Online Appendix

for

Resilience of Global Supply Chains and Generic Drug Shortages

Anaïs Galdin

A Model: Proofs and Derivations **2**
A.1 Proof: Capacity Is Never Set Below Contracted Demand . . . 2
A.2 First-Order Conditions . . . 2
A.3 Derivation of Realized Consumer Surplus under Shortages . . . 4

B Estimation Details **7**
B.1 Defining Products and Drug Markets . . . 7
B.2 Plant Location-Specific Statistics . . . 8
B.3 Parameter Estimates and Elasticities: Tables . . . 9
B.4 Demand Estimation Under Censoring: Algorithm . . . 13
B.5 Gandhi–Houde Differentiation Instruments . . . 14
B.6 Supply Estimation: Moment Conditions and Estimation Algorithms . . . 15
B.7 Nested Fixed Point and Two-Step GMM . . . 16

C Counterfactual Details **20**
C.1 Counterfactual Changes in Realized Consumer Surplus . . . 20
C.2 Reshoring Subsidies: Details . . . 22
C.3 Shortage-Contingent Pricing: Details . . . 24
C.4 Failure-to-Supply Penalties: Details . . . 27
C.5 What It Costs to Reduce Shortages . . . 29

D Market Structure: Institutions and Prices **30**
D.1 How Procurement Auctions Work . . . 30
D.2 Which Drugs Go Short and Why . . . 31
D.3 Prices around Shortage Onset . . . 33
D.4 Do Buyers Discipline Suppliers After a Shortage? . . . 35

E Reduced-Form Results: Details **38**
E.1 Offshoring and Shortages: Correlations . . . 38
E.2 A Dynamic Difference-in-Differences Framework . . . 38
E.3 A Shift-Share Instrument for Offshoring: Details . . . 42

F Data Construction **50**
F.1 Sources and Observation Units . . . 50
F.2 Record Linkage: Techniques and Results . . . 52

A. Model: Proofs and Derivations

A.1. Proof: Capacity Is Never Set Below Contracted Demand

Fix a market and the choices of competing suppliers, and suppress market and time subscripts. Under the maintained logit demand, $Q_j^D(p_j)$ is continuous, strictly decreasing in p_j, and tends to zero as p_j tends to infinity. Assume $\mathbb{E}_{\omega\sim F_\ell}[\omega] > 0$.

Consider a positive-capacity choice (p_j, κ_j) with $\kappa_j < Q_j^D(p_j)$. There exists $p_j' > p_j$ such that $Q_j^D(p_j') = \kappa_j$. Since $\omega \in [0,1]$, sales equal $\kappa_j\omega$ at both prices for every yield realization. Capacity costs and location fixed costs are unchanged, while expected profit increases by

$$(p_j' - p_j)\kappa_j \mathbb{E}_{\omega\sim F_\ell}[\omega] > 0.$$

Hence a positive-capacity best response cannot satisfy $\kappa_j < Q_j^D(p_j)$. At a positive-capacity optimum, $\theta_j = Q_j^D(p_j)/\kappa_j \leq 1$, so the minimum operator in the shortage threshold can be omitted. □

A.2. First-Order Conditions

For a given molecule-form m and period t, manufacturer j chooses price p_{jmt} and planned capacity κ_{jmt}. Using the demand-to-capacity ratio θ_{jmt} and the function $h_\ell(\theta_{jmt})$ defined in (6)–(13), the firm's ex-ante expected profit is

$$\begin{aligned}\mathbb{E}_{\omega_\ell}\left[\pi_{jmt}^\ell\right] &= p_{jmt}\mathbb{E}_{\omega\sim F_\ell}\left[\min\left\{Q_{jmt}^D(\delta), \kappa_{jmt}\omega\right\}\right] - c_{jmt}^\ell\kappa_{jmt} \\ &= p_{jmt}Q_{jmt}^D(\delta)\left[1 - F_\ell(\theta_{jmt})\right] + p_{jmt}\kappa_{jmt}h_\ell(\theta_{jmt}) - c_{jmt}^\ell\kappa_{jmt}. \end{aligned} \tag{A.1}$$

The first revenue term corresponds to states in which realized capacity is sufficient to meet demand, while the second revenue term corresponds to shortage states. In particular,

$$\Pr(\text{Shortage})_{jmt} = F_\ell(\theta_{jmt}),$$

and

$$h_\ell(\theta_{jmt}) = \int_0^{\theta_{jmt}} \omega f_\ell(\omega)d\omega = F_\ell(\theta_{jmt})\mathbb{E}\left[\omega_{jmt} \mid \omega_{jmt} < \theta_{jmt}\right].$$

Thus $h_\ell(\theta_{jmt})$ is the probability-weighted surviving capacity share in shortage states.

First-Order Conditions (FOCs). I first derive the capacity FOC. Since

$$\frac{\partial\theta_{jmt}}{\partial\kappa_{jmt}} = -\frac{Q_{jmt}^D(\delta)}{\kappa_{jmt}^2} = -\frac{\theta_{jmt}}{\kappa_{jmt}}, \qquad h_\ell'(\theta_{jmt}) = \theta_{jmt}f_\ell(\theta_{jmt}),$$

I obtain

$$\begin{aligned} 0 = \frac{\partial\mathbb{E}_{\omega_\ell}\left[\pi_{jmt}^\ell\right]}{\partial\kappa_{jmt}} &= p_{jmt}\left[\theta_{jmt}^2 f_\ell(\theta_{jmt}) + h_\ell(\theta_{jmt}) - \theta_{jmt}^2 f_\ell(\theta_{jmt})\right] - c_{jmt}^\ell \\ &= p_{jmt}h_\ell(\theta_{jmt}) - c_{jmt}^\ell. \end{aligned}$$

Therefore,

$$\frac{c^{\ell}_{jmt}}{p_{jmt}} = h_{\ell}(\theta_{jmt}) = \int_0^{\theta_{jmt}} \omega f_{\ell}(\omega) d\omega. \tag{FOC 1}$$

Next, I derive the price FOC. Since

$$\frac{\partial \theta_{jmt}}{\partial p_{jmt}} = \frac{1}{\kappa_{jmt}} \frac{\partial Q^D_{jmt}(\delta)}{\partial p_{jmt}},$$

differentiating (A.1) with respect to p_{jmt} gives

$$\begin{aligned} 0 = \frac{\partial \mathbb{E}_{\omega_\ell}\left[\pi^{\ell}_{jmt}\right]}{\partial p_{jmt}} &= Q^D_{jmt}(\delta)\left[1 - F_{\ell}(\theta_{jmt})\right] + \kappa_{jmt} h_{\ell}(\theta_{jmt}) \\ &\quad + p_{jmt}\left[\frac{\partial Q^D_{jmt}(\delta)}{\partial p_{jmt}}\left[1 - F_{\ell}(\theta_{jmt})\right] - Q^D_{jmt}(\delta) f_{\ell}(\theta_{jmt}) \frac{\partial \theta_{jmt}}{\partial p_{jmt}}\right. \\ &\quad \left. + \kappa_{jmt} h'_{\ell}(\theta_{jmt}) \frac{\partial \theta_{jmt}}{\partial p_{jmt}}\right]. \end{aligned}$$

Using

$$h'_{\ell}(\theta_{jmt}) = \theta_{jmt} f_{\ell}(\theta_{jmt}) \qquad \text{and} \qquad \theta_{jmt} = \frac{Q^D_{jmt}(\delta)}{\kappa_{jmt}},$$

the two terms involving $f_{\ell}(\theta_{jmt})$ cancel. Hence the price FOC is

$$\left[1 - F_{\ell}(\theta_{jmt})\right]\left[Q^D_{jmt}(\delta) + p_{jmt}\frac{\partial Q^D_{jmt}(\delta)}{\partial p_{jmt}}\right] + \kappa_{jmt} h_{\ell}(\theta_{jmt}) = 0. \tag{FOC 2}$$

I rewrite the second FOC using the log-price logit derivative:

$$\frac{\partial Q^D_{jmt}(\delta)}{\partial p_{jmt}} = -\frac{\alpha}{p_{jmt}} Q^D_{jmt}(\delta)(1 - s^D_{jmt}).$$

This implies that the marginal revenue term is

$$Q^D_{jmt}(\delta) + p_{jmt}\frac{\partial Q^D_{jmt}(\delta)}{\partial p_{jmt}} = Q^D_{jmt}(\delta)\left[1 - \alpha(1 - s^D_{jmt})\right].$$

Combining (FOC 1) and (FOC 2). Substituting the log-price demand derivative into (FOC 2) gives

$$\left[1 - F_{\ell}(\theta_{jmt})\right] Q^D_{jmt}(\delta)\left[1 - \alpha(1 - s^D_{jmt})\right] + \kappa_{jmt} h_{\ell}(\theta_{jmt}) = 0.$$

Dividing by κ_{jmt} and using (FOC 1), namely

$$\frac{c^{\ell}_{jmt}}{p_{jmt}} = h_{\ell}(\theta_{jmt}),$$

yields

$$\underbrace{\left[1 - F_{\ell}(\theta_{jmt})\right]}_{\text{Pr(no shortage)}} \underbrace{\theta_{jmt}}_{\text{DCR}} \left[1 - \alpha(1 - s^D_{jmt})\right] = -\frac{c^{\ell}_{jmt}}{p_{jmt}}.$$

Here $\theta_{jmt} = Q^D_{jmt}(\delta)/\kappa_{jmt}$ is the demand-to-capacity ratio. In equilibrium, $\kappa_{jmt} \geq Q^D_{jmt}(\delta)$, so $\theta_{jmt} \in (0, 1]$.

Define the own-price elasticity under log-price demand as

$$\eta_{jjmt} \equiv \frac{\partial Q^D_{jmt}(\delta)}{\partial p_{jmt}} \frac{p_{jmt}}{Q^D_{jmt}(\delta)} = -\alpha(1 - s^D_{jmt}).$$

Then the combined FOC can be written as

$$\left[1 - F_\ell(\theta_{jmt})\right] \theta_{jmt} \left(1 + \eta_{jjmt}\right) = -\frac{c^\ell_{jmt}}{p_{jmt}}.$$

From these conditions, equilibrium capacity κ^*_{jmt} satisfies

$$\kappa^*_{jmt} = -\frac{p^*_{jmt}}{c^\ell_{jmt}} \left(1 + \eta_{jjmt}\right) Q^D_{jmt}(\delta^*) \left[1 - F_\ell(\theta^*_{jmt})\right],$$

where

$$\theta^*_{jmt} = \frac{Q^D_{jmt}(\delta^*)}{\kappa^*_{jmt}}.$$

Using the capacity FOC,

$$\frac{c^\ell_{jmt}}{p^*_{jmt}} = h_\ell(\theta^*_{jmt}),$$

I can equivalently write

$$\kappa^*_{jmt} = -\left(1 + \eta_{jjmt}\right) \left[\frac{1 - F_\ell(\theta^*_{jmt})}{h_\ell(\theta^*_{jmt})}\right] Q^D_{jmt}(\delta^*).$$

The ratio in brackets compares the probability of no shortage to the probability-weighted surviving capacity share in shortage states.

Finally, solving for the equilibrium price gives the markup form

$$p^*_{jmt} = \underbrace{-\left(\frac{1}{1 + \eta_{jjmt}}\right) c^\ell_{jmt}}_{\text{markup over marginal cost}} \underbrace{\left(\frac{\kappa^*_{jmt}}{Q^D_{jmt}(\delta^*)\left[1 - F_\ell(\theta^*_{jmt})\right]}\right)}_{\text{capacity/no-shortage adjustment}} \tag{A.2}$$

with

$$\eta_{jjmt} = -\alpha(1 - s^{D,*}_{jmt}).$$

A.3. Derivation of Realized Consumer Surplus under Shortages

This appendix derives the conditional utility of buyers displaced into the outside option by a shortage and the realized consumer surplus expression used in Section 4.6. To keep notation light, I suppress market and time subscripts; all objects refer to a given market (m, t).

Let

$$CS^{\text{logit}}(\delta) \equiv \ln\left(1 + \sum_j e^{\delta_j}\right) = -\ln s^D_0(\delta)$$

denote the frictionless logit surplus index, excluding the Euler–Mascheroni constant.

Mass of unaffected and shortage-displaced buyers. Unaffected buyers are buyers whose realized outcome coincides with their ex-ante choice. This group includes buyers who initially chose an inside good and receive it, as well as buyers who initially chose the outside option. Since the realized inside share is $1 - s_0^{\text{real}}$ and the latent outside share is $s_0^D(\delta)$, the mass of unaffected buyers is

$$\left(1 - s_0^{\text{real}}\right) + s_0^D(\delta) = 1 - \left[s_0^{\text{real}} - s_0^D(\delta)\right].$$

Shortage-displaced buyers are buyers who initially chose an inside good but are forced into the outside option after supply shocks realize. Their mass is the excess outside share:

$$s_0^{\text{real}} - s_0^D(\delta) = \sum_j s_j^D(\delta)(1 - r_j).$$

Utility of unaffected buyers. A standard property of the logit model is that the expected utility of the chosen alternative, conditional on that alternative being chosen, equals the inclusive value plus the Euler–Mascheroni constant. Therefore, buyers whose realized outcome coincides with their ex-ante choice obtain

$$E[U_{\text{unaffected}}] = \gamma_E + CS^{\text{logit}}(\delta) = \gamma_E - \ln s_0^D(\delta). \tag{A.3}$$

Utility of shortage-displaced buyers. Shortage-displaced buyers initially choose an inside good but receive the outside option. Their realized utility is the outside-good utility conditional on having rejected the outside option ex ante. Since the outside option has mean utility normalized to zero,

$$u_{i0} = \varepsilon_{i0}, \qquad E[u_{i0}] = \gamma_E.$$

Using the law of iterated expectations for the outside-option utility,

$$E[u_{i0}] = \underbrace{s_{0mt}^D(\delta) E[u_{i0} \mid \text{chose } 0]}_{\text{Those who wanted 0}} + \underbrace{(1 - s_{0mt}^D(\delta)) E[u_{i0} \mid \text{chose } j \neq 0]}_{\text{Those who wanted inside goods}}. \tag{A.4}$$

Conditional on choosing the outside option, the outside option is the utility-maximizing alternative. Hence

$$E[u_{i0} \mid \text{chose } 0] = \gamma_E + CS^{\text{logit}}(\delta).$$

Setting $E[U_{\text{displaced}}] \equiv E[u_{i0} \mid \text{chose } j \neq 0]$, I can thus rewrite:

$$\gamma_E = s_{0mt}^D(\delta)(CS_{mt}^{\text{logit}}(\delta) + \gamma_E) + (1 - s_{0mt}^D(\delta)) E[U_{\text{displaced}}]. \tag{A.5}$$

Solving,

$$E[U_{\text{displaced}}] = \gamma_E - \frac{s_{0mt}^D(\delta)}{1 - s_{0mt}^D(\delta)} CS_{mt}^{\text{logit}}(\delta) = \gamma_E - \left(\frac{s_{0mt}^D(\delta)}{1 - s_{0mt}^D(\delta)}\right)(-\ln s_{0mt}^D(\delta)). \tag{A.6}$$

Thus, shortage-displaced buyers receive less than the unconditional mean outside-option utility, γ_E, because they are forced into an option they rejected ex ante.

Average shortage loss. The loss per shortage-displaced buyer is the difference between the expected utility of receiving the initially chosen alternative and the expected utility of being forced into the outside option:

$$\begin{aligned}\bar{\mathcal{L}}(\delta) &\equiv E[U_{\text{unaffected}}] - E[U_{\text{displaced}}] \\ &= \left[\gamma_E + CS^{\text{logit}}(\delta)\right] - \left[\gamma_E - \frac{s_0^D(\delta)}{1 - s_0^D(\delta)} CS^{\text{logit}}(\delta)\right] \\ &= \frac{CS^{\text{logit}}(\delta)}{1 - s_0^D(\delta)} = \frac{-\ln s_0^D(\delta)}{1 - s_0^D(\delta)}. \end{aligned} \tag{A.7}$$

This is the average welfare loss, in expected-utility units, for each buyer displaced into the outside option by a shortage.

Realized consumer surplus. Using the masses above, realized expected utility is

$$\begin{aligned} W^{\text{real}} &= \underbrace{\left[1 - \left(s_0^{\text{real}} - s_0^D(\delta)\right)\right]}_{\text{mass of unaffected buyers}} E[U_{\text{unaffected}}] + \underbrace{\left[s_0^{\text{real}} - s_0^D(\delta)\right]}_{\text{mass of shortage-displaced buyers}} E[U_{\text{displaced}}] \\ &= \gamma_E + CS^{\text{logit}}(\delta) - \left[s_0^{\text{real}} - s_0^D(\delta)\right] \bar{\mathcal{L}}(\delta). \end{aligned} \tag{A.8}$$

Dropping the Euler constant, which is common across allocations and cancels in counterfactual differences, gives

$$CS^{\text{real}}\left(\delta, s_0^{\text{real}}\right) = CS^{\text{logit}}(\delta) - \left[s_0^{\text{real}} - s_0^D(\delta)\right] \bar{\mathcal{L}}(\delta). \tag{A.9}$$

Using (A.7), this can equivalently be written as

$$CS^{\text{real}}\left(\delta, s_0^{\text{real}}\right) = CS^{\text{logit}}(\delta) \left[\frac{1 - s_0^{\text{real}}}{1 - s_0^D(\delta)}\right]. \tag{A.10}$$

Finally, substituting the fill-rate expression for the excess outside share gives

$$CS^{\text{real}} = CS^{\text{logit}}(\delta) - \left[\sum_j s_j^D(\delta)(1 - r_j)\right] \bar{\mathcal{L}}(\delta). \tag{A.11}$$

B. Estimation Details

B.1. Defining Products and Drug Markets

The final sample is a balanced quarterly panel of NDC-54 drug products and manufacturing plants covering 2004Q1 to 2019Q4. A unit of observation is a manufacturing plant (defined at the labeler-plant level), NDC-54 drug product, molecule-form market, fiscal quarter. An NDC-54 combines the FDA's five-digit labeler code and four-digit product code and therefore identifies a labeler-specific drug product (its active ingredient, strength, dosage form, and formulation) without distinguishing the commercial package size recorded by the final two digits of the NDC. Sales prices and quantities are from IQVIA, which reports them at the NDC-542 package level. I aggregate these observations across package codes to the NDC-54 level and compute price per each as total sales divided by total eaches. For the baseline demand estimation, I aggregate the data to the NDC-54-year level and define a market as an ATC-5 therapeutic market interacted with a year. ATC5 markets may contain several molecule–dosage-form groups; the Molform grouping reported separately in this appendix is therefore distinct from the market definition used in the structural estimation and counterfactuals. The resulting estimation sample covers 2005 to 2019, and an observation is an NDC-54 product-year within an ATC-5-year market.

Table B.1: **Number of NDC-54 Products and Labelers by Drug Market-Year**

		p50	Mean	SD	Min.	p25	p75	p95	Max.	Market-Years
Molform	Number of NDC-54s	2	3.95	5.42	1	1	4	16	39	7,066
	Number of labelers	1	2.05	2.00	1	1	2	7	12	7,076
ATC-4	Number of NDC-54s	6	10.41	11.04	2	3	13	34	68	2,811
	Number of labelers	3	4.44	3.55	1	2	6	12	20	2,815
ATC-5	Number of NDC-54s	4	6.61	6.26	2	2	8	20	42	4,436
	Number of labelers	2	3.11	2.32	1	1	4	8	14	4,440

Notes: The FDA's National Drug Code (NDC) identifies drugs using a unique, three-segment number in the format "5-4-2." The first set of digits identifies the labeler, such as the drug manufacturer, repackager, or distributor. The second set of digits is the product code, which identifies the specific strength, dosage form, and formulation of a drug for a specific labeler. The third set of digits identifies the package size and type. The NDC-54 level therefore defines a drug product at the labeler–molecule–dosage-form level without distinguishing commercial package sizes. Molform markets group products sharing the same molecule and dosage form using drug-name crosswalks constructed from the Orange Book, NDC Directory, FDA Drug Shortages data, and RxMix. ATC-4 and ATC-5 markets use the World Health Organization's Anatomical Therapeutic Chemical classification. ATC-4 markets group products in the same chemical subgroup, whereas ATC-5 markets group products sharing the same chemical substance. The sample covers 2005–2019 and contains 31,493 NDC-54–year observations, 4,064 distinct NDC-54 products, and 332 distinct labelers. For each market definition, I first count the distinct NDC-54 products and five-digit labeler codes within each market m and year t. The reported statistics summarize the resulting distributions across Molform–years, ATC-4–years, and ATC-5–years. Observations above the variable-specific 99th percentile are dropped separately for the NDC-54 and labeler distributions; hence, the reported number of market-years may differ across the two rows. Before this trimming, the sample contains 7,135 Molform-years, 2,837 ATC-4-years, and 4,478 ATC-5-years.

B.2. Plant Location-Specific Statistics

Table B.2: **Shortages, Prices and Sales by Plant Location (Annual Panel used for Estimation)**

Plant location	Variable	Median	Mean	Std	N Obs
	Fraction of Time Short by Year	0.00	0.14	0.30	21,334
	Shortage Dummy	0.00	0.25	0.43	21,334
	Never Short Dummy	0.00	0.17	0.38	15,827
U.S.	IQVIA Price per Each (2010 USD)	8.03	246.54	1,315.28	20,232
	IQVIA Price per Package (2010 USD)	48.69	369.41	1,480.91	20,232
	IQVIA Eaches	45,138.01	558,295.06	2,113,819.00	21,334
	IQVIA Packages	9,451.29	88,914.82	289,545.22	21,334
	IQVIA Sales (2010 USD)	533,426.96	15,007,841.00	76,315,584.00	21,334
	Fraction of Time Short by Year	0.00	0.12	0.27	17,622
	Shortage Dummy	0.00	0.22	0.41	17,622
	Never Short Dummy	0.00	0.28	0.45	12,009
OECD	IQVIA Price per Each (2010 USD)	12.43	312.14	1,288.46	16,686
	IQVIA Price per Package (2010 USD)	68.60	467.61	1,615.22	16,686
	IQVIA Eaches	32,616.40	364,305.12	2,021,358.50	17,622
	IQVIA Packages	6,843.00	91,993.58	681,819.12	17,622
	IQVIA Sales (2010 USD)	469,835.37	20,533,648.00	124,983,744.00	17,622
	Fraction of Time Short by Year	0.00	0.16	0.30	43,615
	Shortage Dummy	0.00	0.30	0.46	43,615
	Never Short Dummy	0.00	0.19	0.39	30,704
Asia	IQVIA Price per Each (2010 USD)	5.13	42.03	173.44	40,513
	IQVIA Price per Package (2010 USD)	38.42	117.48	372.43	40,513
	IQVIA Eaches	11,883.00	161,707.03	705,311.56	43,615
	IQVIA Packages	2,256.61	22,731.08	116,722.71	43,615
	IQVIA Sales (2010 USD)	88,896.05	1,647,463.75	23,782,450.00	43,615

Notes: An observation is an NDC-54 product and one of its manufacturing plants in a year, 2005–2019. Location is the region of the plant's country: U.S. (Home), OECD (Foreign North: OECD and other high-income countries) and Asia (Foreign South, primarily India and China). I split a product's eaches, packages and sales equally across its plants with a known location each year. The never-short indicator comes from the 2002–2019 shortage records and is missing for products absent from them. Prices and sales are in 2010 dollars.

B.3. Parameter Estimates and Elasticities: Tables

Table B.3: **Parameter Estimates of the Structural Model**

	Parameter	Description	Estimate	SE
Demand	α	Log-price sensitivity	1.264	0.279
	β_{brand}	Brand dummy	0.624	0.210
	β_{bg}	Branded-generic dummy	0.860	0.267
	β_A	Number of presentations	0.719	0.028
	Demand Shock Distribution ξ			
	μ_ξ	Mean demand shock	0.054	–
	σ_ξ	SD demand shock	2.343	–
Supply	**Capacity Yield Shock** ω_ℓ			
	$a_{\omega,H}$	Beta shape a (U.S.)	0.815	0.030
	$b_{\omega,H}$	Beta shape b (U.S.)	0.098	0.005
	a_{ω,F_S}	Beta shape a (Asia)	0.736	0.009
	b_{ω,F_S}	Beta shape b (Asia)	0.136	0.005
	a_{ω,F_N}	Beta shape a (OECD)	1.387	0.026
	b_{ω,F_N}	Beta shape b (OECD)	0.077	0.000
	Firm Productivity γ			
	μ_γ	Mean log productivity	23.316	0.105
	σ_γ	SD log productivity	2.284	0.008
	Trade Cost τ_ℓ			
	ϕ^{dist}	Log-distance coefficient	1.916	0.108
	$\phi^{\text{dist}\times\text{oil}}$	Distance × oil-price coefficient	-0.024	0.027
Location Choice	**Location Fixed-Cost Index** $FC_{\ell\mathcal{T}}$			

Relative to Foreign North; estimates and SEs in index units

Period	Home	Foreign South	Foreign South − Home
2005–2010	0.990 (0.201)	8.198 (2.364)	7.208
2011–2013	1.302 (0.204)	5.798 (0.950)	4.495
2014–2016	1.622 (0.182)	3.607 (0.456)	1.985
2017–2019	1.186 (0.111)	1.028 (0.173)	-0.158

	β_π	Inverse shock scale per $1 million	3.934	–

Notes: This table reports parameter estimates for Section 5.4. The first column denotes the estimation step. Demand-coefficient standard errors are heteroskedasticity-robust 2SLS standard errors. The demand shock follows $\xi \sim \mathcal{N}(\mu_\xi, \sigma_\xi^2)$. The capacity-yield shock follows $\omega_\ell \sim \text{Beta}(a_{\omega,\ell}, b_{\omega,\ell})$. Standard errors for the capacity-yield, trade-cost, and productivity parameters use $B = 50$ product–year-cluster bootstrap draws that hold demand fixed and re-estimate Steps 2 and 3. Location-choice standard errors use a cluster-sandwich variance estimator. Fixed-cost parameters enter the multinomial-logit location-choice index and are identified only up to a common additive constant. OECD is normalized to zero in each period, so only differences across locations matter. Fixed-cost estimates and standard errors are reported in index units; the Foreign South-minus-Home column reports the corresponding point difference. The coefficient β_π is calibrated and measured per million dollars.

Table B.4: **Demand Estimation Results**

Parameter	Description	Estimate	SE
Panel A: Sample and identification			
Sample	Demand sample	All drugs	–
Allocation	Manufacturing assignment	All matched plants	–
N	Observations	31,493	–
MT	ATC5 market–year cells	4,478	–
KP F	First-stage statistic	26.28	–
Panel B: Demand parameters			
α	Log-price sensitivity	1.264	0.279
β_{brand}	Brand dummy	0.624	0.210
β_{bg}	Branded-generic dummy	0.860	0.267
β_A	Number of presentations	0.719	0.028
μ_ξ	Demand-shock mean	0.054	–
σ_ξ	Demand-shock standard deviation	2.343	–
Panel C: Stage-2 demand moments			
$\overline{\eta}_{jjmt}$	Mean own-price elasticity	-1.170	–
$\overline{\eta}_{jkmt}$	Mean cross-price elasticity	0.094	–
$\Pr(\eta_{jjmt} < -1)$	Elastic observations (%)	89.3	–
$\overline{s}^{D}_{0mt}$	Mean latent outside share	0.476	–
$\overline{s}^{\text{obs}}_{0mt}$	Mean observed outside share	0.586	–
Iterations	Stage-2 iterations	9	–

Notes: This table reports demand estimates for Section 5.4. Markets are defined at the ATC5–year level, prices are IQVIA prices per each, and quantities are IQVIA eaches. The price coefficient is estimated by two-stage least squares using four lagged ATC4-peer Gandhi–Houde instruments based on same- and other-ATC5-market rival counts and distances. Heteroskedasticity-robust standard errors are reported for the Stage-1 coefficients. In Stage 2, an expectation-maximization (EM) algorithm recovers latent demand and market size using the residual-root correction. Generic products are the omitted brand-status category.

Table B.5: **Demand Elasticities by Manufacturing Location**

		Elasticity with respect to							
		Home		Foreign North		Foreign South		All plants	
Elasticity	**Description**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
$\eta_{jjmt} \equiv -\hat{\alpha}(1 - s^D_{jmt})$	Own-price elasticity	-1.132	0.223	-1.174	0.178	-1.182	0.164	-1.162	0.192
$\eta_{jkmt} \equiv \hat{\alpha} s^D_{kmt}$	Cross-price elasticity	0.108	0.114	0.091	0.109	0.085	0.096	0.095	0.106
$\Pr(\eta_{jjmt} < -1)$	Elastic observations (%)	83.9	–	89.8	–	91.2	–	88.3	–
q^D	Latent quantity share	0.638	–	0.161	–	0.201	–	1.000	–
N	Matched product–plant–year rows	15,823	–	13,561	–	36,784	–	66,168	–

Notes: This table reports plant-share-weighted means and standard deviations of annual product elasticities after the Step-2 expectation-maximization (EM) demand correction. Because utility is linear in log price, own-price elasticity is $\eta_{jjmt} = -\hat{\alpha}(1 - s^D_{jmt})$ and cross-price elasticity is $\eta_{jkmt} = \hat{\alpha} s^D_{kmt}$. The reported cross-price elasticity averages over competing products within the ATC5 market–year. Standard deviations describe cross-sectional dispersion and are not coefficient standard errors. Location-specific moments use observed manufacturing assignments and the manufacturing weights specified for the sample. The table is therefore restricted to products with observed manufacturing assignments, so its all-plants moments need not equal the full-demand-sample moments. Foreign North denotes OECD production and Foreign South denotes Asian production. The all-plants columns aggregate the three manufacturing locations. Latent quantity shares sum to one across locations.

Table B.6: **Step-4 Fixed Costs Estimates by Period**

Period	FC_H	FC_{F_N}	FC_{F_S}	$FC_{F_S} - FC_H$	**Profit-equivalent gap (in $1000)**	t-stat.
2005–2010	0.990	0	8.198	7.208	1,832	3.25
2011–2013	1.302	0	5.798	4.495	1,143	5.57
2014–2016	1.622	0	3.607	1.985	505	5.92
2017–2019	1.186	0	1.028	-0.158	-40	-1.04

Notes: The first four numerical columns report location fixed cost indices from the Step-4 conditional-logit location model. The Type-I extreme-value scale is normalized to one and candidate-location profits are measured in millions of dollars. Foreign North is normalized to zero in every period, so only differences across locations are identified. The profit-equivalent gap divides the index difference $FC_{F_S,\mathcal{T}} - FC_{H,\mathcal{T}}$ by the benchmark calibration $\beta^{\text{cal}}_{\pi} = 3.934$ and converts the result to thousands of dollars. It is the candidate-profit difference that offsets the estimated location wedge under the benchmark calibration; it should not be interpreted as an independently identified plant-construction cost. The final column uses standard errors clustered at the product level (NDC-54 in the data) across years.

The estimated Foreign South–Home fixed-cost index difference declines from 7.208 in 2005–2010 to -0.158 in 2017–2019, although the final-period difference is not statistically distinguishable from zero ($t = -1.04$). These estimates measure residual location-cost wedges conditional on candidate profits and the calibrated profit coefficient, not absolute plant-construction costs.

Table B.7: **Moments: Mean Estimates**

Outcome	Description	Estimate	SE
Capacity Yield			
ω_H	Home	0.8923	0.0058
ω_{F_S}	Asia	0.8439	0.0046
ω_{F_N}	OECD	0.9475	0.0031
Capacity Yield Conditional on a Shortage			
$E[\omega \mid \omega < \theta]_H$	Home	0.5942	0.0018
$E[\omega \mid \omega < \theta]_{F_S}$	Asia	0.4938	0.0010
$E[\omega \mid \omega < \theta]_{F_N}$	OECD	0.7723	0.0010
Expected Share of Excess Demand			
s_H^{excess}	Home	0.0879	0.0004
$s_{F_S}^{\text{excess}}$	Asia	0.1190	0.0003
$s_{F_N}^{\text{excess}}$	OECD	0.0459	0.0001
Expected Share of Excess Demand Conditional on a Shortage			
$E[s^{\text{excess}} \mid \omega < \theta]_H$	Home	0.3584	0.0008
$E[s^{\text{excess}} \mid \omega < \theta]_{F_S}$	Asia	0.4284	0.0004
$E[s^{\text{excess}} \mid \omega < \theta]_{F_N}$	OECD	0.2081	0.0005
Shortage Probability $F_\ell(\theta)$			
$F_H(\theta)$	Home	0.2451	0.0015
$F_{F_S}(\theta)$	Asia	0.2777	0.0009
$F_{F_N}(\theta)$	OECD	0.2208	0.0010
Capacity Buffer			
$\bar{B}_H$	Home	3.121	2.174
$\bar{B}_{F_S}$	Asia	1.406	0.310
$\bar{B}_{F_N}$	OECD	0.097	0.011
Demand-to-Capacity Ratio (Lower θ Means More Spare Capacity)			
θ_H	Home	0.8924	0.0036
θ_{F_S}	Asia	0.8369	0.0022
θ_{F_N}	OECD	0.9616	0.0016
Marginal Cost (Dollars per Each)			
c_H	Home	49.333	3.297
c_{F_S}	Asia	17.830	0.817
c_{F_N}	OECD	128.661	6.977

Notes: A supplier is in shortage when $\omega_{jmt} < \theta_{jmt}$, with probability $F_\ell(\theta_{jmt})$. The expected share of excess demand is $F_\ell(\theta_{jmt}) - h_\ell(\theta_{jmt})/\theta_{jmt}$. Conditional rows are model expectations over shortage states, weighting contracts by $F_\ell(\theta)$. The capacity buffer is $B_{jmt} = \kappa_{jmt} - Q_{jmt}^D$. Standard errors are reported in the final column; buffer standard errors use the delta method.

Table B.8: **Sensitivity of the Foreign South–Home Fixed-Cost Wedge**

Period	Low response	Intermediate response	Benchmark response
2005–2010	0.555	4.102	7.208
2011–2013	-0.139	2.501	4.495
2014–2016	-1.057	0.807	1.985
2017–2019	-1.685	-0.615	-0.158

Notes: The table reports $FC_{F_S,\mathcal{T}} - FC_{H,\mathcal{T}}$ in location-choice index units, with $FC_{F_N,\mathcal{T}} = 0$ in every period, the extreme-value scale normalized to one and candidate-location profits in millions of dollars. The columns differ only in how strongly location choices respond to profit differences: the benchmark response uses the calibrated coefficient $\beta_\pi^{\text{cal}} = 3.934$, the intermediate response half of it, and the low response the upper endpoint of its one-sided positive profile-likelihood interval, close to zero. I re-estimate all eight region–period indices under each response.

Table B.8 reports a declining Foreign South–Home index difference under each calibration. The final-period point estimate is negative in all three cases, while its magnitude and the date of the sign change depend on the maintained responsiveness to profit differences.

B.4. Demand Estimation Under Censoring: Algorithm

This appendix details the censored-demand procedure summarized in Section 5.2. Let $\mathcal{U}_{mt}$ and $\mathcal{C}_{mt}$ denote the sets of unconstrained and constrained (in shortage) products in market-year pair (m, t), and let $Q_{mt} \equiv \sum_j Q_{jmt}^D$ denote total inside latent demand.

1. **Initialization.** Set latent total demand to observed sales, $Q_{mt}^{(0)} = \sum_j q_{jmt}^{\text{real}}$, and obtain initial slopes $(\hat{\alpha}, \hat{\boldsymbol{\beta}})$ by two-stage least squares on equation (27) using unconstrained observations in $\mathcal{U}_{mt}$ only.

2. **Market-size step.** Given $Q_{mt}^{(k)}$ and the current slopes, recover the market size $M_{mt}^{(k)}$ from the residual-root market-size calibration: market by market, $M_{mt}^{(k)}$ solves

$$M_{mt} \sum_{j \in \mathcal{U}_{mt}} s_{jmt}^D = \sum_{j \in \mathcal{U}_{mt}} q_{jmt}^{\text{real}},$$

   so that model-implied demand for unconstrained products, for which observed sales *are* latent demand, aggregates exactly to their observed sales. Holding the slopes and residuals fixed, the shares s_{jmt}^D are functions of mean utilities alone, so the condition is linear in M_{mt} with closed form $M_{mt} = \sum_{j \in \mathcal{U}_{mt}} q_{jmt}^{\text{real}} / \sum_{j \in \mathcal{U}_{mt}} s_{jmt}^D$. Market size still enters the shares indirectly, since the truncation points of the imputed residuals shift with $\log(M_{mt}\, s_{0mt}^D)$; the closed-form update is therefore iterated with the imputation step below in a within-market fixed point. Because the inside shares sum to less than one, $M_{mt} > Q_{mt}^{(k)}$ holds by construction.[1]

[1]The condition uses only unconstrained products and requires no ad hoc assumption on the outside share, in the spirit of the market-size recoveries of Huang and Rojas (2014) and Dubois and Lasio (2018). Since the unconstrained residuals are point residuals of equation (27), the condition is equivalent to $\log(M_{mt}\, s_{0mt}^D) = \hat{\mu}_{mt}$: it recovers the market size embedded in the estimated market fixed effect. Re-inverting the mean utilities at each candidate M_{mt} would instead make the condition hold for every M_{mt}.

3. **Censored-likelihood step.** Given $M_{mt}^{(k)}$, the residual implied by equation (27) is

$$\xi_{jmt}^{(k)} = \ln q_{jmt}^{\text{real}} - \ln\left(M_{mt}^{(k)} - Q_{mt}^{(k)}\right) + \hat{\alpha} \ln p_{jmt} - \hat{\boldsymbol{\beta}}' \mathbf{X}_{jmt}.$$

For $j \in \mathcal{U}_{mt}$, this residual is point identified. For $j \in \mathcal{C}_{mt}$, $q_{jmt}^{\text{real}} \leq Q_{jmt}^{D}$ implies the lower bound $\xi_{jmt} > \underline{\xi}_{jmt}^{(k)}$, where $\underline{\xi}_{jmt}^{(k)}$ is the same expression evaluated at observed (rationed) sales. Assuming $\xi_{jmt} \sim \mathcal{N}(\mu_\xi, \sigma_\xi^2)$, I estimate $\left(\hat{\mu}_\xi^{(k)}, \hat{\sigma}_\xi^{(k)}\right)$ by maximizing the censored log likelihood

$$\mathcal{L}(\mu_\xi, \sigma_\xi) = \sum_{j \in \mathcal{U}} \ln\left[\frac{1}{\sigma_\xi}\phi\left(\frac{\xi_{jmt}^{(k)} - \mu_\xi}{\sigma_\xi}\right)\right] + \sum_{j \in \mathcal{C}} \ln\left[1 - \Phi\left(\frac{\underline{\xi}_{jmt}^{(k)} - \mu_\xi}{\sigma_\xi}\right)\right],$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and distribution functions.

4. **Imputation step.** Update latent total demand by replacing constrained sales with expected latent demand conditional on the truncation event:

$$Q_{mt}^{(k+1)} = \sum_{j \in \mathcal{U}_{mt}} q_{jmt}^{\text{real}} + \sum_{j \in \mathcal{C}_{mt}} \mathbb{E}\left[Q_{jmt}^{D}(\xi) \,\middle|\, \xi > \underline{\xi}_{jmt}^{(k)}; \hat{\mu}_\xi^{(k)}, \hat{\sigma}_\xi^{(k)}\right],$$

where the expectation is evaluated under the corresponding truncated normal distribution.[2]

5. **Slope update and convergence.** Re-estimate equation (27) by two-stage least squares with constrained observations included at their imputed latent quantities and unconstrained observations at observed sales, then return to the market-size step. Iterate until the vector $\{Q_{mt}^{(k)}\}$ converges. Including constrained observations in later passes addresses the selection concern in Step 1 of Section 5.2: the converged slope estimates use all product–years, not only the unconstrained subsample.

Demand estimation results and elasticities are reported in Appendix Section B.3. For counterfactual analysis, I use the point-identified $\hat{\xi}_{jmt}$ for unconstrained observations and, for constrained observations, draws from the truncated distribution $\mathcal{TN}\left(\hat{\mu}_\xi, \hat{\sigma}_\xi; \underline{\xi}_{jmt}, \infty\right)$. Truncated draws are consistent with the occurrence of a shortage and preserve the equilibrium correlation between high demand shocks and supply constraints.

B.5. Gandhi–Houde Differentiation Instruments

To address price endogeneity in demand estimation (Section 5.2), I construct differentiation instruments à la Gandhi and Houde (2019) from the composition of each product's therapeutic neighborhood. The idea is that the number and proximity of rival products in therapeutic space shift the competitive pressure faced by a product and, hence, its equilibrium markup.

[2] Updates are bounded to prevent numerically extreme imputations in thin markets; the bounds never bind at the converged solution for the markets used in estimation.

Let m denote an ATC5 market and let $a(m)$ denote its ATC4 class. Let $\mathcal{J}_{t-1}$ denote the sterile injectable products observed four quarters before period t. I define two sets of rival products:

$$\Omega_{jmt}^{\text{same}} = \{k \in \mathcal{J}_{t-1} : m_k = m,\ \text{labeler}_k \neq \text{labeler}_j\},$$
$$\Omega_{jmt}^{\text{other}} = \{k \in \mathcal{J}_{t-1} : a(m_k) = a(m),\ m_k \neq m,\ \text{labeler}_k \neq \text{labeler}_j\}.$$

The set $\Omega_{jmt}^{\text{same}}$ contains rival products in the same ATC5 market as product j, whereas $\Omega_{jmt}^{\text{other}}$ contains rival products in other ATC5 markets within the same ATC4 class.

For each pair (j,k), let d_{jk} measure the distance between products j and k in the space of observed product characteristics, including route of administration, dosage form, therapeutic classification, brand–generic and marketing status, and presentation breadth. I transform this distance into the proximity weight

$$w_{jk} = \frac{1}{1+d_{jk}},$$

so that products that are closer in observed characteristic space receive greater weight.

For each rival set, I construct a count instrument and a distance-weighted instrument:

$$z_{jmt}^{\text{same,count}} = \left|\Omega_{jmt}^{\text{same}}\right|, \qquad z_{jmt}^{\text{same,dist}} = \sum_{k \in \Omega_{jmt}^{\text{same}}} \frac{1}{1+d_{jk}},$$
$$z_{jmt}^{\text{other,count}} = \left|\Omega_{jmt}^{\text{other}}\right|, \qquad z_{jmt}^{\text{other,dist}} = \sum_{k \in \Omega_{jmt}^{\text{other}}} \frac{1}{1+d_{jk}}.$$

The count instruments measure the number of rival products; the distance-weighted instruments give greater weight to rivals with similar characteristics. Their variation captures differences in competitive pressure within and across ATC5 markets in the same ATC4 class.

The exclusion restriction is that these lagged measures of rival presence and characteristics are orthogonal to ξ_{jmt} conditional on ATC5–year fixed effects and product controls. Rival sets are measured one year earlier and exclude products sold by the same labeler. Their orthogonality to current demand shocks remains an identifying assumption.

The four variables $(z_{jmt}^{\text{same,count}}, z_{jmt}^{\text{same,dist}}, z_{jmt}^{\text{other,count}}, z_{jmt}^{\text{other,dist}})$ are the excluded instruments for $\ln p_{jmt}$ in the baseline 2SLS demand estimation of Section 5.2. They are non-missing for the full estimation sample. First-stage strength is reported with the demand estimates in Appendix Section B.3.

B.6. Supply Estimation: Moment Conditions and Estimation Algorithms

B.6.1. Mapping Quantity Shortfalls into Annual Shortage Shares

The model is static at the contract-period level. To link quantity shortfalls to observed shortage duration, I normalize the period's length to one and assume that demand arrives uniformly and that a supplier remains available until its deliverable quantity is exhausted. Conditional on yield ω, the supplier can meet demand for a fraction $\min\{1, \kappa_{jmt}\omega / Q_{jmt}^{D}\}$ of the period. Under this mapping, the fraction of the period in shortage equals the share of unmet demand in equation (10).

Recall from equation (13) that $h_\ell(\theta) = \int_0^\theta \omega f_\ell(\omega)\, d\omega$. Taking expectations of equation (10) over the yield shock $\omega \sim F_\ell$ gives the model-implied mean shortage share:

$$\begin{aligned}\mathbb{E}_{\omega\sim F_\ell}\left[s_{jmt}^{\text{excess}}(\omega)\right] &= \int_0^{\theta_{jmt}} \left(1 - \frac{\omega}{\theta_{jmt}}\right) f_\ell(\omega)\, d\omega \\ &= F_\ell(\theta_{jmt}) - \frac{h_\ell(\theta_{jmt})}{\theta_{jmt}}. \end{aligned} \tag{B.1}$$

This establishes the mean duration share used in the main text.

Second moment and conditional variance. To construct the conditional-variance moment, define the second truncated raw moment of capacity yield,

$$h_\ell^{(2)}(\theta) \equiv \int_0^\theta \omega^2 f_\ell(\omega)\, d\omega. \tag{B.2}$$

The unconditional second moment of the firm-level excess-demand share is

$$\mathbb{E}_{\omega\sim F_\ell}\left[\left(s_{jmt}^{\text{excess}}(\omega)\right)^2\right] = F_\ell(\theta_{jmt}) - \frac{2h_\ell(\theta_{jmt})}{\theta_{jmt}} + \frac{h_\ell^{(2)}(\theta_{jmt})}{\theta_{jmt}^2}. \tag{B.3}$$

For a given observation, the first two moments conditional on a positive shortage are therefore

$$\mathbb{E}_{\omega\sim F_\ell}\left[s_{jmt}^{\text{excess}}(\omega) \,\middle|\, s_{jmt}^{\text{excess}}(\omega) > 0\right] = \frac{F_\ell(\theta_{jmt}) - \dfrac{h_\ell(\theta_{jmt})}{\theta_{jmt}}}{F_\ell(\theta_{jmt})}, \tag{B.4}$$

$$\mathbb{E}_{\omega\sim F_\ell}\left[\left(s_{jmt}^{\text{excess}}(\omega)\right)^2 \,\middle|\, s_{jmt}^{\text{excess}}(\omega) > 0\right] = \frac{F_\ell(\theta_{jmt}) - \dfrac{2h_\ell(\theta_{jmt})}{\theta_{jmt}} + \dfrac{h_\ell^{(2)}(\theta_{jmt})}{\theta_{jmt}^2}}{F_\ell(\theta_{jmt})}. \tag{B.5}$$

Therefore the conditional variance used in the GMM estimation is

$$\begin{aligned}&\text{Var}_{\omega\sim F_\ell}\left[s_{jmt}^{\text{excess}}(\omega) \,\middle|\, s_{jmt}^{\text{excess}}(\omega) > 0\right] \\ &\quad= \mathbb{E}_{\omega\sim F_\ell}\left[\left(s_{jmt}^{\text{excess}}(\omega)\right)^2 \,\middle|\, s_{jmt}^{\text{excess}}(\omega) > 0\right] - \left(\mathbb{E}_{\omega\sim F_\ell}\left[s_{jmt}^{\text{excess}}(\omega) \,\middle|\, s_{jmt}^{\text{excess}}(\omega) > 0\right]\right)^2. \end{aligned} \tag{B.6}$$

Because θ_{jmt} varies across observations, the location-level conditional variance targeted in estimation is not the average of (B.6). It is the variance in the pooled distribution obtained by drawing a product–market–year observation within location ℓ, drawing ω from F_ℓ, and conditioning on a positive shortage. The next subsection gives this pooled moment explicitly.

B.7. Nested Fixed Point and Two-Step GMM

The parameters of the supply-shock distribution and the equilibrium capacities are jointly determined, creating a nested fixed-point problem. The outer problem estimates the location-specific parameters

$$\Theta_\omega \equiv \{(a_{\omega\ell}, b_{\omega\ell}) : \ell \in \{H, F_N, F_S\}\}. \tag{B.7}$$

while the inner problem computes, for each candidate Θ_ω, the equilibrium demand-to-capacity ratios $\theta_{jmt}(\Theta_\omega)$ and hence the implied capacities $\kappa_{jmt}(\Theta_\omega) = Q_{jmt}^D / \theta_{jmt}(\Theta_\omega)$ for every observation.

Data Moments. Let N_ℓ denote the number of product–market–year observations in the supply-estimation sample with observed plant location $\ell(j) = \ell$. For each location $\ell \in \{H, F_N, F_S\}$, I compute three empirical moments: the empirical shortage-event rate, the mean annual excess-demand share, and the variance of the excess-demand share conditional on a shortage occurring during the period:

$$\widehat{P}_\ell^{\text{data}} = \frac{1}{N_\ell} \sum_{jmt:\ell(j)=\ell} \rho_{jmt}^{\text{obs}}, \tag{B.8}$$

$$\widehat{S}_\ell^{\text{data}} = \frac{1}{N_\ell} \sum_{jmt:\ell(j)=\ell} s_{jmt}^{\text{excess,obs}}, \tag{B.9}$$

$$\widehat{V}_{\ell,+}^{\text{data}} = \text{Var}\left(s_{jmt}^{\text{excess,obs}} \;\middle|\; \rho_{jmt}^{\text{obs}} = 1,\ \ell(j) = \ell \right). \tag{B.10}$$

Model-implied moments. For a candidate Θ_ω and the corresponding values $\theta_{jmt}(\Theta_\omega)$, the model-implied event rate and mean excess-demand share are

$$\widehat{P}_\ell(\Theta_\omega) = \frac{1}{N_\ell} \sum_{jmt:\ell(j)=\ell} F_\ell\big(\theta_{jmt}(\Theta_\omega)\big). \tag{B.11}$$

$$\widehat{S}_\ell(\Theta_\omega) = \frac{1}{N_\ell} \sum_{jmt:\ell(j)=\ell} \left[F_\ell\big(\theta_{jmt}(\Theta_\omega)\big) - \frac{h_\ell\big(\theta_{jmt}(\Theta_\omega)\big)}{\theta_{jmt}(\Theta_\omega)} \right]. \tag{B.12}$$

To construct the model-implied conditional variance of shortage duration among products that go short, note that an observation's probability of being short is $F_\ell(\theta_{jmt}(\Theta_\omega))$. The model-implied conditional first moment is therefore

$$\widehat{S}_{\ell,+}(\Theta_\omega) = \frac{\displaystyle\sum_{jmt:\ell(j)=\ell} \left[F_\ell\big(\theta_{jmt}(\Theta_\omega)\big) - \frac{h_\ell\big(\theta_{jmt}(\Theta_\omega)\big)}{\theta_{jmt}(\Theta_\omega)} \right]}{\displaystyle\sum_{jmt:\ell(j)=\ell} F_\ell\big(\theta_{jmt}(\Theta_\omega)\big)}. \tag{B.13}$$

The corresponding conditional second raw moment is

$$\widehat{S}_{\ell,+}^{(2)}(\Theta_\omega) = \frac{\displaystyle\sum_{jmt:\ell(j)=\ell} \left[F_\ell\big(\theta_{jmt}(\Theta_\omega)\big) - \frac{2h_\ell\big(\theta_{jmt}(\Theta_\omega)\big)}{\theta_{jmt}(\Theta_\omega)} + \frac{h_\ell^{(2)}\big(\theta_{jmt}(\Theta_\omega)\big)}{\theta_{jmt}(\Theta_\omega)^2} \right]}{\displaystyle\sum_{jmt:\ell(j)=\ell} F_\ell\big(\theta_{jmt}(\Theta_\omega)\big)}. \tag{B.14}$$

The model-implied conditional variance targeted by the third GMM moment is thus

$$\widehat{V}_{\ell,+}(\Theta_\omega) = \widehat{S}_{\ell,+}^{(2)}(\Theta_\omega) - \left[\widehat{S}_{\ell,+}(\Theta_\omega)\right]^2. \tag{B.15}$$

This conditional variance captures both variation in capacity-yield realizations for a given observation and variation in equilibrium demand-to-capacity ratios across observations. See also Appendix B.6.1.

Outer loop: two-step GMM. I stack the nine moment differences

$$g_{N,\ell}(\Theta_\omega) \equiv \begin{pmatrix} \widehat{P}_\ell^{\text{data}} - \widehat{P}_\ell(\Theta_\omega) \\ \widehat{S}_\ell^{\text{data}} - \widehat{S}_\ell(\Theta_\omega) \\ \widehat{V}_{\ell,+}^{\text{data}} - \widehat{V}_{\ell,+}(\Theta_\omega) \end{pmatrix}, \qquad \ell \in \{H, F_N, F_S\}. \tag{B.16}$$

Stacking the three location-specific vectors gives

$$g_N(\Theta_\omega) = \begin{pmatrix} g_{N,H}(\Theta_\omega) \\ g_{N,F_N}(\Theta_\omega) \\ g_{N,F_S}(\Theta_\omega) \end{pmatrix}, \tag{B.17}$$

and the estimator solves

$$\widehat{\Theta}_\omega = \arg\min_{\Theta_\omega} g_N(\Theta_\omega)' W_N g_N(\Theta_\omega). \tag{B.18}$$

I implement two-step GMM. The first step uses a positive-definite diagonal weighting matrix W_N; the second updates W_N using the estimated covariance matrix of the moment conditions.

Inner loop: solving for θ_{jmt}. For each candidate Θ_ω, I solve the equilibrium condition in (FOC 2′),

$$\frac{1}{\theta_{jmt}} = -\left(1 + \widehat{\eta}_{jjmt}\right) \frac{1 - F_\ell(\theta_{jmt})}{h_\ell(\theta_{jmt})}, \qquad \ell = \ell(j), \tag{B.19}$$

for every (j, m, t) observation. The latent demand $\widehat{Q}^D_{jmt}$ and own-price elasticity $\widehat{\eta}_{jjmt}$ are obtained from the first-stage demand estimates and are held fixed within each GMM evaluation. (F_ℓ, f_ℓ) are the CDF and PDF implied by the candidate shock parameters. Solving in $\theta_{jmt}(\Theta_\omega) \in (0, 1]$ rather than in capacity κ_{jmt} is numerically convenient because θ_{jmt} is the sufficient statistic entering all three model moments. I then evaluate $g_N(\Theta_\omega)' W_N g_N(\Theta_\omega)$ and update Θ_ω until the outer GMM problem converges.

B.7.1. Estimation of Fixed-Cost Indices

Step 4 recovers the region-by-period component of location fixed costs that remains after conditioning on the variable-profit differences implied by Steps 1–3. Section 5.3.3 describes the conditional-logit location-choice model used for estimation of these location–period fixed cost indices. This appendix describes the calibration of the coefficient on expected variable profits. Appendix Table B.6 reports the fixed cost estimates and Appendix B.7.2 describes the construction of standard errors.

Conditional on the candidate location-specific expected variable profits $\{\mathbb{E}_\ell[\pi^\ell_{jmt}]\}_{\ell \in \{H, F_N, F_S\}}$, where $\mathbb{E}_\ell[\cdot]$ is the expectation over the capacity-yield shock $\omega \sim F_\ell$ with productivity held fixed, and on the calibrated coefficient β_π, I estimate the eight fixed-cost components in equation (39), normalizing $FC_{F_N,\mathcal{T}} = 0$ in each time block. For each period $\mathcal{T}$, the likelihood first-order conditions satisfy

$$\sum_{j,m,t:\, \mathcal{T}(t)=\mathcal{T}} \left[\mathbf{1}\left\{ \ell^{\text{obs}}_{jmt} = \ell \right\} - \widehat{P}^\ell_{jmt} \right] = 0, \qquad \ell \in \{H, F_S\}. \tag{B.20}$$

Thus, within each period, the estimated fixed-cost components reconcile the observed frequency of production in each region with the location probabilities implied by differences in variable profits. Note that this share fit follows from the likelihood score conditions and is not an independent measure of model fit. The estimates should be interpreted as common location-payoff wedges, relative to Foreign North, that remain after accounting for the model-implied differences in variable profits. They need not coincide with accounting measures of physical plant setup costs.

The coefficient β_π is the inverse scale of the monetary location-profitability shock, with profits measured in millions of dollars. I calibrate it using the dispersion of candidate-location expected-profit differences:

$$\beta_\pi^{\text{bench}} = \frac{2}{\text{median}_{jmt,\, r<s} \left| \mathbb{E}_r[\pi_{jmt}^r] - \mathbb{E}_s[\pi_{jmt}^s] \right|} = 3.934 \quad \text{per million dollars,} \tag{B.21}$$

where $r, s \in \{H, F_N, F_S\}$. The calculation uses the three unordered pairwise profit differences for each observation.

I estimate the fixed-cost parameters conditional on this calibrated coefficient. I report sensitivity to smaller values of β_π, which imply weaker responses of production locations to expected-profit differences. Table B.8 reports these sensitivity results. The intermediate specification sets $\beta_\pi = \beta_\pi^{\text{bench}}/2$, and the low-response specification sets β_π equal to the upper endpoint of its one-sided positive profile-likelihood interval.

B.7.2. Inference for Supply and Location Fixed-Cost Indices

For the parameters estimated in supply Steps 2 and 3, I use a nonparametric bootstrap over NDC 5-4 product–year clusters, conditional on the estimated demand system. In each draw, I resample clusters with replacement, re-estimate the location-specific capacity-yield parameters, reconstruct marginal costs, re-estimate the trade-cost regression, and recompute the pooled moments of $\log \gamma_{jmt}$. The reported standard errors are based on $B = 50$ bootstrap replications.

The Step 4 location fixed-cost indices are not re-estimated within this supply bootstrap. Their reported standard errors come from the cluster-robust sandwich covariance matrix of the conditional-logit likelihood, clustered by NDC-54 product across years. Standard errors for $FC_{F_S,\mathcal{T}} - FC_{H,\mathcal{T}}$ use the full covariance between the two estimated indices. This inference is conditional on the candidate-location profit vectors constructed from Steps 1–3 and on the calibrated benchmark value β_π^{bench}.

When constructing candidate-location profits, I hold the recovered productivity realization $\widehat{\gamma}_{jmt}$ fixed across locations. Candidate locations differ through wages, predicted trade costs, regulatory-cost shifters, and capacity-yield distributions; productivity is not redrawn from a location-specific distribution.

C. Counterfactual Details

C.1. Counterfactual Changes in Realized Consumer Surplus

This appendix provides the algebra underlying the decomposition of counterfactual changes in realized consumer surplus used in Section 6.2, and the valuation of the policies' transfers in the second welfare line.

Recall that realized consumer surplus in market (m,t) is

$$CS^{\text{real}}_{mt}(\delta, s^{\text{real}}_{0mt}) = \underbrace{CS^{\text{logit}}_{mt}(\delta)}_{\text{frictionless logit surplus}} - \underbrace{\left[s^{\text{real}}_{0mt} - s^{D}_{0mt}(\delta)\right]\bar{\mathcal{L}}_{mt}(\delta)}_{\text{shortage penalty}}, \tag{C.1}$$

where $s^{D}_{0mt}(\delta)$ is the latent outside share implied by demand, s^{real}_{0mt} is the realized outside share after supply shocks, and

$$\bar{\mathcal{L}}_{mt}(\delta) = \frac{CS^{\text{logit}}_{mt}(\delta)}{1 - s^{D}_{0mt}(\delta)}$$

is the average shortage loss derived in Eq. (22). In the simulations, prices, capacities and locations determine the distribution of realized supply; conditional on these choices, (C.1) is affine in s^{real}_{0mt}, so evaluating it at the expected realized outside share is the same as taking the expectation of realized surplus over yield shocks. For reshoring, this calculation is performed within each simulated location profile before averaging across profiles, which preserves the nonlinearities in the welfare object.

C.1.1. Decomposition

Let $(\delta_0, s^{\text{real}}_{0,0})$ denote the baseline allocation and $(\delta_1, s^{\text{real}}_{0,1})$ a counterfactual allocation. Using Eq. (C.1), the change in realized consumer surplus can be written as

$$\Delta CS^{\text{real}} = CS^{\text{logit}}(\delta_1) - CS^{\text{logit}}(\delta_0) - \left[s^{\text{real}}_{0,1} - s^{D}_{0}(\delta_1)\right]\bar{\mathcal{L}}(\delta_1) + \left[s^{\text{real}}_{0,0} - s^{D}_{0}(\delta_0)\right]\bar{\mathcal{L}}(\delta_0). \tag{C.2}$$

Adding and subtracting $\left[s^{\text{real}}_{0,1} - s^{D}_{0}(\delta_1)\right]\bar{\mathcal{L}}(\delta_0)$ gives

$$\begin{aligned}\Delta CS^{\text{real}} = &\underbrace{\left(\left[s^{\text{real}}_{0,0} - s^{D}_{0}(\delta_0)\right] - \left[s^{\text{real}}_{0,1} - s^{D}_{0}(\delta_1)\right]\right)\bar{\mathcal{L}}(\delta_0)}_{\Delta CS_{\text{access}}\ \text{(access effect)}} \\ &+ \underbrace{CS^{\text{logit}}(\delta_1) - CS^{\text{logit}}(\delta_0) - \left[s^{\text{real}}_{0,1} - s^{D}_{0}(\delta_1)\right]\left[\bar{\mathcal{L}}(\delta_1) - \bar{\mathcal{L}}(\delta_0)\right]}_{\Delta CS_{\text{price}}\ \text{(price/mean-utility effect)}}.\end{aligned} \tag{C.3}$$

The price component is the change in frictionless logit surplus minus the revaluation of shortage displacement. When counterfactual prices rise, mean utilities fall through the log-price term $-\alpha \log p_{jmt}$, reducing the frictionless logit surplus; prices also change the surplus lost when a buyer is displaced into the outside option, and the second term captures this revaluation, holding fixed the counterfactual share of buyers forced outside. The figures labeled "Access and Price Effects" in Section 6 plot $\Delta CS_{\text{access}}$ and ΔCS_{price} as a percent of baseline realized surplus; they sum to ΔCS^{real} at every policy rate.

C.1.2. Counterbalancing Forces

The condition for a net welfare gain ($\Delta CS^{\text{real}} > 0$) can be expressed as

$$\underbrace{\Big(\Big[s^{\text{real}}_{0,0} - s^{D}_{0}(\delta_0)\Big] - \Big[s^{\text{real}}_{0,1} - s^{D}_{0}(\delta_1)\Big]\Big)\bar{\mathcal{L}}(\delta_0)}_{\text{gain from improved access}} \quad > \quad \underbrace{-\Delta CS_{\text{price}}}_{\text{loss from higher prices}} . \tag{C.4}$$

Thus, realized consumer surplus can rise even when prices increase, provided the reduction in shortage displacement is large enough. The relevant access margin is the change in the realized outside share relative to the latent outside share under the corresponding mean utilities, not the raw change in the realized outside share alone.

C.1.3. Marginal Effect of Reducing Shortages

Holding mean utilities fixed, the marginal welfare effect of the realized outside share is

$$\frac{\partial CS^{\text{real}}_{mt}}{\partial s^{\text{real}}_{0mt}} = -\bar{\mathcal{L}}_{mt}(\delta) < 0. \tag{C.5}$$

Thus, a one percentage point reduction in the realized outside share raises realized consumer surplus by $0.01 \times \bar{\mathcal{L}}_{mt}(\delta)$ in expected-utility units. The gain is larger in markets with a low latent outside share because the average shortage loss,

$$\bar{\mathcal{L}}_{mt}(\delta) = \frac{-\ln s^{D}_{0mt}(\delta)}{1 - s^{D}_{0mt}(\delta)},$$

is increasing as the outside option becomes less attractive.[3]

C.1.4. Realized Consumer Surplus at the Price Paid

The second welfare line of Section 6.2 values the policy's transfer as a proportional change in the prices buyers pay in the same market, on the equilibrium allocation. In market (m, t), let $B_{mt} = \sum_j p_{jmt}\, \mathbb{E}[q^{\text{real}}_{jmt}]$ be spending at contract prices and T_{mt} the net transfer paid by buyers, so that the price paid scales every inside contract price by $\lambda_{mt} = 1 + T_{mt}/B_{mt}$. Under reshoring the transfer is the subsidy outlay $s_{\text{reshore}} \sum_j \mathrm{R}^{H}_{jmt} \mathbb{1}\{\ell_j = H\}$, and both quantities are expectations within a location profile. Under the penalty it is minus the fines $\sum_j \mathbb{E}[\mathcal{P}_{jmt}]$, returned in full to buyers. Under shortage-contingent pricing buyers pay $(1 + \psi^{\text{spot}})$ times the contract price on every unit delivered in a shortage state, so $\lambda_{mt} = 1 + \psi^{\text{spot}}$ in those states and $\lambda_{mt} = 1$ otherwise. Because price enters mean utility through $-\alpha \log p$, the latent outside share at the price paid is

$$s^{D,\text{paid}}_{0mt} = \frac{1}{1 + \lambda_{mt}^{-\alpha}\left(1/s^{D}_{0mt} - 1\right)}, \tag{C.6}$$

and realized consumer surplus at the price paid is the realized surplus (C.1) with $s^{D,\text{paid}}_{0mt}$ in place of s^{D}_{0mt} and the realized outside share held at its equilibrium value,

$$CS^{\text{real}}_{mt}\big(\delta, s^{\text{real}}_{0mt}\big)\Big|_{\text{paid}} = -\log s^{D,\text{paid}}_{0mt}\, \frac{1 - s^{\text{real}}_{0mt}}{1 - s^{D,\text{paid}}_{0mt}}. \tag{C.7}$$

[3] This effect is non-linear in the essentiality of the drug: for essential markets with $s^{D}_{0} \to 0$, the average shortage loss becomes very large, so even small reductions in shortages can generate large welfare gains.

I hold the counterfactual allocation fixed and value the transfer using the proportional-price convention above. The calculation is performed state by state for shortage-contingent pricing and within each location profile for reshoring before aggregation.

For interpretation, I also solve for the uniform factor $\bar{\lambda}$ by which baseline contract prices would have to change for buyers, at their baseline allocation, to attain the counterfactual surplus index. I aggregate with the same market-size weights and report $\bar{\lambda} - 1$ as the price-equivalent change. This expresses the comparison in price units, rather than converting the surplus index into dollars or measuring total social welfare.

Table C.1 reports the transfers themselves and manufacturers' operating profit at the policy points of the main comparison. Fines are paid by manufacturers and, in the second welfare line, received by buyers; premia and the subsidy flow to manufacturers, paid by buyers in the second line and by the Treasury or a third party in the first. The subsidy is the one instrument with a public outlay.

Table C.1: **Transfers and Manufacturers' Profits under the Three Policies**

	Reshoring		**Shortage-contingent**		**Penalty**	
	$s = 0.105$	$s = 0.20$	$\psi = 0.25$	$\psi = 0.50$	$\psi = 0.50$	$\psi = 1$
Transfer per year, billion dollars						
Paid by buyers at the price paid	1.7	11.7	12.7	26.2	−2.8	−4.9
Received by manufacturers	1.7	11.7	12.7	26.2	−2.8	−4.9
Public outlay under the first welfare line	1.7	11.7	0	0	−2.8	−4.9
Manufacturers, percent change						
Expected operating profit after transfers	+5.2	+47.3	+30.5	+62.7	−1.1	−1.9

Notes: Transfers are averages per year over 2017 to 2019. A negative buyer entry is a rebate of fines; a negative public entry is fines received by the Treasury under the first welfare line, in which fines are not rebated, the subsidy is a public outlay and premia are paid by a third party. Operating profit is after fines, premia and subsidy, over all modeled manufacturers, foreign ones included, before location fixed costs.

C.2. Reshoring Subsidies: Details

Political and Institutional Context of Reshoring Policies. From 1976 to 2006, the Internal Revenue Code exempted from taxation corporate income generated in U.S. territories. The tax break was estimated to have saved drug companies the 2022 equivalent of $235 million per year, or 70 cents per American (Washington Post, January 2023).

Since 2020, the U.S. and several European governments have started a push towards tax incentives that would encourage reshoring. In the U.S., key legislative proposals, such as the *Domestic Medical and Drug Manufacturing Tax Credit*, discussed offering manufacturers a tax credit of 10.5% of the net income from the sale of critical medical products produced domestically. If passed, this would effectively halve their corporate tax. In 2021, Joe Biden's *American Jobs Plan* pledged an investment of $300 billion towards U.S. manufacturing, which includes a substantial allocation for reinforcing supply chains and ensuring pandemic readiness by reshoring Active Pharmaceutical Ingredients (APIs).[4] In France, Presi-

[4]The White House's June 2021 100-day review of America's Supply Chains: "Building Resilient Supply Chains, Revitaliz-

dent Emmanuel Macron publicly announced in June 2023 that France would reshore the production of 50 essential medicines, backed by significant public funding ($173 million in public money will go to support eight of the new reshoring projects, and companies reshoring to France can apply for a share of an additional 50 million euros in funding).

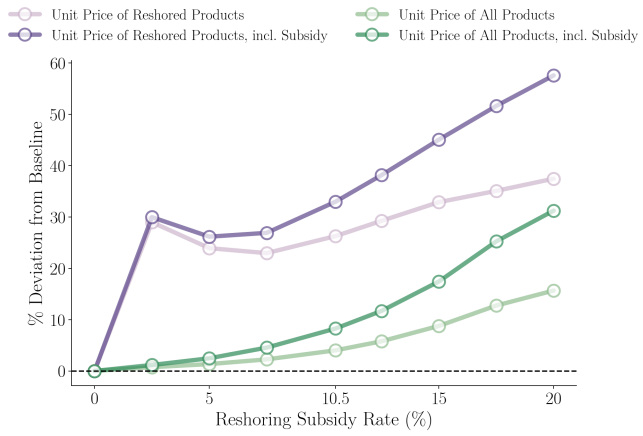
In practice, pre-2006, a large share of U.S. drug manufacturing plants were located in Puerto Rico. As a result, Puerto Rico today has idle plants, a trained pharmaceutical workforce and relatively low wages compared to the mainland United States, leading many voices to call for a "near-shoring" to Puerto Rico specifically. This comes with limitations. First, Puerto Rico is subject to natural disasters: in September 2017, Hurricane Maria caused more than 3,000 fatalities and millions of dollars in damage, destroyed most of the drug manufacturing plants that remained on the island and pushed firms to offshore production entirely.[5] Second, recent power failures underlined the fragility of the electricity infrastructure, which makes it risky to install manufacturing plants on the island. While reshoring to Puerto Rico may help in reducing the frequency of shocks while keeping relative costs lower than in mainland U.S., it may also result in supply disruptions that are bigger in magnitude.

Additional outcomes. Figure C.1 reports the outcomes that complement Figure 6: prices averaged over all products, capacity and buffer, the decomposition of realized consumer surplus, and the relocation of production across regions. Averaged over all products, with products weighted equally, prices rise by 4% at the benchmark rate and 16% at a 20% subsidy, against 26 and 37% for the products that move (panel C.1a). Planned capacity falls slightly at low rates and rises by 14% at the benchmark rate, while the buffer over expected demand is a fifth thinner at low rates and back at its baseline level at the benchmark rate (panel C.1b). The decomposition in panel C.1c attributes the whole change in realized consumer surplus to the price effect of equation (41), with the access effect slightly negative at every rate up to 15%. The expected share of product lines made in the United States rises from 12% in the baseline to 25% at the benchmark rate and 52% at a 20% subsidy (panel C.1d); the relocated lines come from other high-income locations first, whose share falls by a quarter at the benchmark rate and by seven tenths at the largest, and from Asia, whose share falls by 12 and 36%. A larger Home revenue base gives a product a larger increase in its location payoff at a given subsidy rate, but the relocation response also depends on candidate profits and location-cost wedges.

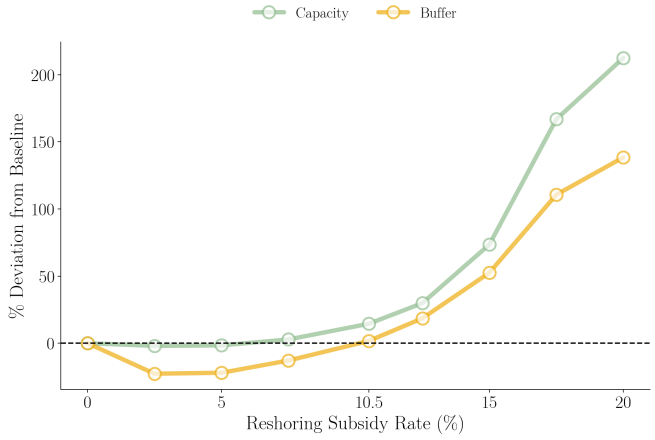
Prices of reshored products by region of origin. The average price change of the products that relocate hides two opposite movements, shown in Figure C.2. Products that move from Asia nearly double their price, by 87% at the benchmark rate, because the model's unit cost at a U.S. plant is 2.7 times the cost at an Asian plant; products that move from other high-income countries cut theirs by more than half, 57% at the benchmark rate, because the U.S. cost is a quarter of theirs. Products from Asia make up about three fifths of relocated products at the benchmark rate and nearly three quarters at a 20% subsidy, so the average rises with the rate, from 24% at 5% to 37% at 20%. The few large products

ing American Manufacturing and Fostering Broad-Based Growth".

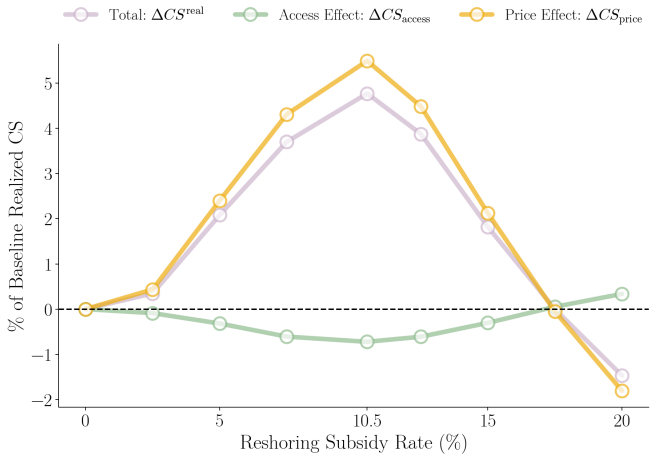
[5] Several pharmaceutical companies with manufacturing and distribution operations on the island were hit hard, including Pfizer, Amgen and Bristol-Myers Squibb.



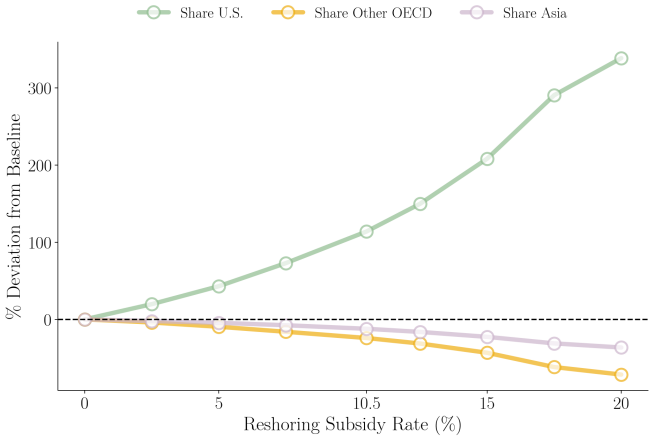
(a) *Unit Prices of Reshored Products and of All Products*



(b) *Capacity and Buffer*



(c) *Access and Price Effects on Realized Consumer Surplus*



(d) *Share of Product Lines by Production Region*

Figure C.1: **Reshoring Subsidy: Prices, Capacity, Welfare Decomposition and Location Shares**

Notes: Percentage deviations from the baseline; the x-axis reports the reshoring subsidy rate in percent of Home revenue. Panel C.1a: average unit price of the products that relocate to the U.S. (purple) and of all products (green), each with and without the subsidy's financing charged to buyers as a proportional surcharge in the same market; products weighted equally. Panel C.1b: planned capacity and buffer over expected demand (capacity minus latent demand). Panel C.1c: change in realized consumer surplus as a percent of its baseline level, split into the access effect $\Delta CS_{\text{access}}$ and the price effect ΔCS_{price} of equation (41); the two components sum to the total at every rate. Panel C.1d: expected share of product lines produced in the U.S., in other high-income locations and in Asia, each relative to its own baseline share (12, 24 and 64%). 2017 to 2019, weighted by market size.

moving from other high-income countries, whose prices fall, are the source of the realized surplus gain in Figure 6.

C.3. Shortage-Contingent Pricing: Details

This appendix derives the closed-form expression for the shortage-state term $\mathbb{E}[q^{\text{real}}_{jmt}\, I^{\text{short}}_{mt}]$ that enters the manufacturer's Step-2 problem under equation (Step-2 Spot). Recall F_ℓ and f_ℓ are the CDF and density of ω_{jmt}, $\theta_{jmt} \equiv Q^D_{jmt}(\delta)/\kappa_{jmt}$ and $h_\ell(\theta) \equiv \int_0^\theta \omega f_\ell(\omega)\, d\omega$. Supplier j is short when $\omega_{jmt} < \theta_{jmt}$, which happens with probability $F_{\ell(j)}(\theta_{jmt})$, and the market is short when at least one supplier is.

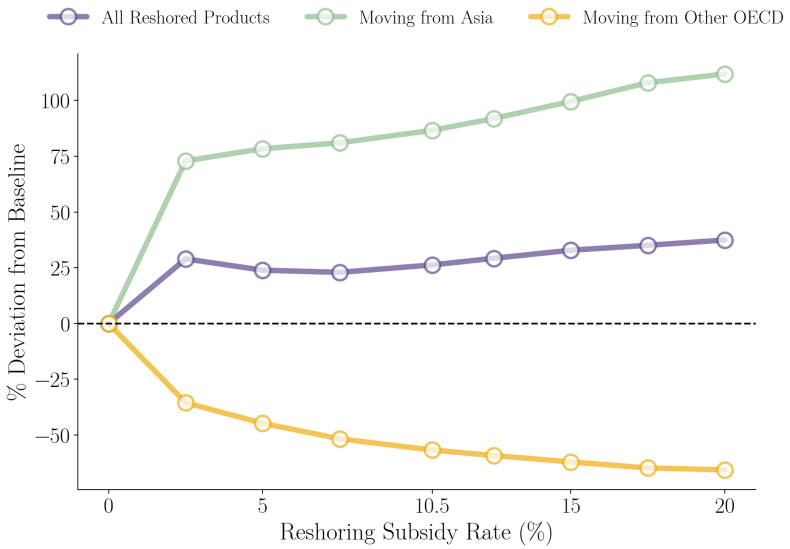


Figure C.2: **Reshoring Subsidy: Unit Prices of Reshored Products by Region of Origin**

Notes: Average change in the contract price of the products that relocate to the U.S., in percent of their own baseline price, for all relocated products and separately for those moving from Asia and from other high-income countries, products weighted equally as in Figure 6; the x-axis reports the reshoring subsidy rate in percent of Home revenue. Averages are taken over the expected set of relocated products at each rate. 2017 to 2019.

Assuming yield shocks are independent across plants conditional on location, the event that the market is short splits into the states in which supplier j is itself short and the states in which it delivers in full while at least one rival is short:

$$\mathbb{E}\left[q_{jmt}^{\text{real}}\, I_{mt}^{\text{short}}\right] = \underbrace{\int_0^{\theta_{jmt}} \kappa_{jmt}\, \omega\, f_{\ell(j)}(\omega)\, d\omega}_{\text{own shortfall}} + \underbrace{Q_{jmt}^{D}(\delta)\left[1 - F_{\ell(j)}(\theta_{jmt})\right]\left[1 - \prod_{k \neq j}\left(1 - F_{\ell(k)}(\theta_{kmt})\right)\right]}_{\text{rival shortfall}}$$

$$= \kappa_{jmt}\, h_{\ell(j)}(\theta_{jmt}) + Q_{jmt}^{D}(\delta)\left[1 - F_{\ell(j)}(\theta_{jmt})\right]\left[1 - \prod_{k \neq j}\left(1 - F_{\ell(k)}(\theta_{kmt})\right)\right]. \tag{C.8}$$

The first term is expected delivery when supplier j is itself short. The second is full delivery by j when at least one rival is short. At an interior capacity choice, holding prices and rivals' choices fixed, the derivative of expected premium-bearing deliveries is

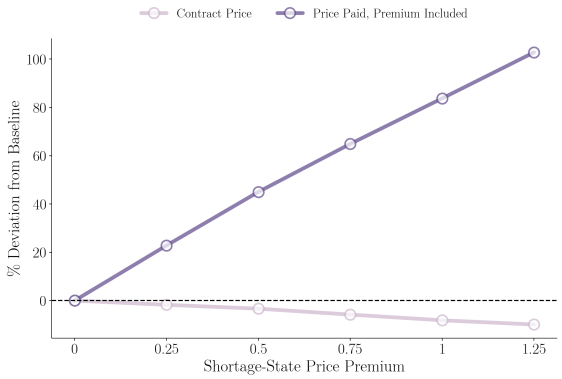
$$\frac{\partial \mathbb{E}[q_{jmt}^{\text{real}}\, I_{mt}^{\text{short}}]}{\partial \kappa_{jmt}} = h_{\ell(j)}(\theta_{jmt}) - \theta_{jmt}^{2} f_{\ell(j)}(\theta_{jmt}) \prod_{k \neq j}\left[1 - F_{\ell(k)}(\theta_{kmt})\right].$$

The first term captures additional deliveries; the second captures the loss of premium eligibility when added capacity eliminates the only shortage in the market. The net effect need not be positive. I evaluate equation (C.8) in the price–capacity problem at observed production locations.

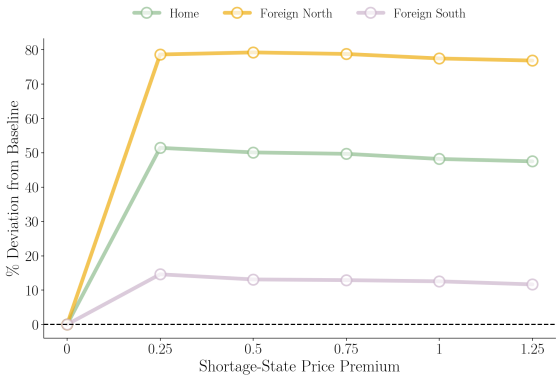
The premia are revenue to manufacturers, whose expected operating profit rises by 30% at a premium of one quarter and by 63% at a premium of one half (Table C.1).

Additional outcomes. Figure C.3 reports the outcomes that complement Figure 7: unit prices with and without the premium, the probability of a shortage by production region, the share of suppliers

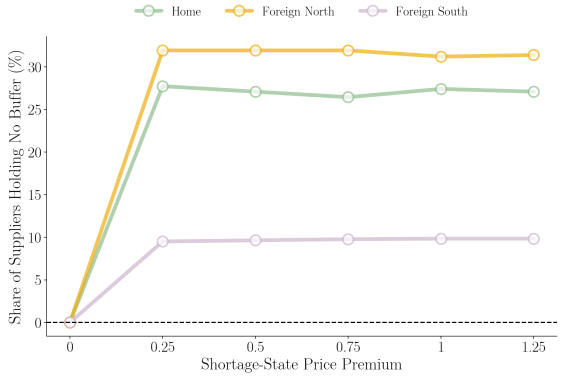
holding no buffer, and the capacity utilization of the suppliers that keep one. Figure C.4 decomposes the change in realized consumer surplus into its access and price effects.



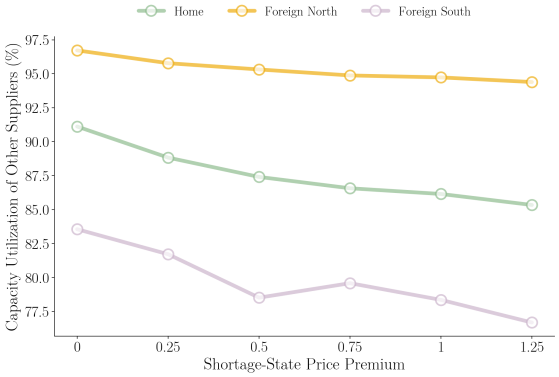
(a) *Unit Prices*



(b) *Probability of Shortage by Production Region*



(c) *Suppliers Holding No Buffer by Production Region*



(d) *Utilization of Suppliers Holding a Buffer, by Production Region*

Figure C.3: **Shortage-Contingent Pricing: Prices, Shortages and Buffers by Production Region**

Notes: Percentage deviations from the baseline unless noted; the x-axis displays the shortage-state premium ψ^{spot} from equation (Eq. Spot Price); production locations are held at their observed values. Panel C.3a: average contract price and average price paid per delivered unit with the premium included, weighted by market size. Panel C.3b: supplier-level probability of a shortage, averaged across suppliers within each production region. Panel C.3c: share of suppliers whose planned capacity equals their contracted demand, so that they are short whenever their yield falls below one, within each production region (percent of suppliers, not a deviation). Panel C.3d: contracted demand over planned capacity, averaged across the suppliers of each production region whose capacity exceeds their contracted demand (percent, not a deviation); a lower value is a larger buffer. Home is the U.S., Foreign North other high-income locations, Foreign South Asia. 2017 to 2019.

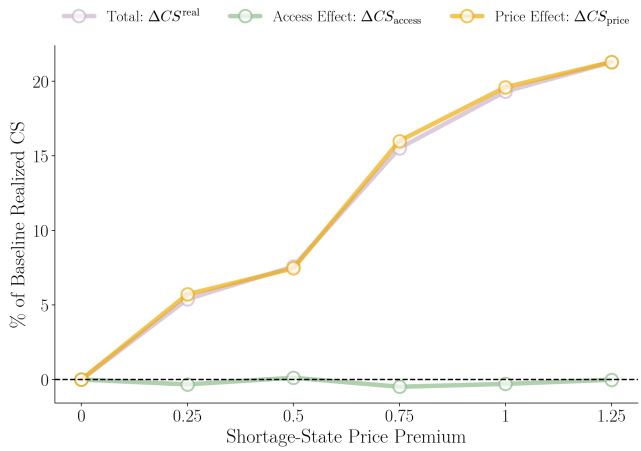


Figure C.4: **Shortage-Contingent Pricing: Access and Price Effects on Realized Consumer Surplus**

Notes: Change in realized consumer surplus at contract prices as a percent of its baseline level, split into the access effect $\Delta CS_{\text{access}}$ and the price effect ΔCS_{price} of equation (41); the two components sum to the total at every premium. The x-axis displays the shortage-state premium ψ^{spot} from equation (Eq. Spot Price); production locations are held at their observed values. 2017 to 2019, weighted by market size.

C.4. Failure-to-Supply Penalties: Details

Expected shortage quantity and its derivative

Recall that latent demand $Q^D_{jmt}(\delta)$ does not depend on the yield shock ω_{jmt} and that the *realized* shortage quantity defined in equation (43) is $Q^{\text{short}}_{jmt} = \max\{Q^D_{jmt}(\delta) - \kappa_{jmt}\,\omega_{jmt}, 0\}$. Taking expectations over ω_{jmt} yields the *expected* shortage quantity in equation (45):

$$\mathbb{E}_{\omega\sim F_\ell}\left[Q^{\text{short}}_{jmt}\right] = \int_0^{\theta_{jmt}} \left(Q^D_{jmt}(\delta) - \kappa_{jmt}\omega\right) f_\ell(\omega)\, d\omega = Q^D_{jmt}(\delta) F_\ell(\theta_{jmt}) - \kappa_{jmt} h_\ell(\theta_{jmt}), \tag{C.9}$$

where $\theta_{jmt} \equiv Q^D_{jmt}(\delta)/\kappa_{jmt}$, F_ℓ is the CDF of ω_ℓ, and $h_\ell(\theta) \equiv \int_0^\theta \omega f_\ell(\omega)\, d\omega$.

Using $d\theta_{jmt}/d\kappa_{jmt} = -\theta_{jmt}/\kappa_{jmt}$, $F'_\ell(\theta) = f_\ell(\theta)$ and $h'_\ell(\theta) = \theta f_\ell(\theta)$, the derivative of the expected shortage with respect to capacity is

$$\begin{aligned}
\frac{\partial}{\partial \kappa_{jmt}} \mathbb{E}_{\omega\sim F_\ell}\left[Q^{\text{short}}_{jmt}\right] &= Q^D_{jmt}(\delta) f_\ell(\theta_{jmt}) \frac{\partial \theta_{jmt}}{\partial \kappa_{jmt}} - h_\ell(\theta_{jmt}) - \kappa_{jmt} h'_\ell(\theta_{jmt}) \frac{\partial \theta_{jmt}}{\partial \kappa_{jmt}} \\
&= -\theta^2_{jmt} f_\ell(\theta_{jmt}) - h_\ell(\theta_{jmt}) + \theta^2_{jmt} f_\ell(\theta_{jmt}) \\
&= -h_\ell(\theta_{jmt}) < 0.
\end{aligned} \tag{C.10}$$

An extra unit of capacity reduces expected shortfall by $h_\ell(\theta_{jmt})$.

Capacity first-order condition with penalties

Expected profit with penalties, equation (44), is the baseline expected profit minus $\psi^{\text{pen}}\, p_{jmt}\, \mathbb{E}_{\omega\sim F_\ell}[Q^{\text{short}}_{jmt}]$. Differentiating with respect to κ_{jmt} and using (C.10),

$$\frac{\partial}{\partial \kappa_{jmt}} \mathbb{E}_{\omega\sim F_\ell}\left[\pi^{\ell,\text{pen}}_{jmt}\right] = \frac{\partial}{\partial \kappa_{jmt}} \mathbb{E}_{\omega\sim F_\ell}\left[\pi^{\ell}_{jmt}\right] + \psi^{\text{pen}}\, p_{jmt}\, h_\ell(\theta_{jmt}). \tag{C.11}$$

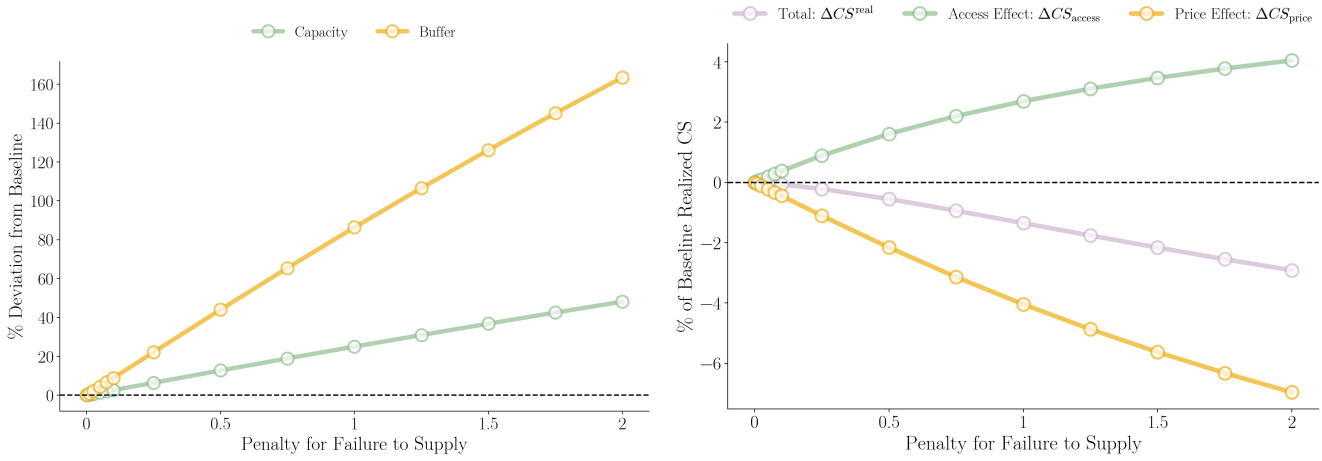
In the baseline model the capacity first-order condition is $p_{jmt} h_\ell(\theta_{jmt}) = c^\ell_{jmt}$. With penalties it becomes

$$p_{jmt} h_\ell(\theta_{jmt}) \left(1 + \psi^{\text{pen}}\right) = c^\ell_{jmt}. \tag{C.12}$$

For given $(p_{jmt}, Q^D_{jmt}, F_\ell)$ and $\psi^{\text{pen}} > 0$, condition (C.12) implies a lower θ_{jmt} (higher κ_{jmt}) than in the baseline, because $h_\ell(\theta)$ is increasing in θ and θ_{jmt} is decreasing in κ_{jmt}. Intuitively, the penalty increases the marginal benefit of capacity by $\psi^{\text{pen}}\, p_{jmt} h_\ell(\theta_{jmt})$, so firms optimally choose larger capacity buffers and reduce expected shortfalls. The rebate of fines to buyers in the second welfare line is applied after equilibrium behavior is determined; it changes the incidence of the policy, not suppliers' incentives.

Manufacturers' operating profit after fines falls by 1% at a rate of one half and 2% at a rate of one, while profit before fines rises, because the repriced buffer is paid for by the buyers who stay (Table C.1).

Additional outcomes. Figure C.5 reports the capacity response and the decomposition of realized consumer surplus that complement Figure 8.



(a) *Capacity and Buffer* (b) *Access and Price Effects on Realized Consumer Surplus*

Figure C.5: **Failure-to-Supply Penalties: Capacity and Welfare Decomposition**

Notes: Percentage deviations from the baseline; the x-axis reports the penalty rate ψ^{pen} from equation (43); production locations are held at their observed values. Panel C.5a: planned capacity and buffer over expected demand (capacity minus latent demand). Panel C.5b: change in realized consumer surplus at contract prices as a percent of its baseline level, split into the access effect $\Delta CS_{\text{access}}$ and the price effect ΔCS_{price} of equation (41); the two components sum to the total at every rate. 2017 to 2019, weighted by market size.

C.5. What It Costs to Reduce Shortages

Table C.2 reports the policy rates associated with selected shortage reductions (Panel A) and with selected annual transfer amounts (Panel B). It also reports prices, capacity buffers, realized consumer surplus, and transfers at those rates. Intermediate rates are interpolated between simulated values. Transfers differ in direction and incidence across policies, so Panel B is not a comparison at equal fiscal or resource cost.

Table C.2: **What It Costs to Reduce Shortages, and What a Given Transfer Buys**

	Rate	Shortages, %		Prices and capacity, %		Realized CS, %		Transfer
		Excess demand	Pr(short)	Contract price	Buffer	Contract prices	Price paid	per year ($bn)
Panel A. Failure-to-supply penalty rate required for a given cut								
Share of excess demand, −10%	0.36	−10.0	−21.2	+6.1	+32	−0.4	+0.5	2.1
Share of excess demand, −25%	1.07	−25.0	−43.6	+19.2	+92	−1.5	+0.7	5.2
Shortage probability, −10%	0.14	−4.3	−10.0	+2.4	+13	−0.1	+0.3	0.9
Shortage probability, −25%	0.44	−12.2	−25.0	+7.6	+39	−0.5	+0.5	2.5
Shortage probability, −50%	1.41	−30.3	−50.0	+25.7	+119	−2.0	+0.6	6.3
Share of excess demand, −50%	not reached at any rate up to two (largest cut 38%)							
Reshoring subsidy, any target	not reached: excess demand rises at every rate; the shortage probability falls by at most 0.3%							
Shortage-contingent pricing	not reached at any premium: the shortage probability rises at every positive rate							
Panel B. Comparisons at common annual transfer amounts								
Penalty, 1 billion	0.17	−4.8	−11.2	+2.8	+15	−0.1	+0.3	1.0
Penalty, 2.5 billion	0.44	−12.2	−25.0	+7.6	+39	−0.5	+0.5	2.5
Penalty, 5 billion	1.02	−24.2	−42.7	+18.3	+88	−1.4	+0.7	5.0
Reshoring subsidy, 1 billion	8%	+2.1	−0.3	+23.2§		+3.8	+2.9	1.0
Reshoring subsidy, 2.5 billion	12%	+7.5	−0.1	+28.5§		+4.1	+1.7	2.5
Reshoring subsidy, 5 billion	15%	+15.9	+0.9	+32.4§		+2.1	−1.2	5.0
Shortage-contingent pricing	a premium of 0.25 already costs 12.7 billion dollars a year and raises the probability of a shortage							

Notes: Percentage deviations from the baseline over 2017 to 2019, weighted by market size, at the policy rate that delivers the target (Panel A) or the annual transfer amount (Panel B); rates between simulated values are linearly interpolated. Shortages: market-level expected share of excess demand and supplier-level probability of a shortage. Contract price is the demand-weighted contract price, or §the average contract price of the products that relocate; the buffer is capacity over expected demand. Realized consumer surplus at the price paid returns the fines to buyers or charges them the subsidy. The transfer is fines collected or subsidy paid out, per year. Equal transfer amounts do not imply equal fiscal or resource costs: fines are paid by suppliers, whereas subsidies are paid to suppliers.

D. Market Structure: Institutions and Prices

D.1. How Procurement Auctions Work

This account draws on health policy reports, documents describing GPO procurement, and my interview with Todd Ebert, then president of the Healthcare Supply Chain Association, on February 22, 2022. The discussion concerns contracting practices in the market studied here.

Auctions and awards. GPOs negotiate pharmaceutical contracts on behalf of member hospitals through competitive procurement auctions (requests for proposals): GPOs solicit supplier bids, evaluate proposals, and negotiate contract terms before awarding contracts. Hospitals generally purchase the drugs through wholesalers at GPO-negotiated prices. Price is a central criterion among suppliers with the required marketing authorization. Sole-source awards were widely used during the study period, although some contracts admit multiple suppliers. A sole-source award concerns a particular purchasing arrangement, not the entire national market. Different purchasing organizations can select different labelers, and one labeler can sell to several buyers.

Figure D.1 depicts the procurement auction.

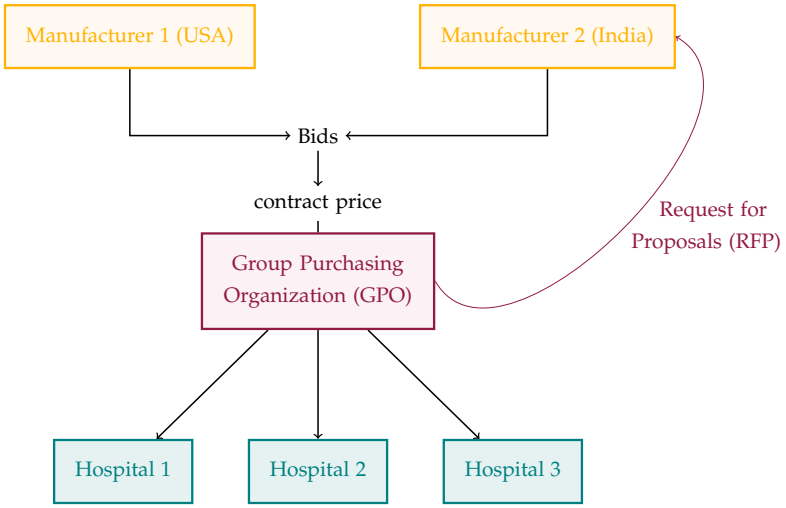


Figure D.1: **U.S. Procurement System: Procurement Auctions**

Notes: This figure depicts hospitals' drug procurement in the U.S. GPOs negotiate pharmaceutical contracts on behalf of member hospitals through procurement auctions: they solicit supplier bids, evaluate proposals, and negotiate contract terms before awarding contracts. Hospitals generally purchase the drugs through wholesalers at GPO-negotiated prices. Drug labelers, the firms holding a drug's license, respond to a request for proposals when they have the capacity to manufacture the drug, in plants they own or outsource, in the U.S. or abroad. Contract prices are not observed in the data: the Red Book reports published list prices (Wholesale Acquisition Cost and Average Wholesale Price), which are benchmarks rather than transaction prices.

Contract terms. Contracts typically last one to two years and require suppliers to honor the negotiated price. GPOs work with suppliers to restore delivery under existing terms, with alternative sourcing used when those arrangements fail. Individual hospitals can purchase outside a GPO agreement, and contracts do not always guarantee purchase volumes. Fixed prices and guaranteed quantities are

therefore distinct contractual provisions. Wholesale distributors handle much of the physical delivery; a GPO's role in negotiating purchasing terms should not be confused with wholesale distribution (Haninger, Jessup, and Koehler, 2011).

Failure-to-supply provisions. Some contracts require a supplier that fails to deliver to reimburse the additional cost of buying the drug elsewhere. Such provisions offer limited protection when alternative supply is unavailable. The ASPE report describes clauses that exclude search costs and restrict the period over which replacement purchases are compensated (Haninger, Jessup, and Koehler, 2011). Ebert reported that failure-to-supply clauses are rarely invoked in practice. These provisions do not make the supplier bear the full cost of unmet treatment needs. Suppliers nevertheless forgo revenue on units they fail to deliver. The penalty counterfactual in Section 6 adds a contractual cost of missing units to this existing lost-sales incentive.

D.2. Which Drugs Go Short and Why

The UUDIS records identify drug-level shortage episodes and, in the detailed availability reports, the affected labeler-products. I use reported onset and resolution dates to measure incidence and duration. Episodes that remain active at the end of a sample are right-censored at that date rather than classified as resolved. These administrative records do not quantify missing units. The model and its estimation provide that quantitative measure; the mapping from delivery shortfalls to the fraction of the contract period in shortage is derived in Appendix B.6.1.

Figure D.2a compares the share of drugs ever reported in shortage across categories in the matched Orange Book and UUDIS data. The denominator is the set of drugs in each category, not prescriptions or sales. Figure D.3 reports causes among episodes with an identified reason. Manufacturing-quality problems account for the largest share. A cause is not reported for 40 percent of episodes.

(a) Share of U.S.-Marketed Drugs in Shortage, by Drug Category

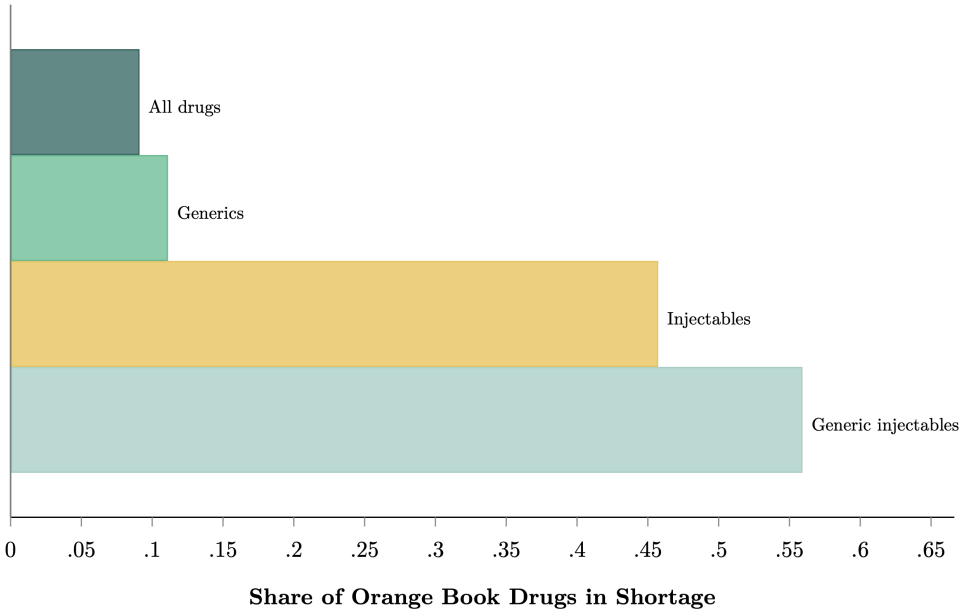


Figure D.2: **Size of Affected Markets**

Notes: Panel D.2a computes the share of all U.S.-marketed drugs that went short, across different drug categories. The list of all U.S.-marketed drugs is from the FDA's Orange Book. While shortage events only affect 9% of all U.S.-marketed drugs and 12% of all generics over the period 2004-2019, the problem becomes more acute if we focus on smaller market segments. 46% of sterile injectable drug products (defined at the molecule-form level) marketed in the U.S. went short at least once over the period, and this number increases to 56% if we focus on the generic versions of injectable drugs. *Source, details:* I use both the January 2020 version of the Orange Book and historical Orange Book records to get the list of marketed drugs in each year over the period 2004-2019. Inside the Orange Book data, I extract the list of i) all generic drugs, ii) all sterile injectable drugs, iii) all generic, sterile injectable drugs marketed in the U.S. I then match these lists to the drugs belonging to the exact same category within the UUDIS shortages data. I then compute, within each category of drug, the share of the market affected by shortages over the period 2004-2019.

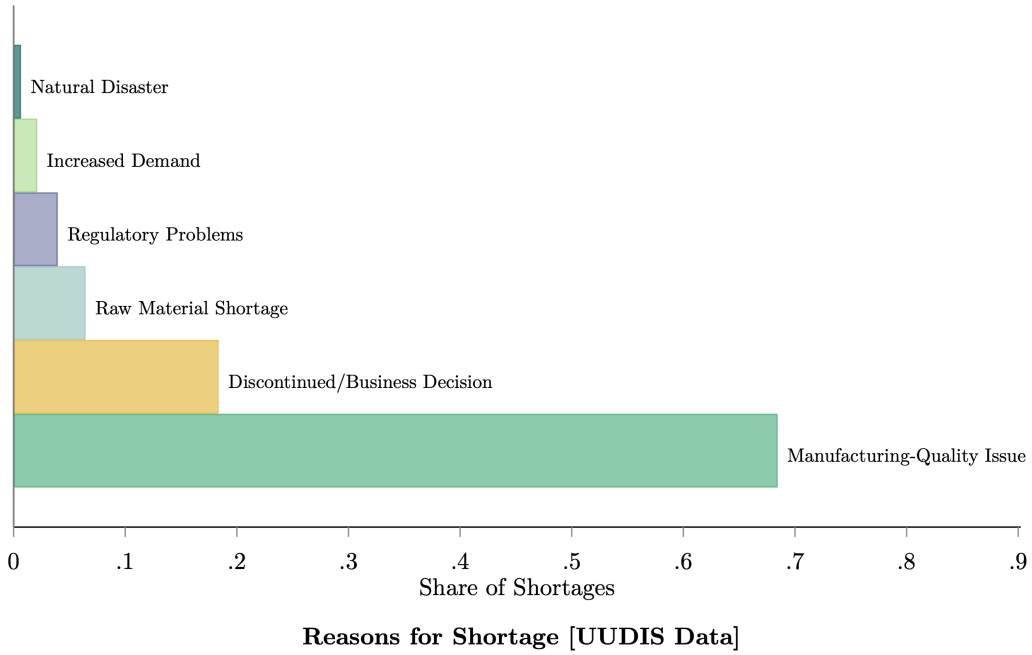


Figure D.3: Reasons for U.S. Shortages

Notes: Drug shortages data were provided by Erin Fox from the University of Utah Drug Information Service (UUDIS). Reasons for shortages are provided to the FDA or the American Society of Health-System Pharmacists (ASHP) by pharmaceutical companies or distributors directly, and are verified by the UUDIS. It is not compulsory to report a cause for a shortage and 40% of shortage events do not report a cause. Shortage reasons are generally very detailed in the data and were aggregated into 6 broader categories. Among those shortages that report a cause, 68% are driven by "Manufacturing-quality issues, 18% are due to "discontinuations or business decisions", 6% are due to "raw material shortages", 4% to "regulatory problems", 3% to "increased demand" and 1% to "natural disaster"

D.3. Prices around Shortage Onset

Figure D.4 plots price profiles around shortage onset using historical Red Book quotations. The two series are Wholesale Acquisition Cost (WAC) and Average Wholesale Price (AWP). They are published price benchmarks, not observed contracting prices between a GPO and a labeler or between a hospital and a GPO. AWP in particular need not equal an actual purchase price (GAO, 2005). The sales-based prices used in structural estimation come from IQVIA, as described in Appendix F.1.

Published prices rise at later horizons following shortage onset. The profiles are consistent with gradual repricing rather than an immediate adjustment to the supply disruption. They do not directly date the renegotiation of individual procurement contracts.

Table D.1 reports NDC-level shortage characteristics. Table D.2 reports the quarterly price statistics, retaining the separate AWP, WAC, and IQVIA measures.

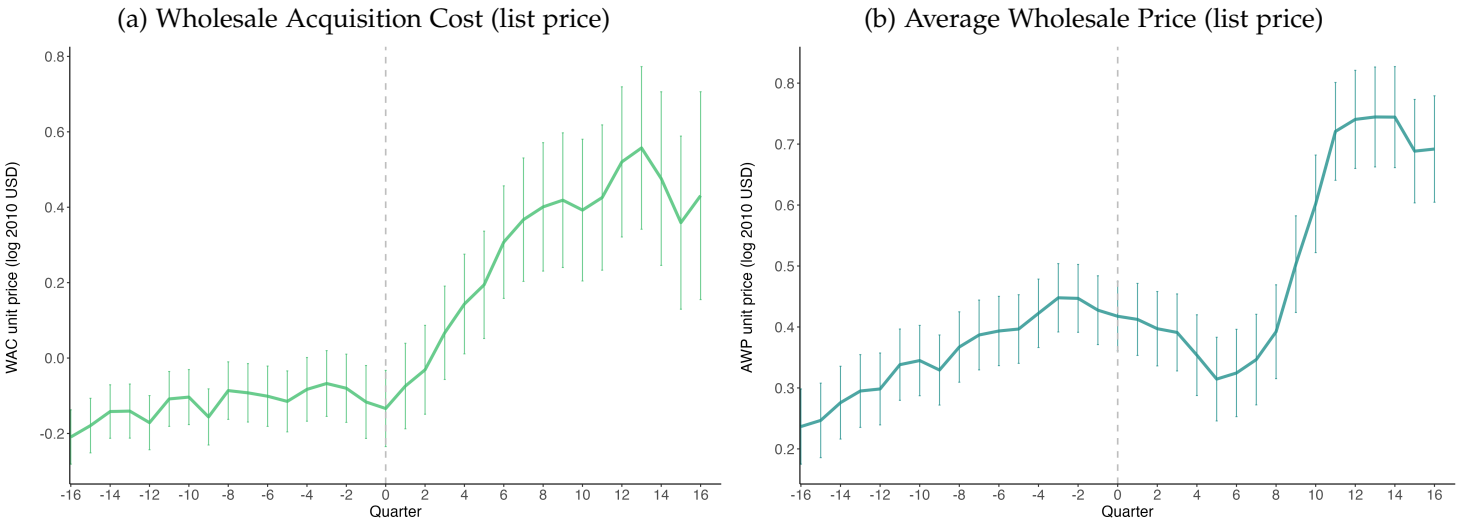


Figure D.4: **Price Variations Around the Start of Shortage Spells**

Notes: Balanced panel of generic injectable drugs in shortage from 2002 to 2019 from the UUDIS. AWP and WAC are published list prices from IBM Micromedex ("RedBook"), not transaction prices. All prices are adjusted for inflation using CPI (base year: 2010). Prices do adjust upwards post-shortage, even if the current structure of contracts between buyer GPOS and suppliers slow down adjustments. For more discussions on pricing, see Appendix D.3.

Table D.1: **NDC-Level Shortages**

	N	p50	mean	sd	min	max	p25	p75
Merger	18,966	0	0.174	0.379	0	1	0	0
Discontinued	18,966	0	0.425	0.494	0	1	0	1
Number of Shortages by NDC (5-4-2)	18,966	3	6.52	7.45	1	37	1	8

Notes: NDC level shortages (5-4-2 format), cross-section from January 2002 to April 2020. Data is from the UUDIS. Information on mergers and discontinuation of products has been extracted from Word Files by the author using text mining techniques

Table D.2: **Quarterly Unit Price of Drugs from the Shortage List**

	Price type	p50	mean	sd	min	p25	p75	max	N
Brand	AWP	1.2	16	46	0.0043	0.14	7.5	557	63,819
	WAC	1.5	14	40	0.0035	0.19	7.7	479	40,067
	Transaction price (IQVIA)	11	82	551	0.28	2.8	39	17,444	19,324
Generic	AWP	1.4	8.1	29	0.0049	0.52	4.3	461	121,780
	WAC	0.67	6.2	22	0.0041	0.18	3.2	351	58,554
	Transaction price (IQVIA)	4.6	23	84	0.11	1.9	14	1,789	38,928
Non-Parenteral	AWP	1.3	7.7	29	0.0056	0.46	3.6	509	123,307
	WAC	0.65	7.5	27	0.0053	0.17	3.1	408	54,909
	Transaction price (IQVIA)	9.7	209	702	0.17	3.5	52	6,233	4,814
Parenteral	AWP	1.5	17	49	0.0043	0.29	9.4	557	143,677
	WAC	1.2	14	38	0.0035	0.21	7.6	479	76,968
	Transaction price (IQVIA)	6.4	52	320	0.11	2.3	23	17,444	65,190
Outside of shortage	AWP	1.3	13	42	0.0043	0.38	5.4	557	225,605
	WAC	0.96	12	35	0.0035	0.18	5.5	479	112,908
	Transaction price (IQVIA)	7.4	72	402	0.086	2.4	28	17,444	53,800
During shortage	AWP	1.2	11	37	0.0043	0.36	5.2	461	41,514
	WAC	0.71	8.1	26	0.0036	0.17	3.8	369	19,054
	Transaction price (IQVIA)	4.8	32	165	0.11	2	15	10,193	16,536

Notes: Quarterly prices in 2010 dollars, 2004 to 2019. Red Book rows cover injectable and non-injectable products of the shortage lists, IQVIA rows sterile injectable drugs. Generic products are those approved under an abbreviated application, brand under a new drug application. AWP (Average Wholesale Price) and WAC (Wholesale Acquisition Cost): list prices per milliliter, gram or each, averaged over priced months. IQVIA: transaction price per each (vial, syringe or bag), sales over items sold. N: product-quarters. Shortage quarter: short at any point.

D.4. Do Buyers Discipline Suppliers After a Shortage?

Most suppliers remain active after a shortage, and I find no evidence of a subsequent decline in their national market shares among products whose supply has recovered. I examine these outcomes as a descriptive check on the treatment of procurement in Section 4.2, using the panel of sterile injectable products over 2004–2019.

Columns 1 and 2 of Table D.3 measure whether a labeler still sells the product one and two years later. Columns 3 and 4 measure whether the plant is still observed supplying the market. The regressions include molecule–form by year fixed effects. Column 5 examines the product's subsequent log market share, adds product fixed effects, and restricts the outcome year to observations without a reported shortage. The shortage indicator and the fraction of the year in shortage enter separate

regressions.

Past shortages are associated with lower subsequent presence, but continued participation remains common. Relative to a next-year presence rate of 99.1 percent, a product-level shortage is associated with a reduction of 0.9 percentage points; a full year in shortage is associated with a reduction of 2.4 points. Plant presence is 3.0 percentage points lower in the next year, relative to a baseline of 92.4 percent. These are changes in observed market presence, not direct observations of contract renewal. Plant nonappearance in customs records can also reflect interrupted shipments rather than permanent exit.

Table D.3: **Continued Sales and Market Shares the Year After a Shortage**

	Labeler still sells the product		Plant still supplies the market		Log market share
	Next year	Two years later	Next year	Two years later	Next year
	(1)	(2)	(3)	(4)	(5)
Product short this year	-0.009^{**}	-0.012^{**}			0.233^{**}
	(0.004)	(0.006)			(0.098)
Fraction of the year short	-0.024^{***}				0.026
	(0.007)				(0.227)
Plant short this year			-0.030^{***}	-0.021^{**}	
			(0.009)	(0.010)	
Baseline mean	0.991	0.984	0.924	0.905	
Observations	29,185	25,739	32,369	24,949	22,771

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered by molecule-form in parentheses. Columns 1 to 4 are linear probabilities with molecule-form by year fixed effects; column 5 adds product fixed effects. Columns 1 and 5 enter each shortage measure separately. Outcomes are measured one year after the shortage year, two years in columns 2 and 4; the baseline mean is the share still present at that horizon. Column 5 keeps product-years without a shortage, so that shares reflect buyers' choices. Sterile injectable drugs, 2004 to 2019; markets are molecule-forms, products the labeler-product pairs.

Figure D.5 extends the comparison to three years after the shortage. It reports continued presence for labeler-products and plants, comparing suppliers whose product was short with those whose product was not short in the initial year, within molecule-form and year.

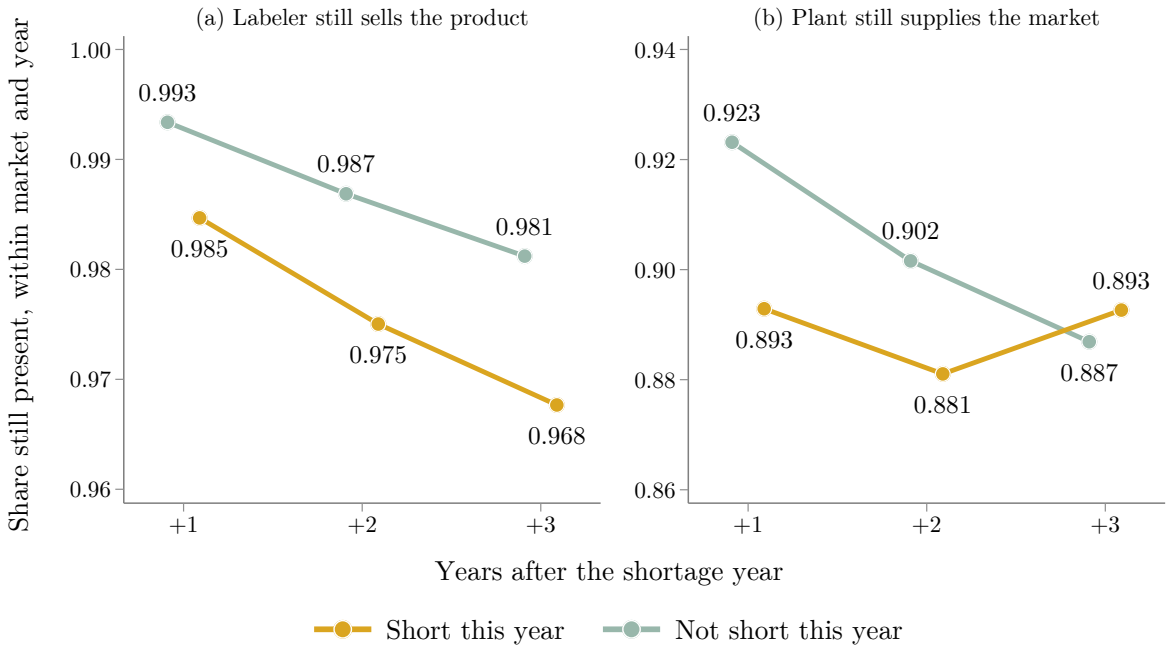


Figure D.5: **Suppliers Still Present One to Three Years After a Shortage**

Notes: Share of labeler-products (panel a) and of plants (panel b) still selling the product or supplying the market one, two and three years after a year in which the product was short, against products not short that year, within molecule-form and year as in Table D.3: the not-short series is the group mean and the short series adds the estimated difference. Sterile injectable drugs, 2004 to 2019.

The market-share estimates do not show a subsequent decline. The coefficient on the shortage indicator is positive; the coefficient on the fraction of the year in shortage is close to zero but imprecisely estimated. The latter estimate does not rule out economically meaningful losses of market share.

These patterns are consistent with limited aggregate market discipline after supply failures. Purchases are observed across all U.S. buyers, so the analysis captures suppliers' net market outcomes rather than individual buyers' responses. A supplier could lose one buyer and replace those sales elsewhere, weakening the effect of that buyer's response on the supplier's incentives. The data do not identify such switching separately from other changes in demand and the recovery of supply. The model of Section 4 therefore leaves shortage history out of buyers' choices, and the failure-to-supply penalty of Section 6 adds a cost that contracts do not impose.

E. Reduced-Form Results: Details

E.1. Offshoring and Shortages: Correlations

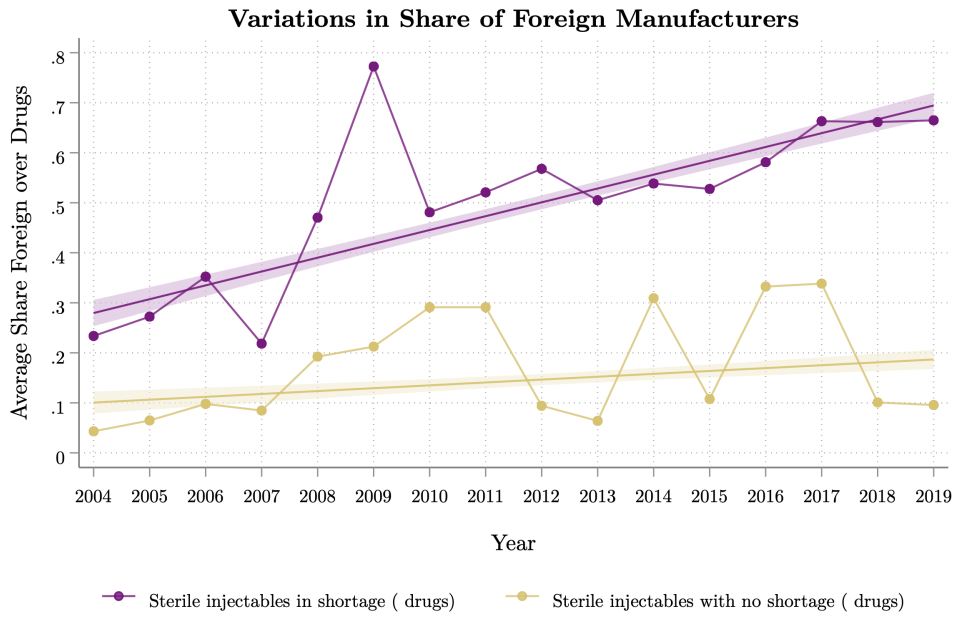


Figure E.1: **Injectable Drugs Going Short are Offshored More**

Notes: This figure shows that injectable drugs going short display higher shares of foreign-based manufacturers than sterile injectables that do not. Leveraging a balanced panel of generic sterile injectable drugs, the share of foreign producers of drugs that go short goes from 23% in 2004 to 67% in 2019. Generic injectable drugs that do not go short go from 4.3% of foreign manufacturers to 33% at their peak in 2017.

Table E.1: **Shortage Duration and Frequency, Imported versus Non-Imported Sterile Injectables**

		N	Mean	SD	p25	p50	p75
Shortage duration (days)	Non-imported	485	487.2	595.1	84.0	243.0	675.0
	Imported	794	616.8	753.8	120.0	329.0	823.0
Shortages per drug	Non-imported	245	1.98	1.40	1.0	1.0	3.0
	Imported	298	2.66	1.59	1.0	2.0	4.0

Notes: Sterile injectable drugs reported in shortage between 2004 and 2019. A drug counts as imported when at least one of its manufacturers is found in the U.S. import customs records (Appendix F). Duration is the number of days between the notification of a shortage and its resolution, measured over 1,279 shortage spells. Shortages per drug counts the distinct spells recorded for each of the 543 drugs. Spells still open at the end of 2019 are right-censored at their last observed month, which shortens recorded duration for 10.2 percent of spells.

E.2. A Dynamic Difference-in-Differences Framework

This appendix details the dynamic difference-in-differences design of Section 3.2, which estimates the effect of moving a drug's production from domestic to offshored plants on the probability and the

duration of its shortages.

E.2.1. Data Sample

The panel is annual, from 2002 to 2019, and covers 480 generic sterile injectable drugs, each defined as a molecule and dosage form. For every drug and year I observe its shortage outcomes, its unit price, its labelers, and its registered manufacturing plants, geo-located to the United States or abroad through the record linkage of Appendix F.2, which applies to domestic and foreign sites alike. The plant records sort drugs into three groups: 324 drugs have a registered foreign plant at some point in the sample; 65 never do and have at least one registered plant in the United States, which I call verified domestic; 91 have no usable plant record. An offshoring event is the first year in which a drug has a registered foreign plant, and a drug counts as offshored in every subsequent year. Among the 324, 25 are produced abroad in 2002, so their record before offshoring is never seen and they are not used; the remaining 299 supply the offshoring events.

E.2.2. Treatment and Control Groups

Treated drugs are the 299 whose production moves abroad during the sample period. In the comparison used throughout, a treated drug is compared in each year after its move with the drugs that have not yet offshored in that year together with the 65 verified-domestic drugs. The 91 drugs without plant information are held out: a drug with no observed plant cannot be verified as domestic, and admitting it would let missing information stand in for domestic production. I also report the comparison with the verified-domestic drugs alone, the never-offshored group, which gives larger estimates.

Table E.2 compares the groups before the move. Treated and never-offshored drugs are alike on price, on the number of years they are observed and, most directly, on their shortage rate before 2005, about 3 percent in both groups. They differ in one respect: treated drugs carry more labelers at baseline.[6]

E.2.3. Empirical Framework

Offshoring events are staggered across the sample period. A two-way fixed effects regression with leads and lags would mix comparisons between drugs treated at different dates, using earlier movers as controls for later ones, and could attribute part of a growing effect to the pre-period (Goodman-Bacon, 2021; Sun and Abraham, 2021). I therefore estimate the group-time effects of equation (1) and aggregate them in two ways. The event-study coefficient $\beta_k = \sum_g w_{gk}\,\mathrm{ATT}(g, g+k)$ plotted in Figure 4 averages the cohort effects at the same distance k from the event, with weights w_{gk} proportional to the number of drugs of cohort g observed at horizon k; the average effect reported in the tables averages $\mathrm{ATT}(g,t)$ over the event year and the ten years after it, weighting cohorts by their size.[7]

[6] Treated drugs also have more registered manufacturing sites, but the site count is partly an outcome of the move, since a foreign plant adds to it, and the labeler count moves with the same decisions that send production abroad. Neither enters the estimation, as conditioning on them would absorb part of the effect being measured.

[7] The number of offshored drugs observed at each horizon falls from 299 in the year of the event to 126 ten years after it, as later movers reach the end of the panel; the tables and figures stop there.

Table E.2: **Summary Statistics: Treated and Control Drugs before Offshoring**

Drug Characteristic	Treated	Control
Log Unit Price (AWP, 2010 dollars) at baseline	2.959 [2.746]	3.047 [3.131]
Number of Labelers at baseline	1.798 [1.332]	1.263 [0.455]
Years Observed	14.867 [4.490]	14.231 [5.147]
Number of Manufacturing Plants	3.340 [3.106]	1.406 [1.157]
Shortage Rate, 2002-2004	0.031 [0.155]	0.028 [0.167]
Drugs	324	65

Notes: Standard deviations in brackets. Treated drugs are the 324 that have a registered foreign plant at some point in the sample, including the 25 already abroad in 2002 that the estimation leaves out; control drugs are the 65 that never do and have at least one registered plant in the United States. The unit price (average wholesale price, in 2010 dollars, in logs) and the number of labelers are averaged over the years before the drug's offshoring event, or over 2002 to 2004 for control drugs; the number of plants is the average number of registered manufacturing sites over all years; the shortage rate is the share of years in shortage over 2002 to 2004, among the drugs observed in those years.

E.2.4. Results

Table E.3 reports the average effects discussed in Section 3.2, against drugs not yet offshored and against never-offshored drugs, next to the mean of each outcome among offshored drugs; the second comparison gives larger estimates for every outcome.

Table E.3: **Average Effect of Offshoring Events on Drugs' Shortage Outcomes**

Dependent Variable	Control: Not Yet Offshored	Control: Never Offshored	Mean (treatment)
Dummy for Shortage	0.089*** [0.031]	0.126*** [0.034]	0.272
Dummy for New Notified Shortage	0.072*** [0.026]	0.102*** [0.026]	0.171
Fraction of Year in Shortage	0.082*** [0.027]	0.110*** [0.029]	0.219
Shortage Length in Days (log)	0.569*** [0.218]	0.836*** [0.237]	301.5 days
Observations	4,752	4,752	
Treated Drugs	299	299	
Never-Offshored Controls	65	65	
Not-Yet-Offshored Drugs as Controls	Yes	No	

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the drug level in brackets. Annual panel of generic sterile injectable drugs, 2002 to 2019. Treatment starts the first year a drug has a registered foreign plant; each coefficient is the group-time estimate of Callaway and Sant'Anna (2021) averaged over the event year and the ten years after it, with no covariates. The control group is either the drugs not yet offshored together with the 65 never-offshored drugs, or the 65 never-offshored drugs alone. Mean (treatment) is the average of the dependent variable among offshored drugs over all their years in the panel, in days for shortage length. The fraction of the year in shortage counts a day as short when at least one of the drug's NDCs is in shortage. Shortage length is the length in days of the shortage event a drug is in during a year, zero in years without a shortage, in logs of one plus the days.

Table E.4 examines the years before the event. Joint tests that the five pre-event coefficients are all zero do not reject for either incidence measure, nor do tests restricted to the three years nearest

the event. Splitting treated drugs by whether they move before 2009, the tests do not reject for earlier movers; among later movers the joint test for newly notified shortages comes closest to rejecting. Averaged over the five pre-event years, the year-on-year differences are 0.018 and 0.015 for the probability of shortage and 0.014 and 0.011 for newly notified shortages, weighting the years equally or as the reported effect weights its components, small against mean shortage probabilities of 27 and 17 percent.

Table E.4: **Pre-Event Differences between Treated and Control Drugs**

Dependent Variable	Shortage Dummy	New Notified Shortage Dummy
	Joint test that the pre-event coefficients are zero	
All five pre-event years	7.56 [0.182]	8.22 [0.145]
Three years nearest the event	4.00 [0.261]	1.32 [0.724]
Drugs offshored earlier	0.55 [0.908]	3.18 [0.672]
Drugs offshored later	8.62 [0.125]	10.45 [0.063]
	Average pre-event difference	
Equal weights	0.018 [0.008]	0.014 [0.007]
Weights of the reported effect	0.015 [0.008]	0.011 [0.007]

Notes: The first panel reports joint Wald tests that the estimated pre-event coefficients β_k of Figure 4 are all zero, with p-values in brackets; the second row uses only the three years nearest the event, and the last two rows split treated drugs by whether their event falls before 2009 and re-estimate each group against the same control drugs. The second panel reports the average of the five pre-event coefficients, weighting the years equally or as the reported effect weights its components, with standard errors clustered at the drug level in brackets.

Table E.5 reports the effect on the length of shortage events in days, in levels, with the longest events capped at the ninety-fifth percentile, and in logs. A few events last more than a decade, so the level estimates are imprecise: only the estimate against never-offshored drugs, about 121 days, is distinguishable from zero, and capping the longest events tightens both estimates without changing that. The log estimates are the ones quoted in the text.

Table E.5: **Effect of Offshoring Events on Shortage Length**

Dependent Variable	Control: Not Yet Offshored	Control: Never Offshored
Shortage Length in Days	68.5 [52.7]	120.9** [58.6]
Shortage Length in Days, Capped at the 95th Percentile	59.1 [41.4]	107.1** [44.1]
Shortage Length in Days (log)	0.569*** [0.218]	0.836*** [0.237]
Observations	4,752	4,752
Treated Drugs	299	299

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the drug level in brackets. The dependent variable is the length in days of the shortage event a drug is in during a year, measured as the difference between the first notification of the shortage and its resolution, zero in years without a shortage; the second row caps events at the ninety-fifth percentile of the distribution, 1,603 days; the third takes the logarithm of one plus the days. Treatment, control groups and averaging as in Table E.3.

E.2.5. Robustness to the Definition of the Control Group

Figure E.2 varies the set of domestic control drugs: the 65 verified-domestic drugs, or widened to drugs with at most 10 or 20 percent of their NDCs produced abroad over the sample period, which are then not counted as offshored. The three paths are nearly indistinguishable, so the result does not depend on where the line between domestic and offshored production is drawn.

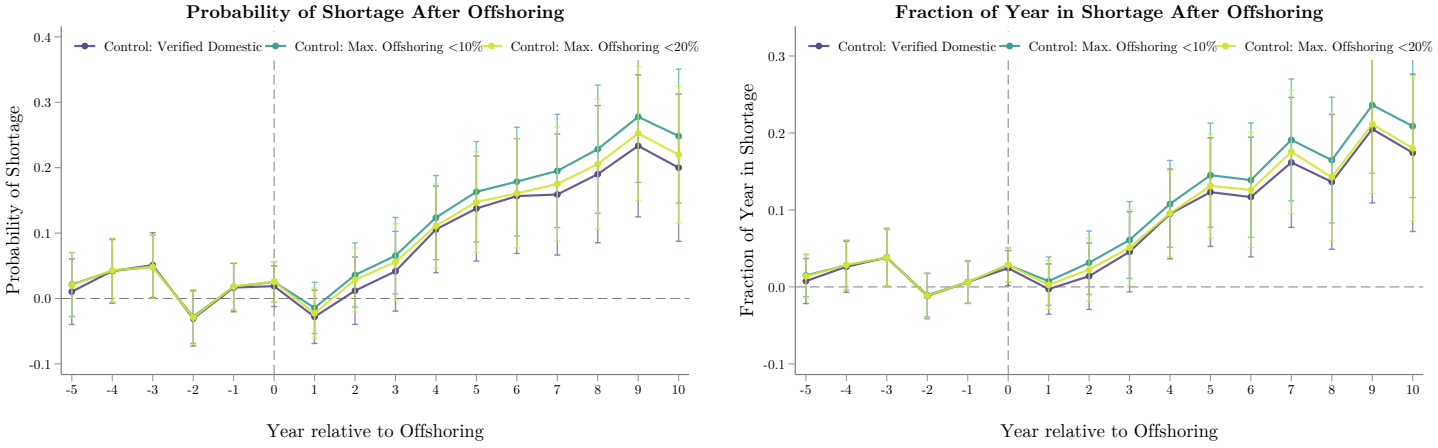


Figure E.2: **Robustness to the Definition of the Control Group**

Notes: Each panel plots the event-study coefficients β_k defined in the note to Figure 4 by year relative to the offshoring event, with 95 percent confidence bars from standard errors clustered by drug. In every line the comparison in a given year is the drugs not yet offshored together with a set of domestic drugs: the 65 verified-domestic drugs for the purple line, widened for the green and yellow lines to drugs with at most 10 or 20 percent of their NDCs produced abroad over the sample period. The left panel uses an indicator equal to one if the drug is in shortage; the right panel the fraction of the year during which at least one of the drug's NDCs is in shortage.

E.2.6. Robustness to the Estimator

Figure E.3 re-estimates the average effect on the probability of shortage, on the probability of a newly notified shortage and on the fraction of the year in shortage with a two-way fixed effects regression with leads and lags, with the group-time estimator against each control group, and with the estimator of de Chaisemartin and D'Haultfœuille (2024), all over the event year and the ten years after it. The averages are close for every outcome: for the probability of shortage, the two-way fixed effects regression gives 0.087 and the estimator of de Chaisemartin and D'Haultfœuille (2024) 0.083, against the reported 0.089. The choice of estimator does not drive the results.

E.3. A Shift-Share Instrument for Offshoring: Details

This appendix details the instrument behind Fact 3 of Section 3.1. The decision of a manufacturing firm to offshore production is endogenous, and there is also selection on the type of drugs and firms that will decide to locate in the Global South. In order to strengthen the evidence for a causal interpretation, I develop a shift-share instrumental variable for the offshoring of sterile injectable drugs.

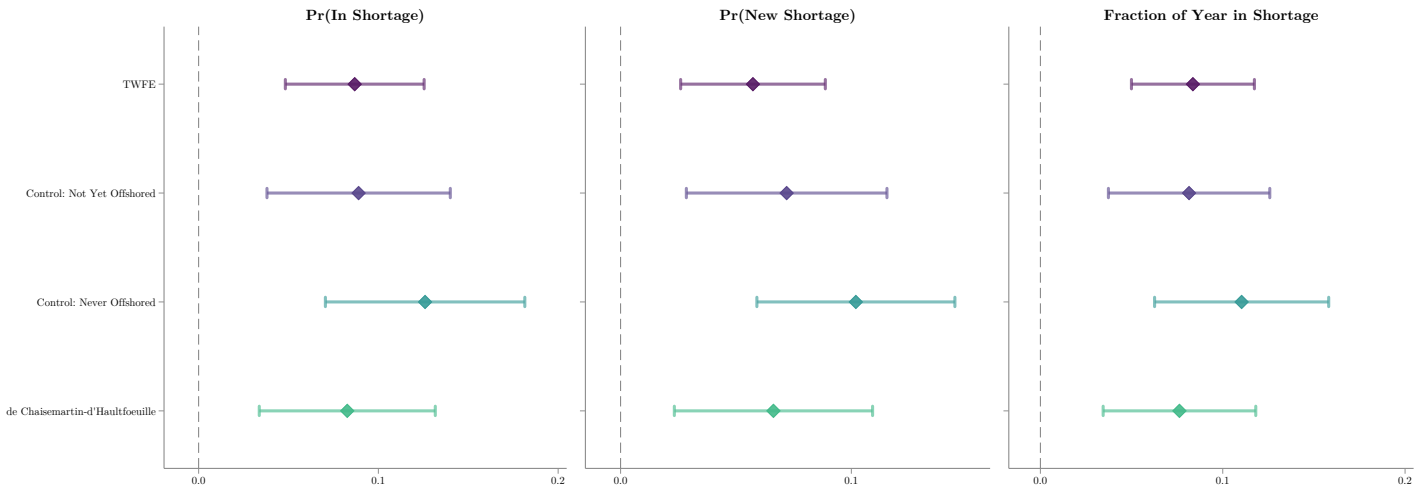


Figure E.3: **Robustness to the Estimator**

Notes: Each panel plots the average effect of the offshoring event over the event year and the ten years after it, with 95 percent confidence bars from standard errors clustered by drug, on the 299 offshored and 65 verified-domestic drugs. The rows are a two-way fixed effects regression with leads and lags, restricted to the same horizons; the group-time estimator of Callaway and Sant'Anna (2021) against drugs not yet offshored together with the verified-domestic drugs, the comparison used throughout, and against the verified-domestic drugs alone; and the estimator of de Chaisemartin and D'Haultfœuille (2024), which averages its dynamic effects over the same horizons. The fraction of the year in shortage counts a day as short when at least one of the drug's NDCs is in shortage.

Instrumental Variable Strategy. Over the past two decades, the global medicine consumption pattern has shifted, largely driven by a boom in the consumption of healthcare and generic drugs in "Pharmerging" giants such as Brazil, Mexico, India, and China.[8] As the world's consumption of drugs shifts, production re-allocates to get closer to this new booming market.

The instrument must isolate a plausibly exogenous source of variation in offshoring rates at the drug product-year-geographical location level. To obtain such variation, I use the idea that the demand for pharmaceutical treatments varies across demographic groups based on the incidence of disease (e.g., some drugs are disproportionately used by men from 25-45). Therefore, changes in demographic composition over time and across countries (the "shifter") interacted with the propensity for a demographic group to use a drug (the "share") provide exogenous variations in demand. Identification in this shift-share design thus rests on the exogeneity of the demographic shifts rather than of the drug-use shares (Borusyak, Hull, and Jaravel, 2022; Goldsmith-Pinkham, Sorkin, and Swift, 2020).

The identifying assumption is that these demographic shifts affect the supply of U.S. hospitals only through where a drug is produced. A rise in demand in India or China could affect that supply directly only if manufacturers located there served their home market first and stopped shipping to the U.S. The U.S. is the most profitable market for generic injectables and is served first; shortages are more frequent and longer in Europe because the U.S. pays more (Dubois, Majewska, and Reig, 2026). Such a diversion also requires that the drug be produced in Asia in the first place, so that it operates through the location of production, which is what the instrument captures, and the production lines that serve

[8] See Figure S3.1 in the Additional Results. IQVIA estimates that the accelerating shift in global spending toward generics, rising from 20% in 2005 to 39% of spending in 2015, was mostly driven by an increase of consumption from these four countries.

the U.S. market are not those that serve the Indian market, as production standards differ.

Predicted Disease Burden. Equation (E.1) captures the construction of the index.

$$(PDB)_{it}^{d} = \sum_{a,g} \left[\text{InsuredPop}_{\text{iagt}} \times \left(\frac{\sum_{k \neq i} \text{disease burden}_{\text{kagt}}^{d}}{\sum_{k \neq i} \text{population}_{\text{kagt}}} \right) \right] \tag{E.1}$$

The last ratio term measures the average disease burden per capita from disease d for gender g and age group a, calculated excluding the country of interest (that is, summing over all other locations k in the Global Burden of Disease data except for country i).[9] This ratio is then weighted by the insured population of gender g and age group a in country i and year t (InsuredPop), that is, the population of the group multiplied by the share of the population insured in that country and year, and summed across age and gender groups. Because the rise in consumption of drug products in the South is in part driven by the rapid access of wealthier populations to health insurance, weighting by the insured rather than the total population provides additional variation at the country-year level. The PDB index eventually varies at the country i-disease d-year t level. Each injectable drug is then matched to the disease (or set of diseases) it treats, and PDB_{jt} averages the index over India and China.

Predicted Disease Burden Measures: Data Construction. To build this predicted disease burden measure, I use the Global Burden of Disease 2019 study of the Institute for Health Metrics and Evaluation (Global Burden of Disease Collaborative Network, 2020). The data provide the country-year-disease measures of burden, broken down into eight demographic groups: four age groups (0–14, 15–49, 50–69, and 70 and over) for each gender.[10] I use as a primary measure of disease burden the number of lost disability-adjusted life years (DALYs), which combines data on the mortality and morbidity caused by each disease. I eventually use population data for each country in each of the eight demographic groups for each year (2004-2019) from the same source, and the share of people insured from its universal health coverage effective coverage index.[11] To link drugs to diseases, I first match each drug product to its therapeutic category in the WHO's ATC classification, using the FDA's and KEGG's drug-to-ATC mappings and RxNorm queries for the remaining drugs, and I create crosswalks that allow linking each therapeutic category in my dataset, based on the Ephmra classification, to the WHO's list of therapeutic categories (ATC). I then map each therapeutic category to a Global Burden of Disease cause, starting from the hand-coded mapping made by Costinot et al. (2019) between therapeutic categories and disease burden, which I crosswalk to the current cause codes of the study and extend to the injectable therapeutic categories it does not cover. The constructed measure displays significant variations over time and locations (Figure E.4).

IV Regression Results. Table E.6 displays the results of a regression of shortage probability on offshoring, in which I instrument for offshoring of drug facilities with the constructed predicted disease

[9] Note that the fact that firms from country i are better at treating a given disease d may cause a lower burden for that disease in country i. Leaving out country i from the average disease burden per capita addresses this endogeneity issue.

[10] Although there may be local variation in the collection of vital statistics that underpin these measures, the study ensures that these data are valid for cross-country and cross-disease comparisons.

[11] The index is published for 1990, 2010 and 2019; I interpolate it linearly between these years.

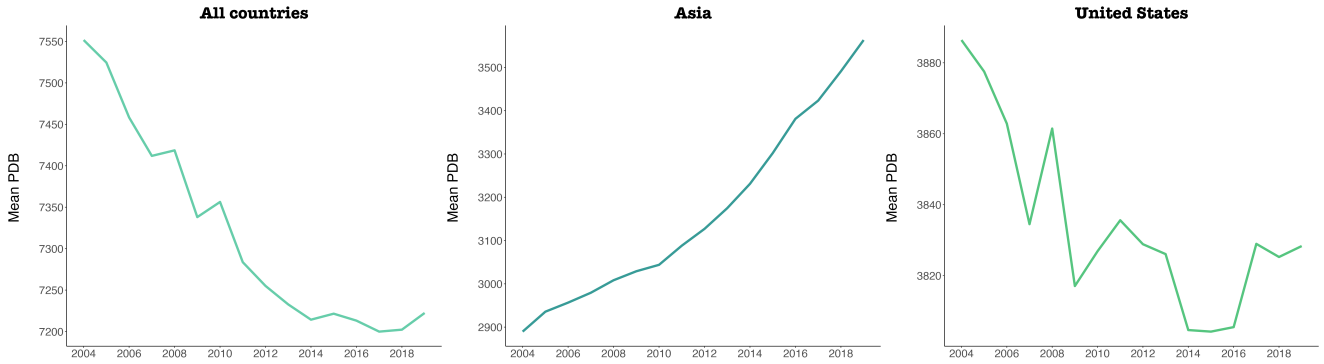


Figure E.4: Variations in Predicted Disease Burden Measures

Notes: Variations over time and across countries in the predicted disease burden measure of equation (E.1), which varies at the country-disease-year level.

burden measure. The dependent variable, Shortage_{jt}, is either (i) a dummy equal to one if a new shortage of drug j is reported in year t (linear probability model), or (ii) the fraction of time, within a year t, that drug j is short:

$$\text{Shortage}_{jt} = \alpha_j + \beta \text{Offshoring}_{jt} + \phi_p \log\left(p_{jt}^{AWP}\right) + \phi_M M_{jt} + \phi_P \mathbb{1}\{\text{PatentStatus}_{jt}\} + \nu_{jt}$$

where the first-stage is

$$\text{Offshoring}_{jt} = a_j + \pi \log(\text{PDB}_{jt}) + \varphi_p \log\left(p_{jt}^{AWP}\right) + \varphi_M M_{jt} + \varphi_P \mathbb{1}\{\text{PatentStatus}_{jt}\} + \omega_{jt}$$

Offshoring_{jt} captures the share of drug j's manufacturing facilities located outside of the U.S. (or in Asia) at time t. $\log\left(p_{jt}^{AWP}\right)$ is the log-median average wholesale price (AWP) per unit (in 2010 dollars), M_{jt} measures the number of labelers for drug j at t and $\mathbb{1}\{\text{PatentStatus}_{jt}\}$ is an indicator variable equal to one if the drug lost its patent status.[12] $\log(\text{PDB}_{jt})$ captures the log predicted disease burden associated with drug j in year t. All regressions include drug fixed effects.[13]

All else equal, an increase in a drug's share of foreign manufacturing facilities is strongly correlated with an increase (i) in the probability that a new shortage occurs and (ii) in the fraction of time in a given year during which a drug is in shortage. Moving from an entirely domestic to an entirely foreign production structure (a 100-percentage-point increase in the share of foreign manufacturers) raises the probability of a new drug shortage by 24 percentage points, all else equal, against a yearly baseline of 7.7 percent. This goes up to 30 percentage points if we only consider Asian manufacturing facilities. Similarly, moving the offshoring share from 0 to 1 raises the fraction of the year in shortage by 74 percentage points, that is, about 270 additional days of shortage a year for a drug short 44 days a year on average (95 percentage points, or about 350 days, for Asian manufacturing facilities). One within-drug standard deviation of predicted disease burden, holding the controls fixed, moves the share of

[12] A drug j in these regressions is an NDC, a labeler's version of a molecule-form drug, and its manufacturing facilities are the plants registered to produce it in a given year. The controls are measured on its molecule-form market. The off-patent indicator equals one from the year of the molecule-form's earliest Orange Book patent expiration onward, and in every year for molecule-forms with no listed patent.

[13] With drug fixed effects, the instrument's variation is the growth of predicted disease burden within a drug over time, as the populations of India and China age and gain insurance coverage.

foreign manufacturers by close to 4 percentage points, the likelihood of a new shortage by close to 1 percentage point, and the time in shortage by about ten days a year. Similar regressions without controls (Additional Results Table S3.3) provide coefficients of similar magnitude; only the effect of Asian production on the fraction of the year in shortage is larger once the controls, which hold fixed a market's price, number of labelers and patent status, are added. The instrumental variable estimates are larger than their least squares counterparts, reported in the table, which I read as measurement error in the observed share of foreign manufacturers and as an effect concentrated among the drugs offshored in response to the growth of demand in Asia.

Robustness. Four further checks, reported in Appendix Tables E.7 and E.8, confirm this reading. First, the estimates reflect demand growth in Asia and not in the United States, whose population ages over the same period. I control for the predicted disease burden of the United States, built with the same leave-one-out construction, and alternatively instrument with the difference between the Asian and the U.S. burdens. The effects of an entirely foreign or entirely Asian production structure remain between 19 and 25 percentage points on the probability of a new shortage and between 77 and 87 percentage points on the fraction of the year in shortage. The U.S. burden itself leaves shortages unchanged once the share of foreign manufacturers is held fixed. Second, the instrument captures the pull of Asian markets specifically, and not a general move of production abroad. A rise in predicted disease burden moves the Asian share of a drug's manufacturing facilities by 0.37 and barely moves the share located in Europe and other high-income countries (Appendix Table E.8). Third, the estimates hold in every part of the period. Excluding 2010 to 2012, the peak years of U.S. shortages, raises the effects of an entirely foreign production structure to 30 and 82 percentage points. The reduced-form effect of predicted disease burden stays positive and close to its full-sample value whichever single year is left out (Additional Results Figure S3.4). Finally, the estimates hold when every molecule-form market counts equally rather than in proportion to its number of NDCs. The effects rise to 37 and 81 percentage points.

Table E.6: **Probability of shortage on share foreign - Instrument: log(PDB)**

Dependent Variable:	**P(New Shortage=1)**		**Fraction of Year Short**	
	(1)	(2)	(3)	(4)
Second Stage†				
Facilities, Share Foreign	**0.2366*****		**0.7395*****	
	[0.0598]		[0.1405]	
Facilities, Share Asia		**0.3039*****		**0.9497*****
		[0.1040]		[0.2738]
Log Median Unit Price	-0.0119	-0.0101	-0.0191	-0.0135
	[0.0063]	[0.0063]	[0.0121]	[0.0135]
Number of Labelers	0.0034	0.0000	-0.0038	-0.0145
	[0.0020]	[0.0035]	[0.0063]	[0.0126]
Off-Patent Dummy	-0.0073	-0.0204	-0.0012	-0.0422
	[0.0139]	[0.0138]	[0.0284]	[0.0433]
First Stage				
log(PDB)	0.4778***	0.3721***	0.4778***	0.3721***
	[0.1126]	[0.0621]	[0.1126]	[0.0621]
F-stat First-Stage‡	18.01	35.85	18.01	35.85
Magnitudes				
Within-drug SD of the instrumented share (pp)	3.85	3.00	3.85	3.00
Effect of a one-SD move in the instrument	0.0091	0.0091	0.0285	0.0285
Mean of the Dependent Variable	0.077	0.077	0.121	0.121
OLS Coefficient	0.0404**	0.0494***	0.0706**	0.0334
	[0.0176]	[0.0114]	[0.0337]	[0.0202]
Inference and Sample				
Wild cluster bootstrap p-value, second stage	0.035	0.035	0.026	0.026
Drug Fixed Effects	✓	✓	✓	✓
Observations	21,502	21,502	21,502	21,502
NDCs	2,559	2,559	2,559	2,559
Molecule-Forms	275	275	275	275
Disease Clusters	39	39	39	39

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered by disease in brackets. Annual panel of NDCs (a labeler's version of a molecule-form drug) of sterile injectable drugs, 2004 to 2019; the first outcome equals one if a new shortage of the NDC is notified in the year, the second is the fraction of the year the NDC is in shortage. †Instrument: log predicted disease burden in India and China. ‡Kleibergen-Paap F statistic. All regressions include drug fixed effects and controls measured on the molecule-form market of the drug. The wild cluster bootstrap p-values are weak-instrument robust (Anderson-Rubin form, Webb weights).

Table E.7: **Probability of shortage on share foreign: robustness**

Dependent Variable:	**P(New Shortage=1)**		**Fraction of Year Short**	
	(1) Foreign	(2) Asia	(3) Foreign	(4) Asia
Panel A. Baseline†				
Facilities, Share	0.2366***	0.3039***	0.7395***	0.9497***
	[0.0598]	[0.1040]	[0.1405]	[0.2738]
F-stat First-Stage‡	18.01	35.85	18.01	35.85
Bootstrap p-value, second stage	0.035	0.035	0.026	0.026
Panel B. Controlling for U.S. Predicted Disease Burden				
Facilities, Share	0.2134**	0.1884**	0.8727**	0.7702***
	[0.1033]	[0.0726]	[0.3467]	[0.2065]
log(PDB), United States	0.0175	0.0679	-0.1005	0.1055
	[0.0654]	[0.0623]	[0.1997]	[0.1390]
F-stat First-Stage‡	31.99	58.87	31.99	58.87
Bootstrap p-value, second stage	0.079	0.079	0.012	0.012
Panel C. Instrument Relative to U.S. Predicted Disease Burden				
Facilities, Share	0.2275***	0.2479***	0.7918***	0.8627***
	[0.0677]	[0.0736]	[0.2109]	[0.2086]
F-stat First-Stage‡	22.81	68.96	22.81	68.96
Bootstrap p-value, second stage	0.013	0.013	0.006	0.006
Panel D. Excluding 2010 to 2012				
Facilities, Share	0.2981***	0.3916***	0.8224***	1.0805***
	[0.0584]	[0.1013]	[0.1170]	[0.2503]
F-stat First-Stage‡	19.99	37.55	19.99	37.55
Bootstrap p-value, second stage	<0.001	<0.001	<0.001	<0.001
Observations	17,947	17,947	17,947	17,947
Panel E. Each Molecule-Form Market Weighted Equally				
Facilities, Share	0.3686***	0.3955***	0.8074***	0.8663***
	[0.1113]	[0.1171]	[0.1627]	[0.2573]
F-stat First-Stage‡	20.44	27.11	20.44	27.11
Bootstrap p-value, second stage	<0.001	<0.001	0.012	0.012
Drug Fixed Effects	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Observations	21,502	21,502	21,502	21,502

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered by disease in brackets. †Instrument: log predicted disease burden in India and China; Panel A reproduces Table E.6. Panel B adds the log predicted disease burden of the United States, built like the instrument and leaving out the United States' own burden; Panel C instruments with the difference between the two log burdens; Panel D drops 2010 to 2012 and the NDCs left with a single year; Panel E weights each NDC-year by one over the number of NDC-years of its molecule-form. F is the Kleibergen-Paap statistic; bootstrap p-values are wild cluster bootstrap, weak-instrument robust (Anderson-Rubin form, Webb weights).

Table E.8: **First stage by location of manufacturing facilities: Asia versus Europe**

Dependent Variable:	**Facilities, Share**			
	(1) Foreign	(2) Asia	(3) Foreign, non-Asia	(4) Europe and High Income
Panel A. With Controls				
log(PDB)	0.4778***	0.3721***	0.1058	0.0481
	[0.1126]	[0.0621]	[0.0821]	[0.0710]
F-stat First-Stage‡	18.01	35.85	1.66	0.46
Wild cluster bootstrap p-value	<0.001	<0.001	0.345	0.646
Mean of Dependent Variable	0.6924	0.3920	0.3004	0.2767
Share of NDC-Years with a Positive Share	0.7922	0.6302	0.5025	0.4820
Share of NDCs whose Share Moves	0.5467	0.7042	0.6409	0.6245
Observations	21,502	21,502	21,502	21,502
NDCs	2,559	2,559	2,559	2,559
Disease Clusters	39	39	39	39
Panel B. Without Controls				
log(PDB)	0.4777***	0.4297***	0.0480	0.0052
	[0.1118]	[0.0790]	[0.0714]	[0.0651]
F-stat First-Stage‡	18.26	29.60	0.45	0.01
Wild cluster bootstrap p-value	<0.001	<0.001	0.589	0.942
Mean of Dependent Variable	0.6804	0.3557	0.3248	0.3021
Share of NDC-Years with a Positive Share	0.7827	0.5902	0.5201	0.5025
Share of NDCs whose Share Moves	0.5205	0.6534	0.6094	0.5956
Observations	25,534	25,534	25,534	25,534
NDCs	2,980	2,980	2,980	2,980
Disease Clusters	45	45	45	45
Drug Fixed Effects	✓	✓	✓	✓

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered by disease in brackets. ‡Kleibergen-Paap F statistic. Each column regresses the share of the drug's manufacturing facilities in the indicated location on log predicted disease burden in India and China, with drug fixed effects. Column (3) counts the foreign facilities not counted in column (2); column (4) counts facilities in Europe, Australia, Canada, Hong Kong, Israel, Japan, New Zealand, Singapore, South Korea and Taiwan. Panel A adds the controls. Bootstrap p-values use Webb weights.

F. Data Construction

F.1. Sources and Observation Units

Figure F.1 summarizes the three blocks of data. Drug-level sources identify marketed products, sales, and shortages. Establishment records identify manufacturing plants. Shipment descriptions and shipper addresses connect the two. I retain product–plant relationships separately from the shipments observed at a particular date.

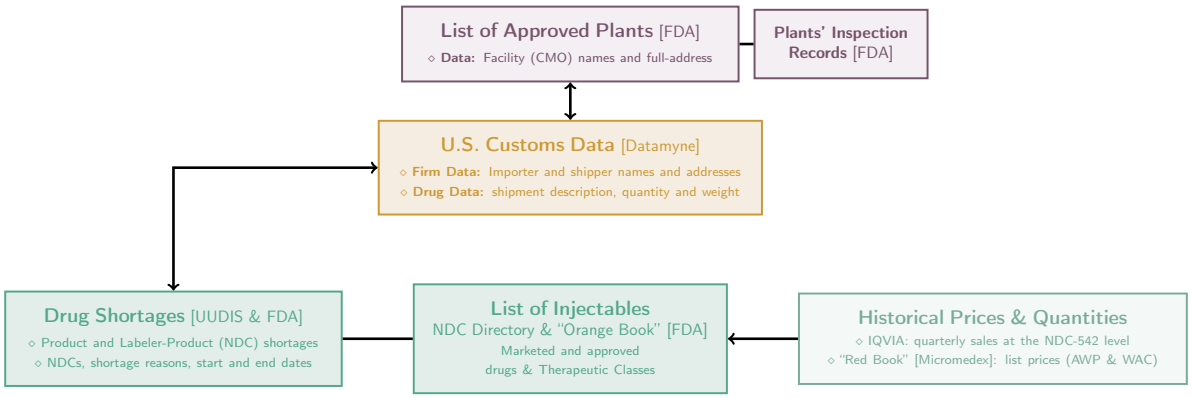


Figure F.1: **Overview of Data Structure**

Notes: Three main blocks of data were used for this paper. The first block captures the characteristics of each sterile injectable drug market (downstream). The main data of this block are the drug shortage data (UUDIS), which are complemented with historical data on drug characteristics (FDA's Orange Book), quarterly sales and pricing data (IQVIA). The second block maps the upstream market and records approved plants and their location (FDA's CDER data). I complement these data with digitized inspection records and warning letters. The third and last block unifies the first two: I use detailed customs data (Descartes Datamyne), U.S. import and export records and Indian export records, to link drug products to manufacturing plants and labelers, using probabilistic record linkage techniques. More details in Appendix F.2.

Drug shortages. Erin Fox at UUDIS provided drug-level shortage histories and detailed labeler-product availability reports. A drug-level record contains the drug name, dosage form, onset date, resolution date when available, reported reason, and American Hospital Formulary Service therapeutic classification. The detailed reports identify which labelers and National Drug Codes (NDCs) are affected and which alternative presentations remain available. I extract these fields from thousands of text documents to construct the product-level shortage histories. The reports provide shortage status and dates, not the volume of missing supply.

Drug characteristics. I combine the FDA's NDC Directory with the Orange Book. The NDC identifies a labeler, product, and package; the Directory also records names, dosage forms, routes of administration, and marketing dates. The Orange Book identifies approved products and application holders, earliest approval dates, generic or brand status, and patent and exclusivity information. I link the sources through application numbers and complement the current Orange Book with historical editions. The resulting product list includes drugs never reported in shortage, which are absent from the UUDIS shortage lists. An establishment, a marketed labeler-product, and a molecule–form market are

distinct units.

IQVIA sales and realized unit prices. IQVIA provides quarterly sales revenue and physical quantities at the package-level NDC (5-4-2), aggregated across U.S. buyers. The data cover 4,981 NDC-54 products, which pool the packages of a labeler's product, in 1,427 molecule-form markets and 682 ATC5 classes. For injectables, the quantity unit is the individual item, such as a vial, ampule, or syringe. In the structural sample, I aggregate packages belonging to a labeler-product (NDC 5-4) by summing revenue and units within each year. The unit price is their ratio. Section 5 defines the market aggregation. These prices are national product-level averages; individual buyers, contracts, and contract-specific rebates are not observed. The data also do not allocate a labeler-product's sales across its manufacturing plants.

Red Book published prices. The historical Red Book pricing tables report quotations submitted by drug manufacturers, with NDCs, labeler names, price types, and effective dates. I complement the historical files with active price records collected from the Red Book. WAC and AWP are published price benchmarks rather than observed GPO contract prices. Within a product's period of validity, I carry forward the last available quotation when no update is recorded in a quarter. The combined files match 97 percent of the 8,106 NDCs on the product-level shortage list; matching at the NDC–year–quarter level gives 393,132 of 688,407 observations. The first statistic measures whether a product has any price record; the second requires a quotation for the particular quarter. Appendix D.3 uses these series for descriptive price profiles, not for the baseline structural price measure.

FDA manufacturing establishments. I obtained from the FDA's Center for Drug Evaluation and Research (CDER), through a Freedom of Information Act request, the yearly records of domestic and foreign establishments registered to make drugs for the U.S. market over 2000–2019. Each record gives the facility name and full address, its FEI or DUNS identifier where recorded, and the operations it declares, such as manufacturing active ingredients or finished dosage forms, repacking, or relabeling. I retain manufacturing establishments and complement the records with facilities reported under the Generic Drug User Fee Amendments (GDUFA), using self-identification records for 2016–2019 and fee-paying facility records for 2013–2019. Registration or fee payment is not conditional on positive production in a given year, and the records do not identify the drugs made at each site.

I also use FDA inspection records and warning letters obtained through Freedom of Information Act requests. I extract establishment information from these records and link it to the plant panel. The regional inspection frequencies in Section 2.2 are calculated from these data over 2004–2019. Additional Results Section S1.3 describes inspection classifications and regulatory follow-up.

U.S. and foreign customs records. Descartes Datamyne supplies shipment-level descriptions, shipper and consignee names and addresses, dates, and quantities or weights. The U.S. import records are maritime bills of lading. I complement them with U.S. export records, Indian trade records, and smaller export files from Argentina, Chile, and Uruguay. Indian exports provide additional links between drugs and Indian producers; U.S. exports help identify domestic plants that are absent from U.S. import manifests. The shipment panel used in the main descriptive analysis spans 2004–2019.

F.2. Record Linkage: Techniques and Results

Date	Consignee Declared	Consignee Declared Address	Shipper Declared	Shipper Address	Short Container Description	Country of Origin	Final Destination	Weight	Quantity
12/29/2019	ASCEND LABORATORIES, LLC	339 JEFFASON ROAD PARSIPPANY NJ USA-07054 PARSIPPANY NJ 07054 US	ALKEM LABORATORIES LTD.	ALKEM HOUSE, DEVASHISH, SENAPATI BAPAT MARG, LOWER PAREL, ADJACENT MUMBAI MAHARASHTRA 400013 IN	1775 SHIPPER PACKED IN 40 PACKAGES PHARMACEUT ICAL PRODUCTS GABAPENTIN TABLETS USP 600MG 50 0S NDC NO : 6787742805 AMLODIPINE BESYLATE TA BLET 2.5MG 500S NDC NO : 6787719705 AMLODIPIN E BESYLATE TABLE [MORE]	INDIA	CINCINNATI-LAWRENCEBURG,OH	10,008.00	1,775.00
12/29/2019	ASCEND LABORATORIES, LLC	339 JEFFASON ROAD PARSIPPANY NJ USA-07054 PARSIPPANY NJ 07054 US	ALKEM LABORATORIES LTD.	ALKEM HOUSE, DEVASHISH, SENAPATI BAPAT MARG, LOWER PAREL, ADJACENT MUMBAI MAHARASHTRA 400013 IN	489 SHIPPER PACKED IN 40 PACKAGES PHARMACEUTI CAL PRODUCTS CEPHALEXIN 500MG CAPS 500S NDC N O : 6787721905 CEPHALEXIN 500MG CAPS 100S NDC NO : 6787721901 DATA LOGGER NO R0121548 DATA LOGGER IS KEPT IN [MORE]	INDIA	CINCINNATI-LAWRENCEBURG,OH	7,568.00	489.00
12/29/2019	ASCEND LABORATORIES, LLC	339 JEFFASON ROAD PARSIPPANY NJ USA-07054 PARSIPPANY NJ 07054 US	ALKEM LABORATORIES LTD.	ALKEM HOUSE, DEVASHISH, SENAPATI BAPAT MARG, LOWER PAREL, ADJACENT MUMBAI MAHARASHTRA 400013 IN	939 SHIPPER PACKED IN 39 PACKAGES PHARMACEUTI CAL PRODUCTS GABAPENTIN 300MG CAPSULES1000S N DC NO : 6787722310 QUETIAPINE TABLETS USP 25M G 100S NDC NO : 6787724201 QUETIAPINE TABLETS	INDIA	CINCINNATI-LAWRENCEBURG,OH	10,264.00	939.00

Figure F.2: **Bill-of-Lading Data for U.S. Imports: Container Descriptions**

Source: Descartes Datamyne. An example of the detailed container descriptions from which drug names and dosage forms are extracted.

F.2.1. Matching Techniques

No common identifier links these datasets. Most fields are text whose format varies within and across sources: drug names rather than drug codes, or codes at different levels that do not map onto one another, manufacturer names and addresses rather than facility numbers. The FDA's facility lists report the same manufacturer under several spellings and addresses across years, and customs records describe shipments in free text from which drug names and dosage forms must first be extracted. No single source provides a comprehensive historical mapping of injectable drugs to their manufacturing plants, a gap the FDA itself faces when it tries to identify fragile supply chains, and the reason quantitative research on drug shortages has been so scarce.[14]

I link the sources with probabilistic record linkage in four steps. First, I parse each text field into its components and standardize common strings, so that the fields to be matched share a format and carry meaning for matching. Second, I geocode every manufacturing facility in the FDA lists (GDUFA and CDER) and every shipper in the customs data by querying Google Maps with the cleaned name and

[14]To take a simplified example, a manufacturer in India such as Aarti Pharmaceuticals Limited appears in the FDA data under this name but also as "aarti pharmaceuticals", "aarti pharma ltd", "india Aartis drug" or, with typos, "street24aaarti"; suppose it has three plants, each with its own address, each reported in several formats and with typos. The goal is one name per manufacturer and one address per plant. The problem is more acute in the customs data, where a shipment description can run over several paragraphs and contain many typos.

address, which returns a longitude, a latitude and a normalized address for each facility; the geocoded locations give me crosswalks of company and facility names and sharpen the matching of the third step. Third, I match each entry to its closest counterpart with a fuzzy matching algorithm: plants first on their cleaned names, then on their coordinates for the plants without an exact match, grouping facilities with similar names and zip codes that lie within ten kilometers of one another.[15] The same algorithm matches drugs across datasets, where a drug may appear under different names or codes; since codes are not always linked to names, I first match drugs to their National Drug Codes and then the codes to the names. Finally, I review by hand the matched pairs with low similarity scores and the unmatched entries. Similarity scores are correlated with correct matches but are an imperfect metric, so the linkage takes several runs, with different combinations of fields, criteria for choosing candidates and weights.

The linkage combines exact matching with approximate comparisons of names and addresses. Depending on the linkage step, I use modified bigram scores, which measure agreement in consecutive character pairs and reward matching initial characters, or Jaro similarity, which accounts for matching characters and differences in their ordering. For facility matches involving several fields, I weight and combine the match and nonmatch contributions into an overall score. To group variant spellings of the same drug within the Orange Book, I use the Levenshtein edit distance, the number of edits needed to turn one name into another. Separately, I extract drug names from shipment descriptions by searching a dictionary of known names and variants.

Plants and observed shipments. Matching a shipper to an FDA manufacturing establishment links the drug in the shipment to its production location. U.S. export records also supply product–plant links. I combine these links with annual registration records to construct the set of registered plants for a drug in a given year. This measure differs from the foreign suppliers observed shipping the drug in a quarter. In particular, a temporary absence from customs records is not treated as sufficient evidence that a registered plant has ceased to be associated with the drug. Similarly, a product without an identified foreign shipment is not classified as domestic without evidence of domestic production.

F.2.2. Matching Results

Shippers in the U.S. customs data are the manufacturing facilities of the origin country, the plant or its administrative office, which I verify by matching the FDA lists of manufacturing facilities (GDUFA and CDER) to the customs records. Every plant that enters the linkage has a city-level geocode. Of the 11,703 drug-manufacturing plants registered with the FDA between 2000 and 2019, 899 make at least one injectable product of my panel. A plant is missed when its shipper does not declare its name to customs, or when the product travels by air rather than by sea, as the most expensive and perishable products do, vaccines for instance.

[15] Some addresses in Asia carry details such as fax numbers or ownership status, so that the same plant can receive different coordinates or none. For each manufacturer I take the most frequent zip code and the most detailed address and coordinates within it, compute the great-circle distance from every geocoded location of that manufacturer to this main location, and assign the main coordinates to locations within ten kilometers.

Nearly three quarters of the injectable products with positive sales between 2004 and 2019 are traced to at least one plant, 3,662 of 4,947, and 73 percent of the products sold in 2019. I trace nearly every drug on the shortage lists to the plants that make it. Of the 2,038 injectable products reported in shortage between 2004 and 2019, 95 percent are linked to at least one manufacturing plant registered with the FDA, and 93 percent once I set aside the plant matches the record linkage flags as doubtful. Of the linked products, 94 percent have at least one plant outside the United States and 89 percent at least one plant in Asia.

Table F.1 summarizes the product-based matching rates. These rates concern their stated product populations, not the share of all FDA establishments matched to customs records.

Table F.1: **Coverage of the Product–Plant Linkage**

Population and linkage criterion	Coverage
Injectable products in IQVIA, 2004–2019	
Products with positive sales linked to at least one plant	3,662 of 4,947
Share of products sold in 2019 linked to a plant	73%
Injectable products on the UUDIS shortage list, 2004–2019	
Products linked to at least one manufacturing plant	95% of 2,038
Same population, excluding plant links flagged as doubtful	93% of 2,038

Notes: Sources: the product–plant linkage, IQVIA, and UUDIS. Products in this coverage exercise are labeler-products, not molecule–form markets. Each row uses its stated population. Percentages are rounded.