Econometrica Supplementary Material

SUPPLEMENT TO "ON THE EFFECT OF PARALLEL TRADE ON MANUFACTURERS' AND RETAILERS' PROFITS IN THE PHARMACEUTICAL SECTOR"

(*Econometrica*, Vol. 88, No. 6, November 2020, 2503–2545)

PIERRE DUBOIS
Toulouse School of Economics

MORTEN SÆTHRE
Department of Economics, Norwegian School of Economics

APPENDIX B

B.1. *Graphic Description of Vertical Chain*

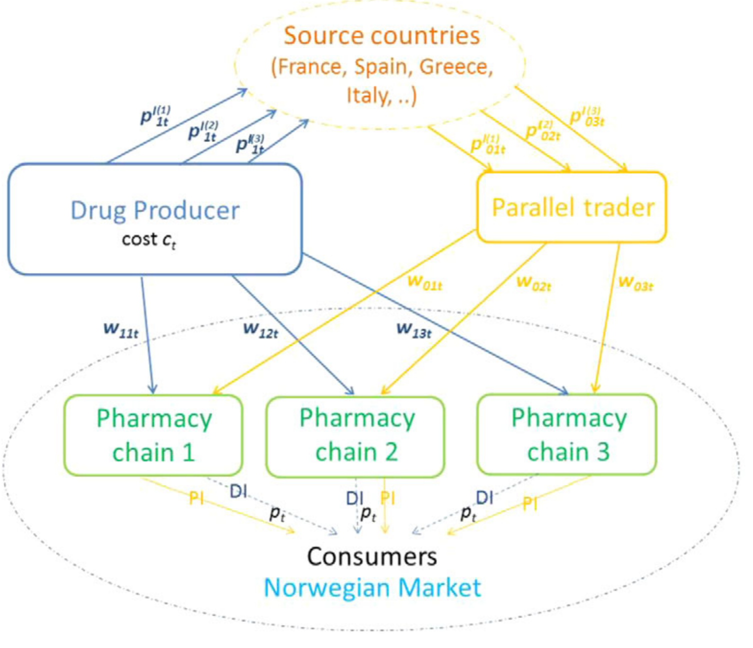

FIGURE B.1.—Vertical chain.

B.2. *Evidence on Switching Behavior Across Chains and Drug Versions*

TABLE B.I SHOWS THE AVERAGE TRANSITION PROBABILITIES by the combination of chain and PI versus DI. It shows that there is much more switching across versions within a chain than switching across chains within a version. This is computed thanks to individual purchases that the anonymous panel allows to follow over time.

Pierre Dubois: pierre.dubois@tse-fr.eu
Morten Sæthre: morten.saethre@nhh.no

 https://doi.org/10.3982/ECTA16039

TABLE B.I

TRANSITION PROBABILITIES ACROSS CHAINS AND DRUG VERSIONS[a]

	Drug Version	Chain 1		Chain 2		Chain 3	
		DI	PI	DI	PI	DI	PI
Chain 1	DI	0.44	0.47	0.04	0.01	0.02	0.03
	PI	0.21	0.69	0.02	0.03	0.02	0.03
Chain 2	DI	0.03	0.04	0.62	0.27	0.02	0.02
	PI	0.02	0.06	0.34	0.54	0.02	0.03
Chain 3	DI	0.04	0.06	0.03	0.02	0.51	0.35
	PI	0.02	0.07	0.02	0.02	0.22	0.64

[a]Notes: Probability to Switch to Purchase Drug Version and Chain in Column Given Last Period Purchase of Version and Chain in Row.

B.3. *Prescription Behavior and Parallel Trade*

One worry for the identification of our model is that doctors will change their prescription behavior if pharmacies induce consumers to consume parallel traded Lipitor more frequently. An example of what we have in mind is that consumers might oppose getting parallel traded drugs, thereby making their doctor prescribe them other types of statins for which there does not exist parallel traded alternatives. Over the sample period, there was an increase in the share of statin prescriptions going to Simvastatin due to new guidelines for statin prescriptions from the Norwegian Medicines Agency and a decrease in the share of statin prescriptions going to atorvastatin (the molecule contained in Lipitor), as shown in Figure B.2. We regard this decrease as a function of the change in policy for statin prescriptions induced by the government, who implemented a lower price cap on simvastatin than atorvastatin, and not necessarily related to the preferences of consumers or doctors for directly imported versus parallel trade drugs.

We investigate the potential endogeneity issues arising from doctors responding to pharmacies strategies for selling parallel traded Lipitor by changing what statin they prescribe. Using data on the prescription behavior of individual doctors, we can look at the share of statin prescriptions going to atorvastatin, together with the behavior of the pharmacies to which each doctor's patients are exposed. This is feasible due to availability of

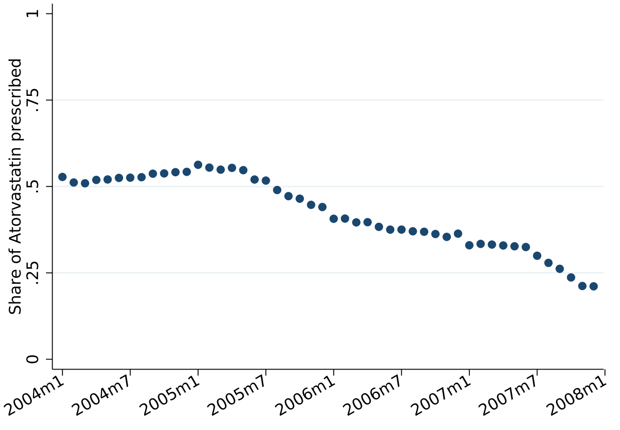

FIGURE B.2.—Physicians' prescription of atorvastatin as share of total statin prescription.

information linking the doctor to the prescription used by a patient for each transaction at each given pharmacy. We use information about the availability of parallel imports (assuming that if a pharmacy did not sell any parallel imports during a month, it means it was not available) and the ratio of margin for parallel and direct imports at a given pharmacy chain. The availability gives a sense of whether the doctor's patient potentially faced foreclosure of direct imports, whereas the margin can be thought of as a reduced-form measure of the pharmacy's decision to foreclose direct imports. We calculate the weighted sum of availability and margin ratio in each chain for each doctor, where the measure is weighted by the share of the doctor's patients patronizing the different chains. More precisely, for doctor d in month t,

$$available_{dt} = \frac{1}{N_{dt}} \sum_{i=1}^{N_{dt}} \mathbf{1}_{\{parallel_{it}\}},$$

where N_{dt} is the number of patients for doctor d in month t, and $\mathbf{1}_{\{parallel_{it}\}}$ is an indicator for whether patient i went to a pharmacy offering parallel traded Lipitor in month t. Similarly,

$$ratio_{dt} = \frac{1}{N_{dt}} \sum_{i=1}^{N_{dt}} \frac{m_{0c(i)t}}{m_{1c(i)t}},$$

where $\frac{m_{0c(i)t}}{m_{1c(i)t}}$ is the ratio of margins for parallel (0) and direct (1) imported Lipitor at the pharmacy chain $c(i)$ visited by patient i in month t. Overall, doctors prescribe Lipitor in 43% of the cases where a statin was prescribed, whereas parallel trade is available for 25% of the patients. The number of unique doctors in our sample is 14,051, who are observed for a maximum of 48 months between January 1, 2004, and December 31, 2007.

Table B.II presents the results of OLS regressions of atorvastatin's share of statin prescriptions on weighted margin ratios and parallel trade availability. The observation unit is a doctor-month. Column (1) shows a large negative coefficient on margin and availability, although this is driven by the overall downward trend in atorvastatin prescriptions,

TABLE B.II

EFFECTS OF MARGINS AND AVAILABILITY OF PARALLEL IMPORTS ON ATORVASTATIN PRESCRIPTION[a]

	(1)	(2)	(3)	(4)	(5)	(6)
$ratio_{dt} * available_{dt}$	−0.052	−0.000	0.003	−0.036	0.005	0.007
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
$ratio_{dt}$	−0.018	−0.010	−0.013	−0.047	−0.022	−0.023
	(0.006)	(0.006)	(0.006)	(0.004)	(0.003)	(0.003)
Time trend		Yes			Yes	
Time FE			Yes			Yes
Physician FE				Yes	Yes	Yes
N	258,281	258,281	258,281	258,281	258,281	258,281
R^2	0.01	0.11	0.13	0.02	0.18	0.20

[a]Note: OLS Regression. Standard Errors Clustered by Doctor. The Dependent Variable Is the Share of Atorvastatin Prescribed by Physician.

together with a tendency for both the margin ratio and the availability of parallel trade to increase over time. This is confirmed by the coefficient on margin ratio going to a quite precisely estimated zero in columns (2) and (3), where we add a linear time trend and time-fixed effects, respectively. When we add doctor-fixed effects together with a time trend or time-fixed effects in columns (5) and (6), we obtain a positive coefficient on the margin ratio and a negative coefficient on availability, both of which are statistically significant. However, considering the size of the coefficients, none of them are economically significant. The coefficient on the margin ratio tells us that the effect of an increase of roughly two standard deviations (the standard deviation of that variable being 0.54), the atorvastatin share of statin prescriptions will increase by roughly one-half percentage point. Similarly for availability, an increase in availability from none to full would yield a decrease in atorvastatin prescriptions by 2.2 percentage points. Considering that the average availability is 25%, this result implies that very large changes in pharmacies behavior is related to relatively small changes in the prescription behavior of doctors in our sample. We thus conclude that we should not be concerned by a potential identification problem due to doctors changing molecule prescriptions in response to pharmacies incentives to sell parallel traded Lipitor more frequently.

B.4. *Details on Derivatives of* $\boldsymbol{\theta}_t^*(\boldsymbol{w}_{0t}, \boldsymbol{w}_{1t})$

In order to obtain how wholesale prices affect the equilibrium $\boldsymbol{\theta}_t^*(\boldsymbol{w}_{0t}, \boldsymbol{w}_{1t})$, let $\mathbf{F}_{\boldsymbol{\theta},t}$ denote the vector of derivatives of pharmacy chain profit with respect to direct import availability at time t, that is,

$$\mathbf{F}_{\boldsymbol{\theta},t} \equiv \left(\frac{\partial \pi_{1t}}{\partial \theta_{1t}}, \frac{\partial \pi_{2t}}{\partial \theta_{2t}}, \dots, \frac{\partial \pi_{Ct}}{\partial \theta_{Ct}} \right)'.$$

Implicit differentiation of the system of first-order conditions $\mathbf{F}_{\boldsymbol{\theta},t} = \mathbf{0}$ yields

$$\left.\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \boldsymbol{\theta}_t'}\right|_{\boldsymbol{\theta}_t = \boldsymbol{\theta}_t^*} \mathrm{d}\boldsymbol{\theta}_t + \left.\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \boldsymbol{w}_{1t}'}\right|_{\boldsymbol{\theta}_t = \boldsymbol{\theta}_t^*} \mathrm{d}\boldsymbol{w}_{1t} = \mathbf{0}.$$

The Jacobian of $\boldsymbol{\theta}_t^*(\boldsymbol{w}_{0t}, \boldsymbol{w}_{1t})$ with respect to $\boldsymbol{w}_{1t}$ is then

$$\frac{\partial \boldsymbol{\theta}_t^*}{\partial \boldsymbol{w}_{1t}'} = -\left(\left.\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \boldsymbol{\theta}_t'}\right|_{\boldsymbol{\theta}_t = \boldsymbol{\theta}_t^*} \right)^{-1} \left.\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \boldsymbol{w}_{1t}'}\right|_{\boldsymbol{\theta}_t = \boldsymbol{\theta}_t^*}.$$

Of course, if some elements of $\boldsymbol{\theta}_t^*$ is not interior, the corresponding elements of $\frac{\partial \boldsymbol{\theta}_t^*}{\partial \boldsymbol{w}_{1t}'}$ will be zero.

Recalling that $\frac{\partial \pi_{ct}}{\partial \theta_{ct}} = (\bar{p}_t - w_{0ct}) \frac{\partial s_{0ct}}{\partial \theta_{ct}}(\boldsymbol{\theta}_t) + (\bar{p}_t - w_{1ct}) \frac{\partial s_{1ct}}{\partial \theta_{ct}}(\boldsymbol{\theta}_t)$, we have that

$$\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \boldsymbol{w}_{1t}'} = - \begin{pmatrix} \frac{\partial s_{11t}}{\partial \theta_{1t}} & 0 & \cdots & 0 \\ 0 & \frac{\partial s_{12t}}{\partial \theta_{2t}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\partial s_{1Ct}}{\partial \theta_{Ct}} \end{pmatrix},$$

while

$$\frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \boldsymbol{\theta}_t'} = \begin{pmatrix} \sum_k m_{k1t} \frac{\partial^2 s_{k1t}}{\partial \theta_{1t}^2} & \sum_k m_{k1t} \frac{\partial^2 s_{k1t}}{\partial \theta_{1t}\,\partial \theta_{2t}} & \cdots & \sum_k m_{k1t} \frac{\partial^2 s_{k1t}}{\partial \theta_{1t}\,\partial \theta_{Ct}} \\ \sum_k m_{k2t} \frac{\partial^2 s_{k2t}}{\partial \theta_{2t}\,\partial \theta_{1t}} & \sum_k m_{k2t} \frac{\partial^2 s_{k2t}}{\partial \theta_{2t}^2} & \cdots & \sum_k m_{k2t} \frac{\partial^2 s_{k2t}}{\partial \theta_{2t}\,\partial \theta_{Ct}} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_k m_{kCt} \frac{\partial^2 s_{kCt}}{\partial \theta_{Ct}\,\partial \theta_{1t}} & \sum_k m_{kCt} \frac{\partial^2 s_{kCt}}{\partial \theta_{Ct}\,\partial \theta_{2t}} & \cdots & \sum_k m_{kCt} \frac{\partial^2 s_{kCt}}{\partial \theta_{Ct}^2} \end{pmatrix}.$$

Then all the derivatives of market shares with respect to w_{1ct} in equation (3.5) can be obtained from elements of the stacked vector $\frac{\partial s_{kt}}{\partial \boldsymbol{w}_{1t}'}$ for $k = 0$ or 1 and which satisfies

$$\frac{\partial \boldsymbol{s}_{kt}}{\partial \boldsymbol{w}_{1t}'} = \left(\frac{\partial \boldsymbol{s}_{kt}}{\partial \boldsymbol{\theta}_t'} \bigg|_{\boldsymbol{\theta}_t = \boldsymbol{\theta}_t^*} \right) \frac{\partial \boldsymbol{\theta}_t^*}{\partial \boldsymbol{w}_{1t}'} = \left(\frac{\partial \boldsymbol{s}_{kt}}{\partial \boldsymbol{\theta}_t'} \bigg|_{\boldsymbol{\theta}_t = \boldsymbol{\theta}_t^*} \right) \left(- \frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \boldsymbol{\theta}_t'} \bigg|_{\boldsymbol{\theta}_t = \boldsymbol{\theta}_t^*} \right)^{-1} \frac{\partial \mathbf{F}_{\boldsymbol{\theta},t}}{\partial \boldsymbol{w}_{1t}'} \bigg|_{\boldsymbol{\theta}_t = \boldsymbol{\theta}_t^*},$$

which shows that the change in a given market share, s_{kct}, caused by the change in a given wholesale price, $w_{1\tilde{c}t}$, will depend on the change in the full vector of θ's following from the change in the Nash equilibrium in the competition between chains.

B.5. *Asymptotic Distribution of θ_{ct} Estimates*

In the case of interior solution to the Nash equilibrium in θ_{ct}, θ_{ct} must satisfy first-order condition (3.2) given other $\theta_{\tilde{c}t}$. Let us denote $\theta_{ct}^u(\boldsymbol{\beta})$ the solution to the first-order condition whether it belongs to the [0,1] interval or not. Then we know that the solution of the Nash equilibrium is $\theta_{ct}(\boldsymbol{\beta}) = \theta_{ct}^u(\boldsymbol{\beta})\mathbf{1}_{\{\theta_{ct}^u(\boldsymbol{\beta}) \in (0,1)\}} + \mathbf{1}_{\{\theta_{ct}^u(\boldsymbol{\beta}) \geq 1\}}$. Using the Delta method, we can first find the asymptotic law of $\theta_{ct}^u(\boldsymbol{\beta})$. We need the gradient of $\theta_{ct}^u(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$. Fully differentiating the first order condition determining $\theta_{ct}^u(\boldsymbol{\beta})$, we obtain for all c:

$$\sum_{c'} \left(m_{0ct} \frac{\partial^2 s_{0ct}}{\partial \theta_{ct}^u\,\partial \theta_{c't}^u} + m_{1ct} \frac{\partial^2 s_{1ct}}{\partial \theta_{ct}^u\,\partial \theta_{c't}^u} \right) \frac{\partial \theta_{c't}^u(\boldsymbol{\beta})}{\partial \beta} + m_{0ct} \frac{\partial^2 s_{0ct}}{\partial \theta_{ct}^u\,\partial \beta} + m_{1ct} \frac{\partial^2 s_{1ct}}{\partial \theta_{ct}^u\,\partial \beta} = 0,$$

where

$$\begin{aligned} \frac{\partial^2 s_{0ct}}{\partial \theta_{ct}^u\,\partial \theta_{c't}^u} &= \int -\rho_{ict} \frac{\partial s_{ict}}{\partial \theta_{c't}^u} - \mathbf{1}_{\{c=c'\}} \rho_{ict} \delta_{ict} s_{ict} (1 - s_{ict}) \\ &\quad + \left(1 - \theta_{ct}^u \rho_{ict}\right) \delta_{ict} [1 - 2 s_{ict}] \frac{\partial s_{ict}}{\partial \theta_{c't}^u}\, dF(V_{it} | \boldsymbol{\beta}), \\ \frac{\partial^2 s_{1ct}}{\partial \theta_{ct}^u\,\partial \theta_{c't}^u} &= \int \rho_{ict} \frac{\partial s_{ict}}{\partial \theta_{c't}^u} + \mathbf{1}_{\{c=c'\}} \rho_{ict} \delta_{ict} s_{ict} (1 - s_{ict}) \\ &\quad + \theta_{ct}^u \rho_{ict} \delta_{ict} [1 - 2 s_{ict}] \frac{\partial s_{ict}}{\partial \theta_{c't}^u}\, dF(V_{it} | \boldsymbol{\beta}) \end{aligned}$$

with

$$\frac{\partial s_{ict}}{\partial \theta_{c't}^u} = \delta_{ic't} (\mathbf{1}_{\{c=c'\}} - s_{ict}) s_{ic't}$$

and $\frac{\partial^2 s_{1ct}}{\partial\theta^u_{ct}\partial\beta}$ and $\frac{\partial^2 s_{0ct}}{\partial\theta^u_{ct}\partial\beta}$ come from taking derivatives with respect to β of

$$\frac{\partial s_{0ct}}{\partial\theta^u_{ct}} = \int \big(-\rho_{ict}s_{ict} + \big(1-\theta^u_{ct}\rho_{ict}\big)\delta_{ict}s_{ict}(1-s_{ict})\big)\,dF(V_{it}|\boldsymbol{\beta}), \quad \text{and}$$

$$\frac{\partial s_{1ct}}{\partial\theta^u_{ct}} = \int \big(\rho_{ict}s_{ict} + \theta^u_{ct}\rho_{ict}\delta_{ict}s_{ict}(1-s_{ict})\big)\,dF(V_{it}|\boldsymbol{\beta}).$$

Then we know that $\hat{\theta}^u_{ct}(\boldsymbol{\beta}) \hookrightarrow N(\theta^u_{ct}(\hat{\boldsymbol{\beta}}), \operatorname{var}(\hat{\theta}^u_{ct}(\boldsymbol{\beta})))$ with $\operatorname{var}(\hat{\theta}^u_{ct}(\hat{\boldsymbol{\beta}})) = [\frac{\partial\boldsymbol{\theta}^u_t(\hat{\boldsymbol{\beta}})}{\partial\beta}]' \operatorname{var}(\hat{\boldsymbol{\beta}}) \times [\frac{\partial\boldsymbol{\theta}^u_t(\hat{\boldsymbol{\beta}})}{\partial\beta}]$ where $[\frac{\partial\boldsymbol{\theta}^u_t(\boldsymbol{\beta})}{\partial\beta}]$ is the Jacobian matrix of $\boldsymbol{\theta}^u_t(\boldsymbol{\beta}) = (\theta^u_{1t}(\boldsymbol{\beta}), \dots, \theta^u_{Ct}(\boldsymbol{\beta}))'$ with respect to the vector of parameters $\boldsymbol{\beta}$, and $\operatorname{var}(\hat{\boldsymbol{\beta}})$ is the variance-covariance matrix of $\hat{\boldsymbol{\beta}}$.

As $\theta_{ct}(\boldsymbol{\beta}) = \theta^u_{ct}(\boldsymbol{\beta})\mathbf{1}_{\{\theta^u_{ct}(\boldsymbol{\beta})\in(0,1)\}} + \mathbf{1}_{\{\theta^u_{ct}(\boldsymbol{\beta})\geq 1\}}$, we obtain directly the asymptotic law of $\theta_{ct}(\boldsymbol{\beta})$ using the one of $\theta^u_{ct}(\boldsymbol{\beta})$. θ_{ct} is censored normally distributed. With ϕ the $N(0,1)$ c.d.f., the c.d.f. of $\hat{\theta}_{ct}(\hat{\boldsymbol{\beta}})$ is $P(\hat{\theta}_{ct}(\hat{\boldsymbol{\beta}}) \leq a) = \phi(\frac{a-\hat{\theta}^u_{ct}(\hat{\boldsymbol{\beta}})}{\operatorname{var}(\hat{\theta}^u_{ct}(\boldsymbol{\beta}))})\mathbf{1}_{\{a\in(0,1)\}} + \mathbf{1}_{\{a\geq 1\}}$ which allows construct the confidence interval of $\theta_{ct}(\boldsymbol{\beta})$.

B.6. *Inner Loop Algorithm of Demand Estimation*

In each period t, given the other chains choices for $\theta_{\tilde{c}t}$ ($\tilde{c} \neq c$), each pharmacy chain c solves the constrained maximization problem:

$$\max_{\theta_{ct}} \pi_{ct} \quad \text{s.t.} \quad 0 \leq \theta_{ct} \leq 1.$$

Letting μ^L_{ct} and μ^H_{ct} denote the multipliers associated with the lower and upper bound on θ_{ct} respectively, the necessary conditions for maximization of the corresponding Lagrangian are

$$\begin{aligned} &\frac{\partial\pi_{ct}}{\partial\theta_{ct}} + \mu^L_{ct} - \mu^H_{ct} = 0, \\ &\mu^L_{ct} \geq 0, \qquad \mu^L_{ct}\theta_{ct} = 0, \qquad \theta_{ct} \geq 0, \\ &\mu^H_{ct} \geq 0, \qquad \mu^H_{ct}(1-\theta_{ct}) = 0, \qquad \theta_{ct} \leq 1. \end{aligned}$$

The equilibrium in each period t is given by the solution to these equations for each chain c. This equilibrium can be redefined as the solution to the following constrained minimization problem:

$$\begin{aligned} &\min_{\{\theta_{ct},\mu^L_{ct},\mu^H_{ct}\}_{c\in\{1,\dots,C\}}} \sum_{c\in\{1,\dots,C\}} \big(\mu^L_{ct}\theta_{ct} + \mu^H_{ct}(1-\theta_{ct})\big) \\ &\text{s.t.} \\ &\frac{\partial\pi_{ct}}{\partial\theta_{ct}} + \mu^L_{ct} - \mu^H_{ct} = 0 \quad \forall c \in \{1,\dots,C\}, \\ &0 \leq \theta_{ct} \leq 1, \qquad \mu^L_{ct} \geq 0, \qquad \mu^H_{ct} \geq 0 \quad \forall c \in \{1,\dots,C\}. \end{aligned}$$

The objective function in this minimization problem is the sum of the complementary slackness condition corresponding to the bounds on θ for each chain, and will thus be zero at the solution, while the constraints ensure that the solution is a Nash-equilibrium.

The full problem in the inner loop of the estimation also includes fitting the mean utility parameters α_{jct} for each product j at each chain c in each period t. These mean utility parameters are set such that observed shares $\hat{s}_{jct}$ are equal to predicted shares $s_{jct}(\boldsymbol{\theta}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta})$, where $\boldsymbol{\theta}_t$ is the vector of θ_{ct} for all chains c, and $\boldsymbol{\alpha}_t$ is the vector of mean utility parameters for all products in period t. We can write this restriction as

$$\boldsymbol{s}_t(\boldsymbol{\theta}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta}) = \hat{\boldsymbol{s}}_t,$$

where $\boldsymbol{s}_t(\boldsymbol{\theta}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta})$ is the vector of predicted market shares from the model, and $\hat{\boldsymbol{s}}_t$ is the vector of observed market shares. These conditions can then be included as constraints in the minimization problem characterizing the market equilibrium in period t.

The full constrained minimization problem can then be written (in vector notation)

$$\begin{aligned} &\min_{\boldsymbol{\alpha}_t, \boldsymbol{\theta}_t, \boldsymbol{\mu}_t^L, \boldsymbol{\mu}_t^H} \boldsymbol{\mu}_t^L \cdot \boldsymbol{\theta}_t + \boldsymbol{\mu}_t^H \cdot (\mathbf{1} - \boldsymbol{\theta}_t) \\ &\text{s.t.} \\ &\boldsymbol{s}_t(\boldsymbol{\theta}_t, \boldsymbol{\alpha}_t, \boldsymbol{\beta}) = \hat{\boldsymbol{s}}_t, \\ &\frac{\partial \boldsymbol{\pi}_t}{\partial \boldsymbol{\theta}_t} + \boldsymbol{\mu}_t^L - \boldsymbol{\mu}_t^H = \mathbf{0}, \\ &\mathbf{0} \leqq \boldsymbol{\theta}_t \leqq \mathbf{1}, \qquad \boldsymbol{\mu}_t^L \geqq \mathbf{0}, \qquad \boldsymbol{\mu}_t^H \geqq \mathbf{0}. \end{aligned}$$

Informally, we can think about the market share constraint as particularly informative about the mean utility parameters $\boldsymbol{\alpha}_t$, while the constraints corresponding to first-order conditions for profit maximization are particularly informative about $\boldsymbol{\theta}_t$, though in practice they will jointly inform all parameters. Note that observed margins only enter each chain's profit maximization problem (i.e., it does not have a direct effect on demand), which serves as an exclusion restriction in our model.

The solution to this constrained minimization program is computed for each market t using a sequential quadratic method.[18] In our empirical estimates, we find a solution satisfying all constraints with the minimized value equal to 0 for all markets. As further checks on the solutions of the inner loop, we perform several additional tests (at the parameter vector estimated in the outer loop). One is a check of the second- order conditions of the firms' maximization problems, to verify that the θ's constitute a maximum. Another is a check of whether a firm would profit by unilaterally deviating by setting θ to one of the corners (if θ is already at a corner, only the other is tested), as a test of the Nash equilibrium. Also, we perform two tests for multiple equilibria. First, we recalculate the solution to the system of equations given above for the cases where we fix a firm's θ at each of the corners (for each firm separately), thus removing the constraint corresponding to this firm's profit maximization problem. We then check whether any of the solutions satisfies the full set of equations. Second, we solve the system of equations for many different starting values, checking whether we obtain nonunique solutions.

B.7. *Pharmacy Retail Pricing With Price Ceiling*

Here, we show that a pharmacy chain offering two goods, PI ($j = 0$) and DI ($j = 1$), subject to a common price ceiling $\bar{p}$ will sometimes choose to price both goods at the

[18]See, for example, Judd (1998, Chapter 4.7). Specifically, we use sequential least squares programming (SLSQP), as implemented in the optimization routines of the Python package SciPy.

price ceiling, even if consumers have a preference for one of the two. Let us assume that consumers have a preference for DI, such that PI will be bought to a lower extent if prices and availability are equal. It can be shown that the chosen prices will both sometimes be at the price ceiling and that the extent of pharmaceutical coverage and "tightness" of the price ceiling will make this even more likely.

Let the demand for each good j at pharmacy c be given by $q_{jc}(p_{0c}, p_{1c}, p_{0-c}, p_{1-c})$, where p_{jc} is the price paid by the consumer for good j in pharmacy c. The price set by the firm, r_{jc} is related to the price paid by the consumer through $p_{jc} = \tau r_{jc}$, where $0 \leq \tau \leq 1$ is the co-payment rate. The profits of pharmacy chain c is given by

$$\pi_c = q_{0c}(r_{0c} - w_{0c}) + q_{1c}(r_{1c} - w_{1c}),$$

where w_{jc} is the pharmacy chain's wholesale price for good j. In a Nash equilibrium, given prices in other chains, the pharmacy chain solves the problem:

$$\max_{r_{0c}, r_{1c}} \pi_c \quad \text{s.t.} \quad r_{0c}, r_{1c} \leq \bar{p},$$

with the corresponding complementary slackness conditions for each $j \in \{0, 1\}$

$$q_{jc} + \tau \frac{\partial q_{0c}}{\partial p_{jc}}(r_{0c} - w_{0c}) + \tau \frac{\partial q_{1c}}{\partial p_{jc}}(r_{1c} - w_{1c}) \geq 0, \quad r_{jc} \leq \bar{p}.$$

Assume that the price ceiling is sufficiently low to bind for good 1 ($r_{1c} = \bar{p}$), which is the one which consumers value the most and will command the highest price in the absence of the price ceiling. To see that the pharmacy could find it optimal to price at the ceiling also for the other product, note that the unconstrained price for good 0 in this case would be

$$r_{0c}^* = w_{0c} + \frac{q_{0c}}{-\tau \dfrac{\partial q_{0c}}{\partial p_{0c}}} + \frac{\dfrac{\partial q_{1c}}{\partial p_{0c}}}{-\dfrac{\partial q_{0c}}{\partial p_{0c}}}(\bar{p} - w_{1c}).$$

It is straightforward to see that r_{0c}^* could exceed $\bar{p}$ if the price ceiling is tight enough. From the second term, we see that the lower the copayment rate τ, the less responsive consumers are to any change in the retail price p_{0c}, thus increasing the optimal unconstrained price r_{0c}^*. A lower price ceiling will tend to reduce r_{0c}^* through the reduced sales of 0, since the price of good 1 becomes lower, and through reducing the profit margin on good 1, which lowers the value of the diverted sales to good 1 with an increase in the price of good 0, but unless the price of good 0 responds too much to a change in the price of good 1 (i.e., the slope of the "reaction function" is too large), it will be possible for both goods to be constrained by the price ceiling simultaneously.

B.8. *Outside Option in the Demand for Atorvastatin in Norway*

Concerning the importance of the outside option of not purchasing the drug, as we cannot observe if some patients are prescribed atorvastatin but are not buying the drug. We thus investigate whether improved access to pharmacies increases or not the purchases of the drug. In a country like Norway with large health insurance coverage in the case of treatment for cholesterol control, we can expect that only a very small share of the population in need is not taking the treatment. However, we evaluate the effect of improved

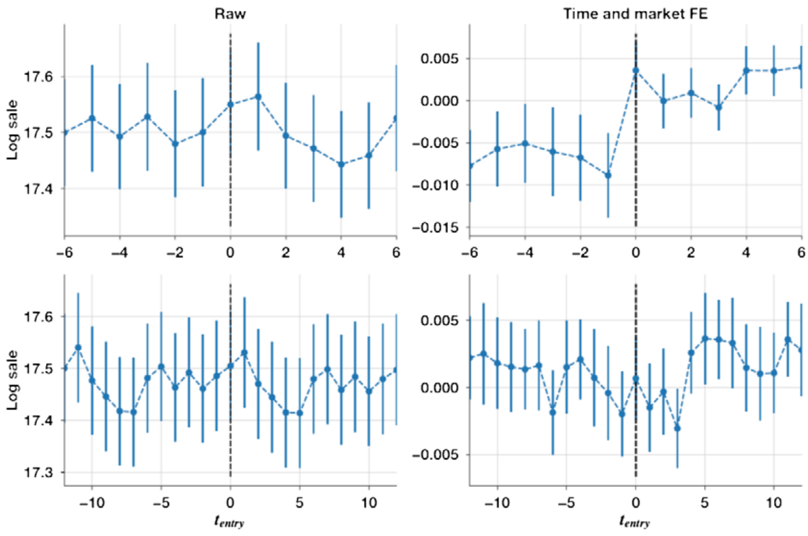

FIGURE B.3.—Effect of entry of a pharmacy outlet on sales in a given area.

access which should affect the value of an outside option by measuring the changes in sales of atorvastatin in each predefined area when one pharmacy enters the market. We regress the log sales in a given area on monthly dummies relative to the date of entry of a pharmacy and find no significant effect of entry on sales, showing that it is unlikely that the outside option of not purchasing a drug is very important. Graph B.3 shows that this change is very small and almost zero statistically and economically.

On the left panel, the top graph shows the log sales quantity in a given area before and after entry of a pharmacy outlet with a 6 weeks window. The bottom graph of the left panel shows the same with a 10- week window. Vertical bars correspond to the standard deviation of log sales across areas at each week. On the right panel, we have the same means and standard deviations estimates after removing market and time fixed effects. We can see that the effect of entry on a 10-week window seem null and if slightly positive on a 6-week window, its magnitude is less than 1% of sales quantities.

B.9. *Details on Counterfactual Computations When Foreclosure Is Banned*

In the case of no foreclosure, the bargaining equations are

$$\frac{\partial \ln \Delta_1 \pi_{ct}}{\partial w_{1ct}} = \frac{1-b_{1c}}{b_{1c}} \frac{\partial \ln \Delta_c \Pi_t}{\partial w_{1ct}} \quad \text{and} \quad \frac{\partial \ln \Delta_0 \pi_{ct}}{\partial w_{0ct}} = \frac{1-b_{0c}}{b_{0c}} \frac{\partial \ln \Delta_c \Pi_t^{\mathrm{PI}}}{\partial w_{0ct}},$$

where

$$\Delta_c \Pi_t = \sum_{\tilde{c}} \big(w_{1\tilde{c}t} \Delta_{1c} s_{1\tilde{c}t} + p_{1t}^{I(\tilde{c})} \Delta_{1c} s_{0\tilde{c}t} \big),$$

$$\Delta_c \Pi_t^{\mathrm{PI}} = \sum_{\tilde{c} \neq c} \big(w_{0\tilde{c}t} - p_{0\tilde{c}t}^{I(\tilde{c})} \big) \Delta_{0c} s_{0\tilde{c}t},$$

$$\Delta_1 \pi_{ct} = (\bar{p}_t - w_{1ct}) s_{1ct} + (\bar{p}_t - w_{0ct}) \Delta_{1c} s_{0ct},$$

$$\Delta_0 \pi_{ct} = \pi_{ct} - \pi_{-0,ct} = (\bar{p}_t - w_{1ct}) \Delta_{0c} s_{1ct} + (\bar{p}_t - w_{0ct}) s_{0ct}.$$

where market shares are integrals of individual choice probabilities computed as

$$s_{ikct} = \frac{\left(\sum_{k'\in\{0,1\}} e^{V_{ik'ct}/\lambda_c}\right)^{\lambda_c - 1} e^{V_{ikct}/\lambda_c}}{\sum_{\tilde{c}}\left(\sum_{k'\in\{0,1\}} e^{V_{ik'\tilde{c}t}/\lambda_{\tilde{c}}}\right)^{\lambda_{\tilde{c}}}} \tag{B.1}$$

and where

$$\Delta_{1c} s_{ik\tilde{c}t} = s_{ik\tilde{c}t} - \frac{\left(\sum_{k'\in\{0,1\}} e^{V_{ik'\tilde{c}t}/\lambda_{\tilde{c}}}\right)^{\lambda_{\tilde{c}} - 1} e^{V_{ik\tilde{c}t}/\lambda_{\tilde{c}}}}{e^{V_{i0ct}} + \sum_{c'\neq c}\left(\sum_{k'\in\{0,1\}} e^{V_{ik'c't}/\lambda_{c'}}\right)^{\lambda_{c'}}} \quad \text{if } \tilde{c} \neq c,$$

$$\Delta_{1c} s_{i1ct} = s_{i1ct} \quad \text{and} \quad \Delta_{1c} s_{i0ct} = s_{i0ct} - \frac{\left(\sum_{k'\in\{0,1\}} e^{V_{ik'ct}/\lambda_c}\right)^{\lambda_c - 1} e^{V_{i0ct}/\lambda_{\tilde{c}}}}{e^{V_{i0ct}} + \sum_{c'\neq c}\left(\sum_{k'\in\{0,1\}} e^{V_{ik'c't}/\lambda_{c'}}\right)^{\lambda_{c'}}}.$$

B.10. *Demand and Price Setting Model for Statins in France*

We use quarterly data on sales volumes and values of statin drugs in France and the US from IMS Health (now called IQVIA) for the years 2004 to 2007. The French data provide sales volumes and values for the retail sector and hospital sector. We model the demand for the retail sector which is the one for which the French regulation plays a role. We use the US retails sales to construct instrumental variables for prices in France. Our random coefficient logit model for French statin demand assume that, for any product a, the random utility for consumer i in France is

$$U_{iat} = \beta_p^i p_{at} + \beta_g g_{at} + \beta_a + \beta_t + \xi_{at} + \varepsilon_{iat},$$

where p_{at} is the price of statin a at time t, β_a is a molecule-strength fixed effect, g_{at} is a dummy variable indicating if the molecule of product a has lost patent exclusivity (changes over time), β_t are quarter specific fixed effects and ξ_{at} is an unobserved demand shock.

Specifying the random coefficient distribution of β_p^i as normal $N(\beta_p, \sigma_p^2)$, and ε_{iat} as extreme value distributed, we obtain the usual market shares

$$s_{at} = \int \frac{\exp(\beta_p^i p_{at} + \beta_g g_{at} + \beta_a + \beta_t + \xi_{at})}{1 + \sum_{\tilde{a}} \exp(\beta_p p_{\tilde{a}t} + \beta_g^i g_{\tilde{a}t} + \beta_{\tilde{a}} + \beta_t + \xi_{\tilde{a}t})} \, dF(\beta_p^i).$$

We estimate this model for the retail pharmaceutical market in France using the usual instrumental variables methods. As excluded instruments, we use the retail prices of drugs in the US and the number of drugs in the corresponding markets in the US. Table B.III show the estimation results and that heterogeneity in the price coefficient matters and is precisely estimated.

TABLE B.III
DEMAND MODEL FOR STATINS IN FRANCE

		BLP	
		Coef.	(Std Err.)
Patent loss of exclusivity dummy variable	β_g	0.0336	(0.2603)
Price	β_p	-2.3291	(0.2346)
	σ_p	1.2651	(0.4311)
Quarter fixed effects		Yes	
Molecule-strength fixed effects		Yes	
N		322	

The demand model is estimated using the 20 mg, 40 mg, and 80 mg markets. In France, there are two molecules for the statin 80 mg market (atorvastatin and fluvastatin), four molecules for the 40 mg market (atorvastatin, fluvastatin, pravavastatin, simvastatin), and five for 20 mg (includes rosuvastatin in addition).

Then, the first-order bargaining equation (A.5) is used to recover marginal cost, but only for branded statins as the prices of generics is fixed by regulation as a percentage of the branded version of the drug (Dubois and Lasio (2018)). This percentage was 60% during the time period of our data, and descriptive statistics on prices show that it was strictly implemented. For the counterfactual equilibrium prices of generics, we assume the same regulatory rule still applies and, therefore, the prices of generics of a molecule will be set as 60% of the bargained price of the branded product.

This implies that, in the equilibrium condition to be satisfied by the counterfactual prices, the semi-elasticity of demand and semielasticity of welfare account for this constraint ($\frac{dp_{g(a)t}}{dp_{at}} = 0.6$), that is,

$$\frac{\partial \ln q_{at}(\overrightarrow{p}_t^I)}{\partial p_{at}} = \left.\frac{\partial \ln q_{at}(\overrightarrow{p}_t^I)}{\partial p_{at}}\right|_{p_{g(a)t}} + \frac{\partial \ln q_{at}(\overrightarrow{p}_t^I)}{\partial p_{g(a)t}} \underbrace{\frac{dp_{g(a)t}}{dp_{at}}}_{=0.6},$$

$$\frac{\partial \ln \Delta_a W_{\mathrm{FR}}(\overrightarrow{p}_t^I)}{\partial p_{at}} = \left.\frac{\partial \ln \Delta_a W_{\mathrm{FR}}(\overrightarrow{p}_t^I)}{\partial p_{at}}\right|_{p_{g(a)t}} + \frac{\partial \ln \Delta_a W_{\mathrm{FR}}(\overrightarrow{p}_t^I)}{\partial p_{g(a)t}} \underbrace{\frac{dp_{g(a)t}}{dp_{at}}}_{=0.6},$$

where $g(a)$ denotes the generic version of drug a, and $\frac{\partial \ln q_{at}(\overrightarrow{p}_t^I)}{\partial p_{g(a)t}}$ denotes the cross price semielasticity of q_{at} with respect to the price of the generic version $p_{g(a)t}$.

To identify the bargaining parameter in France, we minimize the residual sum of squares

$$\sum_{a,t} \bigl(c_{at}(b) - \delta_a - \delta_s - \delta_t\bigr)^2$$

with respect to b, δ_a, a molecule-fixed marginal cost term; δ_s, a strength-fixed term; and δ_t, a quarter-fixed term, where $c_{at}(b) = p_{at} + \frac{1}{\frac{\partial \ln q_{at}(\overrightarrow{p}_t^I)}{\partial p_{at}} + \frac{1-b}{b}\frac{\partial \ln \Delta_a W_{\mathrm{FR}}(\overrightarrow{p}_t^I)}{\partial p_{at}}}$.

To find the new equilibrium for a counterfactual policy in Norway, we simultaneously solve the first-order conditions of other statins wholesale prices in France given by

$$p_{at} = c_{at} - \frac{1}{\dfrac{\partial \ln q_{at}(\overrightarrow{p}_t^I)}{\partial p_{at}} + \dfrac{1-b_{\mathrm{FR}}}{b_{\mathrm{FR}}} \dfrac{\partial \ln \Delta_a W_{\mathrm{FR}}(\overrightarrow{p}_t^I)}{\partial p_{at}}} \quad \text{for all } a \text{ except atorvastatin,}$$

where the right-hand side of the equation has all statin prices in France as argument ($\overrightarrow{p}_t^I$), including atorvastatin (Lipitor).

B.11. *Entry Decisions of PI*

We show indeed that when the wholesale prices in potential source countries decrease, there is entry of PI, which makes sense as it is when the procurement price of PI is lower that it is more profitable to have PI in Norway. We show this across all ATC5 therapeutic classification molecules-strength products for which there is no generics yet but for which the same product exists in one of the source countries during some time in the 2004–2007 period for which we have the Norwegian data. We denote entry of parallel imports in market t for ATC code j in any of the pharmacist chain in Norway as $y_{jt} = 1$ if entry and $y_{jt} = 0$ otherwise. We estimate a logit model of entry as a function of the retail price

$$P(y_{jt} = 1) = \frac{\exp\left(\alpha_j + p_{jt}\beta + \sum_s w_{jst}\beta_s\right)}{1 + \exp\left(\alpha_j + p_{jt}\beta + \sum_s w_{jst}\beta_s\right)},$$

where p_{jt} is the observed mean retail price across the three chains for product j, and w_{jst} is the wholesale price of product j in country s. We estimate the simple logit where $\alpha_j = \alpha$ is common across ATC codes and also the fixed effect logit with unrestricted α_j. Table B.IV shows the results for the 298 products. Column (2) shows the results with fixed effects on fewer observations because all markets for which there is no variation over time or strength of parallel imports presence are dropped because of the ATC5 fixed effect. The results show that the higher is the retail price in Norway of the product and the lower is the wholesale price in source countries the more likely is the entry of parallel imports. The only positive effect of wholesale price in France when not controlling for ATC5 fixed effect disappears in the second column with ATC5 fixed effects.

We also estimate the same logit model at the product-chain level, with the following entry probability of parallel imports in chain c:

$$P(y_{jct} = 1) = \frac{\exp\left(\alpha_j + \alpha_c + p_{jct}\beta + \sum_s w_{jst}\beta_s\right)}{1 + \exp\left(\alpha_j + \alpha_c + p_{jct}\beta + \sum_s w_{jst}\beta_s\right)},$$

where p_{jct} is the observed retail price in chain c for product j. Table B.V shows also that the higher is the retail price of the product in Norway and the lower is the wholesale price in source countries the more likely is the entry of parallel imports.

TABLE B.IV

PROBABILITY OF MARKET ENTRY OF PARALLEL IMPORT

Logit Model	(1)	(2)
Retail price Norway	0.038	0.056
	(0.003)	(0.012)
Wholesale prices source countries (in NOK)		
Czech Republic	−0.302	0.031
	(0.028)	(0.072)
France	−0.040	−1.463
	(0.032)	(0.123)
Greece	−0.265	−0.805
	(0.032)	(0.080)
Poland	0.071	−0.082
	(0.022)	(0.089)
Spain	0.255	−0.061
	(0.039)	(0.090)
UK	−0.127	0.040
	(0.022)	(0.044)
ATC5 fixed effects	No	Yes
Year-Month fixed effects	Yes	Yes
N	8554	4179

TABLE B.V

PROBABILITY OF MARKET CHAIN ENTRY OF PARALLEL IMPORT

Logit Model	(1)	(2)
Retail price Norway	0.019	0.029
	(0.002)	(0.004)
Wholesale prices source countries (in NOK)		
Czech Republic	−0.234	−0.041
	(0.019)	(0.042)
France	−0.064	−1.054
	(0.022)	(0.064)
Greece	−0.307	−0.470
	(0.022)	(0.045)
Poland	0.171	0.052
	(0.016)	(0.050)
Spain	0.143	−0.111
	(0.026)	(0.058)
UK	−0.125	−0.008
	(0.015)	(0.029)
ATC5 fixed effects	No	Yes
Year-Month fixed effects	Yes	Yes
Chain fixed effects	Yes	Yes
N	25,222	12,489

B.12. *Effects of Counterfactual Policies on Consumer Welfare*

Aggregation of consumer welfare across individuals relies on transforming the welfare changes in monetary terms. As we do not identify the price sensitivity of Norwegian consumers in the data, we can aggregate individual changes only under the assumption that there is no heterogeneity of price sensitivity and we cannot transform the changes in monetary terms. We however are able to provide distributions of individual relative changes. Results on consumer welfare[19] show that the share of consumers who would be better off with a ban on parallel imports is of 39% on average across markets (dosage–month) with some variation across markets such that the median across markets of the percentage of consumers who are better off is 34%. Thus, while banning one version of the product removes variety, a substantial portion of consumers are better off on average because most consumers prefer direct imports to parallel imports. In the case of banning foreclosure, effects are small and not reported.

B.13. *General Bargaining Model*

We show here that we can rewrite the bargaining stage model with both θ^0_{ct} and θ^1_{ct} meaning that we do not impose which drug will have a θ at a corner in the bargaining model (while we know which one will be zero in equilibrium, which makes the model simpler to present as we do in the main part of the paper).

When the manufacturer bargains over wholesale price, if he proposes a low enough price to the pharmacy chain, it may not be profitable for the parallel trader to enter and in that case the pharmacy chain has no other supply channel. This changes the pharmacy chain profit function which cannot use parallel imports as an alternative. At equilibrium, we are never in the region of low enough manufacturer wholesale price $w_{\mathbf{1ct}}$ such that the parallel imports are not possible.[20] Thus, the Nash surplus of the pharmacy chain can be written in the case where, at the margin, it remains possible to use parallel imports and the Nash surplus of the manufacturer still needs to take into account the fact that the retailer can use parallel imports. We thus can write the Nash surplus of the manufacturer and the retailer as with both PI and DI (checking empirically that the wholesale prices are high enough).

Then, which θ is at a corner will depend on the wholesale prices resulting from the bargaining game. We show that the bargaining equations look similar, but the derivatives of demands with respect to wholesale prices have more complicated expressions even if they boil down to the one we have in the paper when wholesale prices of PI are lower than the ones of DI.

We denote by θ^0_{ct} and θ^1_{ct} the probabilities that the choice sets are {PI} or {DI}, respectively, and thus $1-\theta^0_{ct}-\theta^1_{ct}$ the probability that the choice set is $B=\{\text{DI}, \text{PI}\}$. The utility of consumer i is given by

$$u_{ikct} = V_{ikct} + \varepsilon_{ict} + \lambda_c \epsilon_{ikct},$$

[19] Our consumer welfare measure is based on expected utility that is $E[u_{ikct}] = \sum_{k,c} s_{ikct} V_{ikct}$ where s_{ikct} is the choice probability of the appropriate case. In the current situation s_{ikct} is given by equation (4.1), in the case of a parallel trade ban, it is given by equation (5.1), in the case of no foreclosure it is given by equation (5.3). Consumer welfare change is the change of this expected utility between two situations.

[20] In fact, this would bring much lower profit to the manufacturer as to prevent parallel trade it would have to sell in Norway at almost lower price than any other source country price. Since other countries have low price (e.g., France can negotiate lower price with the manufacturer because it has a large demand) it is more profitable to sell at higher price in Norway even if it induces parallel trade.

where V_{ikct} is the mean utility consumer i obtains from choosing the drug of origin k in pharmacy chain c in market t and ε_{ict} and ϵ_{ikct} are chain-specific and product-specific sequentially observed shocks, respectively.

The probability that consumer i chooses $k \in \{0, 1\}$ conditional on choice of pharmacy chain c when both products are available in the pharmacy is given by

$$s_{ikt|c,B} = \frac{e^{V_{ikct}/\lambda_c}}{e^{V_{i0ct}/\lambda_c} + e^{V_{i1ct}/\lambda_c}} = \frac{1}{1 + e^{V_{ikct}/\lambda_c - V_{ik'ct}/\lambda_c}} \quad \text{with } k' = 1 - k$$

because ϵ_{ikct} is i.i.d. extreme value distributed.

The choice probability of product k conditional on the choice of pharmacy c is then

$$s_{ikt|c}\left(\theta^0_{ct}, \theta^1_{ct}\right) = \theta^k_{ct} + \left(1 - \theta^0_{ct} - \theta^1_{ct}\right) s_{ikt|c,B}$$

that is, the probability of drug k being the only one available plus the probability that both are available times the probability that drug k is chosen when both are available.

The shock ϵ_{ikct} is observed after choosing a pharmacy chain and the consumer chooses a chain using the expected utility of choosing a pharmacy by taking expectations with respect to the possible choice sets and with respect to the shock ϵ_{ikct}. The consumer utility of visiting pharmacy c is then $I_{ict} + \varepsilon_{ict}$, where

$$I_{ict}\left(\theta^0_{ct}, \theta^1_{ct}\right) \equiv \sum_{k \in \{0,1\}} \theta^k_{ct} V_{ikct} + \left(1 - \theta^0_{ct} - \theta^1_{ct}\right) E_{\epsilon_{ikct}}\left[\max_{k \in \{0,1\}} (V_{ikct} + \lambda_c \epsilon_{ikct})\right]$$

with the log-sum formula for the inclusive value in case the choice set contains both products:

$$E_{\epsilon_{ikct}}\left[\max_{k \in \{0,1\}} (V_{ikct} + \lambda_c \epsilon_{ikct})\right] = \lambda_c \ln\left(\sum_{k \in \{0,1\}} e^{V_{ikct}/\lambda_c}\right)$$

which is always greater than $\max(V_{i0ct}, V_{i1ct})$.

Then, as ε_{ict} is extreme value distributed independently across chains, patient i chooses chain c with probability

$$s_{ict}\left(\boldsymbol{\theta}^0_t, \boldsymbol{\theta}^1_t\right) = \frac{e^{I_{ict}(\theta^0_{ct}, \theta^1_{ct})}}{\sum_{\tilde{c}} e^{I_{i\tilde{c}t}(\theta^0_{\tilde{c}t}, \theta^1_{\tilde{c}t})}}.$$

Denoting by $F(.|\boldsymbol{\beta})$ the cumulative distribution function of consumer preferences $V_{it} \equiv (V_{i01t}, \dots, V_{i0Ct}, V_{i11t}, \dots, V_{i1Ct})$ conditional on the parameter vector β, we can write the aggregate choice probability or market share of drug k sold by c in period t as

$$s_{kct}\left(\boldsymbol{\theta}^0_t, \boldsymbol{\theta}^1_t\right) = \int s_{ikct}\left(\boldsymbol{\theta}^0_t, \boldsymbol{\theta}^1_t\right) dF(\mathbf{V}_{it}|\boldsymbol{\beta}) = \int s_{ict}\left(\boldsymbol{\theta}^0_t, \boldsymbol{\theta}^1_t\right) s_{ikt|c}\left(\theta^0_{ct}, \theta^1_{ct}\right) dF(\mathbf{V}_{it}|\boldsymbol{\beta}), \qquad \text{(B.2)}$$

and the aggregate market share of drug k within the pharmacy chain c as

$$s_{kt|c}\left(\theta^0_{ct}, \theta^1_{ct}\right) = \int s_{ikt|c}\, dF(\mathbf{V}_{it}|\boldsymbol{\beta}) = \theta^k_{ct} + \left(1 - \theta^0_{ct} - \theta^1_{ct}\right) \int s_{ikt|c,B}\, dF(\mathbf{V}_{it}|\boldsymbol{\beta}).$$

Let us now turn to the behavior of the pharmacy chains. The profits of chain c normalized by market size in time t are

$$\pi_{ct} = \sum_{k\in\{0,1\}} (p_{kct} - w_{kct})s_{kct},$$

where p_{kct} is the retail price and w_{kct} the wholesale price of drug k in pharmacy c at t. For almost all drugs under patent (including the one used in the structural model estimation), the retail prices happen to always be equal to the price ceiling ($p_{kct} = \bar{p}_t$). We denote by $m_{kct} \equiv \bar{p}_t - w_{kct}$ the product price-cost margin.

Appendix A.1 of the paper shows that if $m_{1ct} < m_{0ct}$ then $\theta^1_{ct} = 0$ and if $m_{1ct} > m_{0ct}$ then $\theta^0_{ct} = 0$. We reproduce this proof here before continuing the exposition of the general bargaining model.

When the pharmacy chain procures the drug from both direct and parallel imports, both margins m_{0ct} and m_{1ct} must be positive and necessary first-order conditions for an interior solution of θ's are

$$0 = \frac{\partial \pi_{ct}}{\partial \theta^0_{ct}} = \frac{\partial \pi_{ct}}{\partial \theta^1_{ct}}.$$

For θ^0_{ct}, the first-order condition is (the equivalent condition for θ^1_{ct} is not shown):

$$\begin{aligned} 0 &= \sum_k m_{kct} \frac{\partial s_{kct}}{\partial \theta^0_{ct}} \\ &= \int \sum_k m_{kct} \Bigg[\underbrace{\frac{\partial s_{ikt|c}}{\partial \theta^0_{ct}}}_{\substack{\text{change in probability}\\ \text{to choose } k \text{ in } c}} \underbrace{s_{ict}}_{\substack{\text{probability to}\\ \text{choose chain } c}} + \underbrace{s_{ikt|c}}_{\substack{\text{probability to choose}\\ k \text{ in chain } c}} \underbrace{\frac{\partial s_{ict}}{\partial \theta^0_{ct}}}_{\substack{\text{change in probability}\\ \text{to choose chain } c}} \Bigg] dF(\mathbf{V}_{it}|\boldsymbol{\beta}), \end{aligned}$$

which shows that θ^0_{ct} has substitution effects within and across chains for both versions of the drug.

Developing the first-order conditions using the effects of θ's on the demand, we show below that it must be that $\theta^k_{ct} = 0$ if m_{kct} is the lowest of the two margins. As

$$\frac{\partial s_{ikt|c}}{\partial \theta^{k'}_{ct}} = 1_{\{k=k'\}} - s_{ikt|c,B} \quad \text{and} \quad \frac{\partial s_{ict}}{\partial \theta^{k'}_{ct}} = \left[V_{ik'ct} - \lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c} \right) \right] s_{ict}(1 - s_{ict}) \le 0,$$

using the fact that

$$\begin{aligned} &\frac{\partial s_{ik't|c}}{\partial \theta^0_{ct}} s_{ict} + s_{ik't|c} \frac{\partial s_{ict}}{\partial \theta^0_{ct}} \\ &\quad = (1_{\{k'=0\}} - s_{ik't|c,B}) s_{ict} + s_{ik't|c} \left[V_{i0ct} - \lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c} \right) \right] (1 - s_{ict}) s_{ict}, \end{aligned}$$

we obtain that the first-order condition for optimal θ_{ct}^0 implies

$$\frac{m_{0ct}}{m_{1ct}} = \frac{\displaystyle\int s_{i1t|c,B}s_{ict} + s_{i1t|c}\left[\lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c}\right) - V_{i0ct}\right](1 - s_{ict})s_{ict}\, dF(\mathbf{V}_{it})}{\displaystyle\int s_{i1t|c,B}s_{ict} - s_{i0t|c}\left[\lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c}\right) - V_{i0ct}\right](1 - s_{ict})s_{ict}\, dF(\mathbf{V}_{it})} \tag{B.3}$$

because $1 - s_{i0t|c,B} = s_{i1t|c,B}$ and $1 - s_{i0t|c} = s_{i1t|c}$.

Similarly, the first-order condition with respect to θ_{ct}^1 (for an interior solution) can be written

$$\frac{m_{1ct}}{m_{0ct}} = \frac{\displaystyle\int s_{i0t|c,B}s_{ict} + s_{i0t|c}\left[\lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c}\right) - V_{i1ct}\right](1 - s_{ict})s_{ict}\, dF(\mathbf{V}_{it}|\boldsymbol{\beta})}{\displaystyle\int s_{i0t|c,B}s_{ict} - s_{i1t|c}\left[\lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c}\right) - V_{i1ct}\right](1 - s_{ict})s_{ict}\, dF(\mathbf{V}_{it}|\boldsymbol{\beta})}. \tag{B.4}$$

We can see that only one of the first-order conditions will be satisfied. Indeed, as $1 - s_{i0t|c} = s_{i1t|c}$,

$$\begin{aligned}
&s_{i1t|c,B}s_{ict} + s_{i1t|c}\left[\lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c}\right) - V_{i0ct}\right](1 - s_{ict})s_{ict} \\
&\quad = s_{i1t|c,B}s_{ict} - s_{i0t|c}\left[\lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c}\right) - V_{i0ct}\right](1 - s_{ict})s_{ict} \\
&\qquad + \left[\lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c}\right) - V_{i0ct}\right](1 - s_{ict})s_{ict} \\
&\quad > s_{i1t|c,B}s_{ict} - s_{i0t|c}\left[\lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c}\right) - V_{i0ct}\right](1 - s_{ict})s_{ict},
\end{aligned}$$

and similarly,

$$\begin{aligned}
&s_{i0t|c,B}s_{ict} + s_{i0t|c}\left[\lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c}\right) - V_{i1ct}\right](1 - s_{ict})s_{ict} \\
&\quad > s_{i0t|c,B}s_{ict} - s_{i1t|c}\left[\lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c}\right) - V_{i1ct}\right](1 - s_{ict})s_{ict}.
\end{aligned}$$

Thus, equation (B.3) cannot be true if $m_{1ct} > m_{0ct}$, and equation (B.4) cannot be true if $m_{1ct} < m_{0ct}$.

In the case in which $m_{1ct} < m_{0ct}$, there is no interior solution for θ_{ct}^1, and thus we will have $\theta_{ct}^1 = 0$, meaning that the pharmacy chain never proposes the drug with the lowest margin alone. Then θ_{ct}^0 is a solution of equation (B.3). Thus if $m_{1ct} < m_{0ct}$ then $\theta_{ct}^1 = 0$ and if $m_{1ct} > m_{0ct}$ then $\theta_{ct}^0 = 0$.

These conditions lead to demands as $s_{ikct}(\theta_t^0, \theta_t^1)$.

Thus, there are two cases depending on which good is the high margin one. If parallel imports (good 0) is the high-margin product for chain c, then $\theta_{ct}^1 = 0$, in the other case,

$\theta^0_{ct} = 0$. We can now simplify the previous expression of the expected inclusive value as

$$I_{ict}\big(\theta^0_{ct}, \theta^1_{ct}\big) = 1_{\{w_{0ct} \le w_{1ct}\}}\left[\theta^0_{ct} V_{i0ct} + \big(1 - \theta^0_{ct}\big)\lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c}\right)\right] + 1_{\{w_{0ct} > w_{1ct}\}}\left[\theta^1_{ct} V_{i1ct} + \big(1 - \theta^1_{ct}\big)\lambda_c \ln\left(\sum_k e^{V_{ikct}/\lambda_c}\right)\right]$$

and the individual choice probabilities are

$$s_{i1ct}\big(\boldsymbol{\theta}^0_t, \boldsymbol{\theta}^1_t\big) = \frac{e^{I_{ict}(\theta^0_{ct}, \theta^1_{ct})}}{\sum_{\tilde{c}} e^{I_{i\tilde{c}t}(\theta^0_{\tilde{c}t}, \theta^1_{\tilde{c}t})}}\left[\theta^1_{ct} + \big(1 - \theta^0_{ct} - \theta^1_{ct}\big)\frac{e^{V_{i1ct}/\lambda_c}}{e^{V_{i0ct}/\lambda_c} + e^{V_{i1ct}/\lambda_c}}\right]$$

and

$$s_{i0ct}\big(\boldsymbol{\theta}^0_t, \boldsymbol{\theta}^1_t\big) = \frac{e^{I_{ict}(\theta^0_{ct}, \theta^1_{ct})}}{\sum_{\tilde{c}} e^{I_{i\tilde{c}t}(\theta^0_{\tilde{c}t}, \theta^1_{\tilde{c}t})}}\left[\theta^0_{ct} + \big(1 - \theta^0_{ct} - \theta^1_{ct}\big)\frac{e^{V_{i0ct}/\lambda_c}}{e^{V_{i0ct}/\lambda_c} + e^{V_{i1ct}/\lambda_c}}\right],$$

where $\theta^0_t \equiv (\theta^0_{0t}, \dots, \theta^0_{Ct})'$ and $\theta^1_t \equiv (\theta^1_{0t}, \dots, \theta^1_{Ct})'$.

Integrating over the distribution of preferences, we obtain the market share of each product as

$$s_{kct}\big(\boldsymbol{\theta}^0_t, \boldsymbol{\theta}^1_t\big) = \int s_{ikct}\big(\boldsymbol{\theta}^0_t, \boldsymbol{\theta}^1_t\big)\, dF(\mathbf{V}_{it}|\boldsymbol{\beta}).$$

The profit maximization problem for each chain c at t now implies the following optimality condition:

$$\frac{\partial \pi_{ct}}{\partial \theta^k_{ct}}\big(\boldsymbol{\theta}^0_t, \boldsymbol{\theta}^1_t\big) \begin{cases} \le 0 & \text{if } \theta^k_{ct} = 0, \\ = 0 & \text{if } 0 < \theta^k_{ct} < 1, \\ \ge 0 & \text{if } \theta^k_{ct} = 1. \end{cases} \tag{B.5}$$

The derivative of profits with respect to the θ's are

$$\frac{\partial \pi_{ct}}{\partial \theta^k_{ct}} = m_{0ct}\frac{\partial s_{0ct}\big(\boldsymbol{\theta}^0_t, \boldsymbol{\theta}^1_t\big)}{\partial \theta^k_{ct}} + m_{1ct}\frac{\partial s_{1ct}\big(\boldsymbol{\theta}^0_t, \boldsymbol{\theta}^1_t\big)}{\partial \theta^k_{ct}}. \tag{B.6}$$

This implies that

$$\frac{\partial \theta^{0*}_{ct}(\mathbf{w}_{0t}, \mathbf{w}_{1t})}{\partial w_{kct}} \begin{cases} = 0 & \text{if } \theta^{0*}_{ct}(\mathbf{w}_{0t}, \mathbf{w}_{1t}) = 0 \Leftrightarrow \boldsymbol{w}_{0ct} > \boldsymbol{w}_{1ct}, \\ = \dfrac{\partial \theta^{0*}_{ct}(\mathbf{w}_{0t}, \mathbf{w}_{1t})}{\partial w_{kct}} & \text{if } \theta^{0*}_{ct}(\mathbf{w}_{0t}, \mathbf{w}_{1t}) > 0 \Leftrightarrow \boldsymbol{w}_{0ct} \le \boldsymbol{w}_{1ct}, \end{cases}$$

$$\frac{\partial \theta^{1*}_{ct}(\mathbf{w}_{0t}, \mathbf{w}_{1t})}{\partial w_{kct}} \begin{cases} = 0 & \text{if } \theta^{1*}_{ct}(\mathbf{w}_{0t}, \mathbf{w}_{1t}) = 0 \Leftrightarrow \boldsymbol{w}_{1ct} > \boldsymbol{w}_{0ct}, \\ = \dfrac{\partial \theta^{1*}_{ct}(\mathbf{w}_{0t}, \mathbf{w}_{1t})}{\partial w_{kct}} & \text{if } \theta^{1*}_{ct}(\mathbf{w}_{0t}, \mathbf{w}_{1t}) > 0 \Leftrightarrow \boldsymbol{w}_{1ct} \le \boldsymbol{w}_{0ct}. \end{cases}$$

We now model the behavior of the manufacturer and parallel importer that bargain with pharmacy chains over wholesale prices. Using the simpler notation θ^{0*}_t, θ^{1*}_t for

$\theta_t^{0*}(w_{0t}, w_{1t})$, $\theta_t^{1*}(w_{0t}, w_{1t})$, the profits of the manufacturer are given by

$$\Pi_t\big(\mathbf{w}_{1t}, \boldsymbol{\theta}_t^{0*}, \boldsymbol{\theta}_t^{1*}\big) = \sum_c \underbrace{(w_{1ct} - c_t)s_{1ct}\big(\boldsymbol{\theta}_t^{0*}, \boldsymbol{\theta}_t^{1*}\big)}_{\substack{\text{Profit to Manufacturer of Direct Imports}\\ \text{profit in chain } c}} + \underbrace{\big(p_{1ct}^{I(c)} - c_t\big)s_{0ct}\big(\boldsymbol{\theta}_t^{0*}, \boldsymbol{\theta}_t^{1*}\big)}_{\substack{\text{Profit to Manufacturer of Parallel Import}\\ \text{in chain c at wholesale source price } p_{1ct}^I}},$$

where c_t is the marginal cost of production, $p_{1ct}^{I(c)}$ is the manufacturer price in the source country of the parallel importer supplying chain c, and, as before, w_{1ct} is the wholesale prices charged for direct imported drugs to chain c at time t.

Each pairwise negotiation with the pharmacy chains, the manufacturer and pharmacy chain c set wholesale prices to maximize the Nash-product

$$\max_{w_{1ct}}(\Pi_t - \Pi_{-c,t})^{b_{1c}}(\pi_{ct} - \pi_{-1,ct})^{1-b_{1c}}, \tag{B.7}$$

where b_{1c} is the bargaining weight of the manufacturer when negotiating with chain c, $\Pi_{-c,t}$ is the manufacturer's profit in absence of an agreement with pharmacy chain c, and $\pi_{-1,ct}$ is likewise pharmacy chain c's profit in absence of an agreement with the manufacturer. The first-order condition for a solution to equation (B.7) is

$$b_{1c}\frac{\partial \Pi_t/\partial w_{1ct}}{\Pi_t - \Pi_{-c,t}} + (1 - b_{1c})\frac{\partial \pi_{ct}/\partial w_{1ct}}{\pi_{ct} - \pi_{-1,ct}} = 0. \tag{B.8}$$

In maximizing the Nash product, there will be an effect on the manufacturer's profit from how changes in wholesale prices affect the equilibrium $\theta_t^*(w_{0t}, w_{1t})$ in the next stage of the game.

Denote the net value of agreement for the manufacturer and chain c as $\Delta_c\Pi_t \equiv \Pi_t - \Pi_{-c,t}$ and $\Delta_1\pi_{ct} \equiv \pi_{ct} - \pi_{-1,ct}$, respectively. The derivative of the manufacturer's profit with respect to the wholesale price is

$$\begin{aligned}
&\frac{\partial \Pi_t\big(\mathbf{w}_{1t}, \boldsymbol{\theta}_t^{0*}(\mathbf{w}_{0t}, \mathbf{w}_{1t}), \boldsymbol{\theta}_t^{1*}(\mathbf{w}_{0t}, \mathbf{w}_{1t})\big)}{\partial w_{1ct}} \\
&= s_{1ct}\big(\boldsymbol{\theta}_t^{0*}, \boldsymbol{\theta}_t^{1*}\big) + \sum_{\tilde{c}}\Bigg[(w_{1\tilde{c}t} - c_t)\frac{\partial s_{1\tilde{c}t}\big(\boldsymbol{\theta}_t^{0*}(\mathbf{w}_{0t}, \mathbf{w}_{1t}), \boldsymbol{\theta}_t^{1*}(\mathbf{w}_{0t}, \mathbf{w}_{1t})\big)}{\partial w_{1ct}} \\
&\quad + \big(p_{1\tilde{c}t}^{I(\tilde{c})} - c_t\big)\frac{\partial s_{0\tilde{c}t}\big(\boldsymbol{\theta}_t^{0*}(\mathbf{w}_{0t}, \mathbf{w}_{1t}), \boldsymbol{\theta}_t^{1*}(\mathbf{w}_{0t}, \mathbf{w}_{1t})\big)}{\partial w_{1ct}}\Bigg] \\
&= s_{1ct}\big(\boldsymbol{\theta}_t^{0*}, \boldsymbol{\theta}_t^{1*}\big) + \sum_{\tilde{c}}\Bigg[w_{1\tilde{c}t}\frac{\partial s_{1\tilde{c}t}\big(\boldsymbol{\theta}_t^{0*}(\mathbf{w}_{0t}, \mathbf{w}_{1t}), \boldsymbol{\theta}_t^{1*}(\mathbf{w}_{0t}, \mathbf{w}_{1t})\big)}{\partial w_{1ct}} \\
&\quad + p_{1\tilde{c}t}^{I(\tilde{c})}\frac{\partial s_{0\tilde{c}t}\big(\boldsymbol{\theta}_t^{0*}(\mathbf{w}_{0t}, \mathbf{w}_{1t}), \boldsymbol{\theta}_t^{1*}(\mathbf{w}_{0t}, \mathbf{w}_{1t})\big)}{\partial w_{1ct}}\Bigg],
\end{aligned}$$

where we use the fact that aggregate demand is fixed, and thus $\sum_{\tilde{c}}(\frac{\partial s_{1\tilde{c}t}}{\partial w_{1ct}} + \frac{\partial s_{0\tilde{c}t}}{\partial w_{1ct}}) = 0$.

Similarly, the derivative of chain c's profits with respect to the wholesale price w_{1ct} is

$$\frac{\partial \pi_{ct}\big(\mathbf{w}_{0ct}, \mathbf{w}_{1ct}, \boldsymbol{\theta}_t^{0*}(\mathbf{w}_{0t}, \mathbf{w}_{1t}), \boldsymbol{\theta}_t^{1*}(\mathbf{w}_{0t}, \mathbf{w}_{1t})\big)}{\partial w_{1ct}}$$
$$= -s_{1ct}\big(\boldsymbol{\theta}_t^{0*}, \boldsymbol{\theta}_t^{1*}\big) + (\bar{p}_t - w_{1ct})\frac{\partial s_{1ct}\big(\boldsymbol{\theta}_t^{0*}(\mathbf{w}_{0t}, \mathbf{w}_{1t}), \boldsymbol{\theta}_t^{1*}(\mathbf{w}_{0t}, \mathbf{w}_{1t})\big)}{\partial w_{1ct}}$$
$$+ (\bar{p}_t - w_{0ct})\frac{\partial s_{0ct}\big(\boldsymbol{\theta}_t^{0*}(\mathbf{w}_{0t}, \mathbf{w}_{1t}), \boldsymbol{\theta}_t^{1*}(\mathbf{w}_{0t}, \mathbf{w}_{1t})\big)}{\partial w_{1ct}}.$$

In the two expressions above, the derivatives of market shares with respect to wholesale prices depend on the derivatives of market shares with respect to equilibrium $\theta' s$ and the derivatives of equilibrium $\theta' s$ with respect to wholesale prices, which can be obtained using the optimal behavior of pharmacies, as detailed below.

Using vector notation for market shares $s_{0t} = (s_{01t}, \dots, s_{0Ct})$ and $s_{1t} = (s_{11t}, \dots, s_{1Ct})$, we can then rewrite equation (B.8) governing the solution to the bargaining between the manufacturer and chain c as

$$s_{1ct} + \mathbf{w}_{1t}' \frac{\partial \boldsymbol{s}_{1t}}{\partial w_{1ct}} + \boldsymbol{p}_{1t}^{I} \frac{\partial \boldsymbol{s}_{0t}}{\partial w_{1ct}} = \frac{1 - b_{1c}}{b_{1c}} \frac{\Delta_c \Pi_t}{\Delta_1 \pi_{ct}} \left(s_{1ct} - m_{1ct} \frac{\partial s_{1ct}}{\partial w_{1ct}} - m_{0ct} \frac{\partial s_{0ct}}{\partial w_{1ct}} \right). \tag{B.9}$$

Letting $s_{j\tilde{c}t\backslash 1c}$ denote the share of chain $\tilde{c}$'s product j in t when direct imports are not available at chain c, we can express the net value for the manufacturer, suppressing arguments θ_t^*, as

$$\Delta_c \Pi_t = \sum_{\tilde{c}} \big[(w_{1\tilde{c}t} - c_t)s_{1\tilde{c}t} + \big(p_{1\tilde{c}t}^{I(\tilde{c})} - c_t\big)s_{0\tilde{c}t}\big] - \sum_{\tilde{c}} \big[(w_{1\tilde{c}t} - c_t)s_{1\tilde{c}t\backslash 1c} + \big(p_{1\tilde{c}t}^{I(\tilde{c})} - c_t\big)s_{0\tilde{c}t\backslash 1c}\big]$$
$$= \sum_{\tilde{c}} \big(w_{1\tilde{c}t}\Delta_{1c}s_{1\tilde{c}t} + p_{1\tilde{c}t}^{I(\tilde{c})}\Delta_{1c}s_{0\tilde{c}t}\big),$$

because $s_{jct\backslash 1c} = 0$, and defining $\Delta_{1c}s_{j\tilde{c}t} \equiv s_{j\tilde{c}t} - s_{j\tilde{c}t\backslash 1c}$, that is, the difference in share of product j in chain $\tilde{c}$ between the case of agreement and disagreement in the negotiations between the manufacturer and chain c.

Similarly, the net value for the chain is

$$\Delta_1 \pi_{ct} = (\bar{p}_t - w_{1ct})s_{1ct} + (\bar{p}_t - w_{0ct})\Delta_{1c}s_{0ct}.$$

Note that the derivatives of market shares with respect to wholesale price follow from the chain rule and the implicit function theorem governing the change in equilibrium θ_t^* when wholesale prices change due to pharmacy chains' optimal behavior.

Parallel Importers Behavior:

The parallel importer profits is given by

$$\Pi_t^{\mathrm{PI}} = \sum_c \big(w_{0ct} - p_{0ct}^{I(c)}\big)s_{0ct}\big(\boldsymbol{\theta}_t^{0*}, \boldsymbol{\theta}_t^{1*}\big),$$

where w_{0ct} is the wholesale price paid for parallel imported drugs by chain c at time t and $p_{0ct}^{I(c)}$ is the price that the importer has to pay for the drug in the source country, which we allow to vary across chains c for full generality because each chain may require different source countries.

We assume that the parallel importer bargains over the wholesale price with each pharmacy chain c, where they take as given the negotiated wholesale prices of originator products to each pharmacy chain $w_{1t} = (w_{11t}, w_{12t}, \dots, w_{1Ct})$. When bargaining over the wholesale prices charged to the chains, w_{0t}, the parallel importer will also take into account how changes in these prices will affect the equilibrium $\theta_t^{0*}(w_{0t}, w_{1t})$, $\theta_t^{1*}(w_{0t}, w_{1t})$. Similar to equation (B.8), the first-order conditions for the solution to the Nash bargaining between each pharmacy chain c and the parallel importer is

$$b_{0c}\frac{\partial \Pi_t^{\mathrm{PI}}/\partial w_{0ct}}{\Pi_t^{\mathrm{PI}} - \Pi_{-c,t}^{\mathrm{PI}}} + (1-b_{0c})\frac{\partial \pi_{ct}/\partial w_{0ct}}{\pi_{ct} - \pi_{-0,ct}} = 0, \tag{B.10}$$

which can be rewritten, following the previous approach using vector notations for prices and market shares stacked over the chains c, as

$$s_{0ct} + \left(\mathbf{w}_{0t} - \boldsymbol{p}_{0t}^I\right)' \frac{\partial \boldsymbol{s}_{0t}}{\partial w_{0ct}} = \frac{1-b_{0c}}{b_{0c}} \frac{\Delta_c \Pi_t^{\mathrm{PI}}}{\Delta_0 \pi_{ct}} \left(s_{0ct} - m_{1ct}\frac{\partial s_{1ct}}{\partial w_{0ct}} - m_{0ct}\frac{\partial s_{0ct}}{\partial w_{0ct}} \right), \tag{B.11}$$

where the left-hand side is the derivative of parallel importer profits with respect to the wholesale price w_{0ct} and we denoted $\Delta_c \Pi_t^{\mathrm{PI}} = \Pi_t^{\mathrm{PI}} - \Pi_{-c,t}^{\mathrm{PI}}$ with $\Pi_{-c,t}^{\mathrm{PI}} = \sum_{\tilde{c}\neq c}(w_{0\tilde{c}t} - p_{0\tilde{c}t}^I) s_{0\tilde{c}t\backslash 0c}$ and

$$\Delta_0 \pi_{ct} = \pi_{ct} - \pi_{-0,ct} = (\bar{p}_t - w_{1ct})\Delta_{0c} s_{1ct} + (\bar{p}_t - w_{0ct}) s_{0ct},$$

where as defined previously $\Delta_{0c} s_{1ct}$ corresponds to the market share of the direct imports at chain c when there are no parallel trade version at chain c.

REFERENCES

DUBOIS, P., AND L. LASIO (2018): "Identifying Industry Margins With Price Constraints: Structural Estimation on Pharmaceuticals," *American Economic Review*, 12 (108), 3685–3724. [11]

JUDD, K. L. (1998): *Numerical Methods in Economics*. Cambridge, MA: MIT Press. [7]

Co-editor Fabrizio Zilibotti handled this manuscript.

Manuscript received 23 January, 2018; final version accepted 14 June, 2020; available online 14 July, 2020.