N05AX16) but are not in our sample.

$$u_{ijt} = \beta \log(p_{jt}^{FR}) + \theta x_j + \gamma_t + \xi_{jt} + \zeta_{igt} + (1-\omega)\,\varepsilon_{ijt} \tag{17}$$

$$u_{ijt} = \delta_{jt} + \zeta_{igt} + (1-\omega)\,\varepsilon_{ijt} \tag{18}$$

The individual chooses $j : \arg\max_{j \in J_t} u_{ijt}$. Assume $\zeta_{igt} + (1-\omega)\varepsilon_{ijt}$ has Type 1 extreme value distribution, then the probability of choosing drug j (for each t):

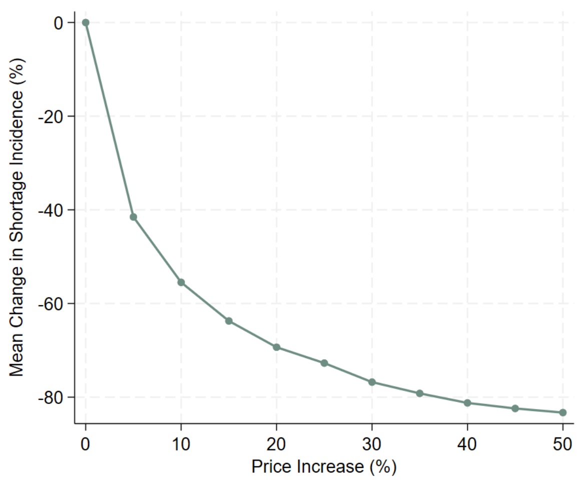
$$\sigma_j = \frac{\exp\left(\frac{\delta_j}{1-\omega}\right)}{\left(\sum_{k \in g} \exp\left(\frac{\delta_k}{1-\omega}\right)\right)^{\omega} \left(1 + \sum_{g'} \left(\sum_{k \in g'} \exp\left(\frac{\delta_k}{1-\omega}\right)\right)^{1-\omega}\right)} \tag{19}$$

The market is defined at the month-year level, and we use the quantities predicted from our demand predictive model in Section 2.2. We assume the quantity demanded is equal to the supply or observed quantities if there is no shortage according to Section 2.2, and equal to the predicted quantity of demand if there is a shortage. We aggregate some products with different box sizes by using drug-daily-doses. This information comes from `https://atcddd.fhi.no/atc_ddd_index/?code=N05AX&showdescription=no` The market size calculations are based on the number of patients diagnosed with some psychotic conditions by year according to the French Public Insurance information (`https://data.ameli.fr/explore/dataset/effectifs/table/?refine.patho_niv1=Maladies+psychiatriques&refine.patho_niv2=Troubles+psychotiques&sort=-prev.`) We assume each person takes 30 daily doses per month. For 2024 there is no information and we project the market size as 6% larger than 2023. The market size numbers are in line to what Kohn et al. (2004) compute: approximately the share of untreated schizophrenia is around 30%.

Construction of instruments. Potential instruments include: Number of competitors in genericized market for each month-year. Number of labs within ATC5 class for each month-year. Strength of the product minus mean strength within ATC5 class.

Counterfactuals results.

Figure 8: Decrease in shortages due to price increase for a set of drugs



Notes: Percentage decrease in number of shortages by ATC5 and time, averaged for each percentage increase in price. 0% increase represents the baseline. The price increase happens only for the products with the largest baseline shortages within ATC5 for the whole sample period.