Online Appendix

Quality Regulation and Competition: Evidence from Pharmaceutical Markets

Juan Pablo Atal, José Ignacio Cuesta and Morten Sæthre

A Effects on Drug Quality

We study whether the bioequivalence regulation affected observable quality of drugs in the market. The effects of the reform on quality are key to understand its welfare consequences. Unfortunately, direct measures of quality (e.g., results from drug testing) are not available in our setting, so in this section we use adverse health events associated with specific drugs and drug recalls as indirect measures. Note that both outcome variables are mostly related to drug safety—which was already covered by the quality standards in place before the reform we study—whereas bioequivalence certification mostly targets drug efficacy. We see this analysis as a complement to our main empirical strategy in Section 6, where we estimate changes in quality due to the regulation using a structural model.

Let the quality outcome for market m at time t be $y_{mt} = \mu_{mt} sales_{mt}$, where μ_{mt} is the probability of an adverse effect associated with drugs in market m, and $sales_{mt}$ are sales of drugs in m. Similar to Jin and Leslie (2003), we model the probability of an adverse outcome—either an adverse health event or a drug recall—as $\mu_{mt} = \mu_{m0} + \gamma_t + \theta T_{mt} + \varepsilon_{mt}$, which combines a baseline probability μ_{m0}, with time shocks common to all markets γ_t, a shifter related to quality regulation θT_{mt}, and a random shock ε_{mt}. This simple framework motivates the estimating equation:

$$\frac{y_{mt}}{sales_{mt}} = \mu_{m0} + \gamma_t + \theta T_{mt} + \varepsilon_{mt} \tag{11}$$

where θ measures the effect of stronger quality regulation on the number of adverse outcomes per unit of sales, whereas μ_{m0} and γ_t are captured by market and time fixed effects.

A.1 Evidence from Adverse Health Events

A first set of outcomes related to quality are the adverse health events associated with drug consumption. We collect data on yearly clinical outcomes between 2010 and 2017 for ICD-10 diagnosis codes associated with molecules in our sample. We exploit public records collected by DEIS (2019), which cover admissions, days of hospitalization, and the number of surgeries across all hospitals in Chile. We link diagnoses to molecules using a crosswalk between American Hospital Formulary Service (AHFS) and ICD codes that tracks adverse health events associated

with the consumption of drugs (WHO, 2007).[32,33] We focus on the 71 molecules with at least one listed adverse effect.[34] In our setting, these events are rare. In 2010, there were on average 7.3 admissions, 13.2 hospital days, and 0.002 surgeries per 100,000 daily doses sold across all markets.

We estimate equation (11) for these outcomes. Columns 1–3 in Table A.2-A display the results across all markets. We find no evidence suggesting that stronger quality regulation decreased the number of discharges and the number of days associated with them. Moreover, we find no evidence of heterogeneous effects on these outcomes across small and large markets in Table A.2-B. These results suggest that stronger quality regulation was not able to reduce adverse health effects of drugs.

A.2 Evidence from Drug Recalls

To study effects on drug recalls, we collect data on the 209 recalls for prescription drugs that occurred during our period of study. Recalls are implemented by ISP as preventative sanitary measures upon notice of adverse events linked to licensed drugs.[35] In the period we cover, there is an average of 1.9 recalls per month, of which 1.4 relate to molecules without bioequivalence requirements and 0.5 relate to molecules with bioequivalence requirements.[36]

Our estimates of equation (11) in the sample of molecules with bioequivalence requirements provide no evidence suggesting that stronger quality regulation improved drug quality as measured by recalls. Columns 4 and 5 in Table A.2-A display the results across all markets, while Table A.2-B does so by market size. Our point estimates are close to zero across both specifications. These results suggest that imposing bioequivalence requirements was unable to decrease drug recalls relative to the quality standards previously in place in our setting.

[32]When several ICD codes capture adverse effects associated with the same molecule, we assign outcomes to molecules using weights for sales volume across molecules within each ICD code.

[33]As an example, admissions coded under "T455 - Poisoning by, adverse effect of and underdosing of anticoagulants and antithrombotic drugs" are attributed to the consumption of Acenocoumarol, an anticoagulant.

[34]The results from the regressions on market outcomes are very similar when restricted to this sample. The results are available from authors upon request.

[35]The reasons for these recalls can be categorized broadly into manufacturing defects including chemical defects and contamination (71%); efficacy concerns or side effects (19 %); or others, which mostly correspond to counterfeit drugs or mislabeling (20 %). Due to the small number of recall events, we use all data irrespective of the specific reason.

[36]Figure A.4 shows the monthly recall frequency, split into drugs with bioequivalence requirements in our sample, and drugs without bioequivalence requirement. We cannot reject the hypothesis of a same trend in recalls over time across these two groups. We estimate an OLS regression for recall rates on an indicator for requirements and its interaction with a time trend, and find that the coefficient on the latter is not statistically different to zero.

B Description of Consumer Survey

B.1 Methodology and Results

To aid the interpretation of our descriptive evidence and guide our model assumptions, we collect additional survey data in which we interview consumers and gather information on perceived quality, perceived price differences, relationship between physician prescription behavior and consumer choices and some additional characterization variables. The questionnaire is displayed in Section B.2 below.

A surveying team composed by 6 members conducted surveys in 4 counties in the city of Santiago, namely Ñuñoa, Providencia, Puente Alto and Santiago. Within such counties, surveyors recruited consumers outside pharmacies, where consumers were purchasing drugs. This recruiting strategy aimed at constructing a sample of consumers familiar with the pharmaceutical market. Recruited participants were asked to participate in a survey with a duration of between 5 and 10 minutes, and were offered no compensation for it.

To collect data on perceived quality and price differences, we focus on a particular market, Atorvastatin, a molecule commonly prescribed as a treatment to cholesterol. Within that market, we focus on 4 drugs that are relevant products in this market. In particular, we work with (i) a popular innovator drug called Lipitor, which is produced by Pfizer, (ii) a bioequivalent branded generic called Lipoten, produced by Pharmavita, (iii) a bioequivalent unbranded generic called simply Atorvastatina, produced by Mintlab, and (iv) and a non-bioequivalent unbranded generic also called Atorvastatina and produced by Mintlab. For reference, the prices of these drugs in the market are around \$50,000 CLP, \$10,000 CLP, \$2,500 CLP and \$2,500 CLP respectively (\$77.5, \$15.5 and \$7.8 USD respectively). Perceived quality and price differences are elicited using a paper sheet that showed the 4 drugs, which is displayed in Figure A.5.

The final sample includes $N = 401$ consumers. Table A.3 provides summary statistics for the main variables in the survey. Among consumers in the sample, 62% report having a household member with a chronic disease, and 36% report purchasing Atorvastatin for a household member. In terms of purchase behavior, 41% often purchase innovator drugs, 21% often purchase branded generics, and the remainder 38% often purchase unbranded generics. The main results of the survey and their relationship to the results in our main analysis are discussed in Section 5. We code observations in which a consumer answered "I don't know" or "I don't recall" as missing. Finally, the questions regarding physicians' prescription behavior have less observations because they were added to the survey with a lag and are therefore not available for around a fourth of the sample.

B.2 Questionnaire

We are conducting a survey about the quality perception of drugs sold in pharmacies. We will ask you a few questions regarding the quality and prices of drugs. In all examples, we will focus in a drug called Atorvastatin, which is commonly used to control cholesterol levels. While we understand that it may be that no one in your household takes Atorvastatin, we ask that you consider it as an example and think as if you had to acquire it for a family member.

1. [Show pictures of four drugs] Consider a scale of 1 to 7, where 1 is a drug of the minimum quality and that does not have the desired therapeutic effects and 7 is a drug of the highest quality that has exactly the expected therapeutic effects. What level of quality do you think the following drug has?

    - Innovator
    - Bioequivalent unbranded generic
    - Unbranded generic
    - Bioequivalent branded generic

2. [Show pictures of 4 drugs] If the price of the innovator drug is $50,000. What price do you think each of these drugs has?

    - Bioequivalent unbranded generic
    - Unbranded generic
    - Bioequivalent branded generic

3. [Show pictures of innovator and bioequivalent unbranded generic] If you were buying a box of Atorvastatin and were offered these two drugs. The innovator is priced at $50,000 in pharmacies. What do you think is the price of this generic?

4. [Show pictures of innovator and unbranded generic] If you were buying a box of Atorvastatin and were offered these two drugs. The innovator is priced at $50,000 in pharmacies. What do you think is the price of this generic?

5. [Show pictures of innovator and bioequivalent branded generic] If you were buying a box of Atorvastatin and were offered these two drugs. The innovator is priced at $50,000 in pharmacies. What do you think is the price of this generic?

6. [Show bioequivalence label] Have you ever seen this label on a drug before this survey?

    - Yes

- No

7. [Do not read, use the following scale] Do you know what it means for a generic drug to be bioequivalent?
   - Very good response: Bioequivalence implies that two drugs have exactly the same therapeutic effects as the original
   - Good response: The generic is the same as the innovator
   - Regular response: A vague answer in terms of the quality of both drugs
   - Bad response: They are part of the same group of medications (e.g. both are Atorvastatin)
   - He has no idea: He does not know, he has no idea, he has not heard
8. When doctors deliver prescriptions, do they generally prescribe drugs by specifying a particular brand or without specifying a brand?
   - Prescribe drug without a specific brand
   - Prescribe drug with a specific brand
   - Does not know
9. When buying a prescription drug at a pharmacy, how much does your doctor, the pharmacist who serves you, and yourself weigh in deciding which version of the drug to buy? In particular, on a scale of 1 to 5, where 1 is no power and 5 is a lot of power, how much power they have:
   - Doctor
   - Pharmacist
   - Customer
10. What type of drug did you buy the last time you needed one?
    - Innovator
    - Bioequivalent unbranded generic
    - Unbranded generic
    - Bioequivalent branded generic
    - Do not remember
    - Never purchased
11. Do you or anyone in your home take any drug for a chronic illness?

- Yes
- No

12. Do you or anyone in your household take any drug to control cholesterol?
    - Yes
    - No
13. What type of drug do you choose when you buy this medication for cholesterol control?
    - Innovator
    - Bioequivalent unbranded generic
    - Unbranded generic
    - Bioequivalent branded generic
    - Do not remember
    - Never purchased

C Empirical Model Details

C.1 The Constant Expenditure Model

Consider the direct utility function for consumer i over quantities q_{ij} of products indexed by j with quality index ψ_{ij} and the amount z_i of the numeraire:

$$u(\sum_j \psi_{ij} q_{ij}, z_i) = \varphi \ln(\sum_j \psi_{ij} q_{ij}) + (1 - \varphi) \ln z_i$$

where the quality index is specified as:

$$\psi_{ij} = e^{\frac{\delta_j + \epsilon_{ij}}{\theta}}$$

with δ_j being a utility term common across consumers for product j, while ϵ_{ij} is an individual-specific random utility term. This utility function implies that the goods indexed by j are perfect substitutes, where the one with the highest ratio of quality over price ψ_{ij}/p_j will be chosen by the consumer. The consumer has Cobb-Douglas preferences over z_i and $\sum_j \psi_{ij} q_{ij}$, implying constant budget shares $1 - \varphi$ and φ allocated to z_i and the products indexed by j. The form of the quality index implies that θ governs the relative importance of quantities z_i and q_{ij} compared with determinants of quality (δ and ϵ in the consumer's preferences.

The conditional direct utility function when choosing product j—which implies $q_{ij} > 0$ and $q_{i,k\neq j} = 0$—can be written as:

$$u_{ij} = u(\psi_{ij} q_{ij}, z_i) = (1-\varphi)\ln z_i + \varphi \ln q_{ij} + \ln \psi_{ij}$$

which implies that quantity q_{ij} and the quality components—δ_j and ϵ_{ij} from $\ln \psi_{ij}$—are additively separable in utility for a given choice j. Maximizing the conditional direct utility subject to a budget constraint $p_j q_{ij} + z_i = y_i$ gives the conditional ordinary demand functions:

$$\begin{aligned} q_{ij}(p_j, y_i) &= \frac{\varphi y_i}{p_j} \\ z_i(p_j, y_i) &= (1-\varphi) y_i \end{aligned}$$

Inserting into the conditional direct utility, we obtain the conditional indirect utility function:

$$\tilde{v}_{ij} = \ln y_i - \varphi \ln p_j + \frac{\delta_j + \epsilon_{ij}}{\theta}$$

which after rescaling by θ becomes:

$$v_{ij} = \theta \ln y_i - \theta\varphi \ln p_j + \delta_j + \epsilon_{ij}$$

where the consumer chooses the product j as $\max_j v_{ij}$, while quantities demanded conditional on this choice are given by the conditional ordinary demand functions above and zero for all non-chosen products.[37] Defining $\alpha = \theta\varphi$ and substituting out θ in the conditional indirect utility function, specifying δ_j as a function of observed and unobserved product characteristics, and appropriately redefining the random utility term, we obtain the utility function from equation (4) in Section 6.1.

C.2 Market Size Approximation

In our demand model, we normalize utility of the outside good to zero, and model the expenditure share of each product as a function of observable characteristics. From observed data on quantities and prices, one needs to take a stance on the market size in terms of total potential budget spent in market m at time t, $B_{mt} \equiv \varphi_{mt} Y_t$. This variable plays a fully analogous role in our discrete-continuous demand model as quantity-based market size in the commonly used unit-demand models.

Since the potential budget is not observed and we do not have any measures that would

[37] Note that we ignore the constant $\varphi \ln \varphi + (1-\varphi)\ln(1-\varphi)$, which is common across all alternatives in the conditional indirect utility function.

allow us to directly infer it, such as the number of individuals with a diagnosis for which a given drug would be relevant, we approximate B_{mt} for each ATC-5 code m and year t using the insights of Huang and Rojas (2013, 2014). Specifically, we adapt the procedure used by Dubois and Lasio (2018) to the specific model we use. We start by specifying a simple logit version of our model, where we use the expenditure share equations:

$$\ln s_{jmt} - \ln s_{0mt} = v_{jmt} - \alpha \ln p_{jmt} + x'_{jmt}\beta + \xi_{jmt}, \qquad s_{jmt} = \frac{p_{jmt}q_{jmt}}{B_{mt}}$$

for two observed, distinct goods j and j' to derive:

$$\ln p_{jmt}q_{jmt} - \ln p_{j'mt}q_{j'mt} = v_{jmt} - v_{j'mt} + -\alpha(\ln p_{jmt} - \ln p_{j'mt}) + (x_{jmt} - x_{j'mt})'\beta + \xi_{jmt} - \xi_{j'mt} \quad (12)$$

which does not depend on the potential budget. The parameters of this equation can be consistently estimated by 2SLS using similar instruments to those introduced in Section 6.3. In a second step, we use the equation:

$$\ln p_{jmt}q_{jmt} - \ln p_{0mt}q_{0mt} = \hat{v}_{jmt} - \hat{\alpha} \ln p_{jmt} + x'_{jmt}\hat{\beta} + \xi_{jmt} \quad (13)$$

where we take the parameters estimated from equation (12) as given, and $\ln p_{0mt}q_{0mt}$ is the log of the expenditure allocated to the outside good, which is common for all products j in market m in year t. Note that the potential budget by definition is equal to the expenditure on all observed alternatives in addition to the expenditure on the outside good, therefore knowing the latter is equivalent to knowing the potential budget with typical price-quantity data.[38] We specify expenditures on the outside good as:

$$\ln p_{0mt}q_{0mt} = \sum_{r=0}^{4} \varsigma_{rm} t^r + \varsigma_p \overline{\ln p}_{mt} + \overline{x}'_{mt}\varsigma_x \quad (14)$$

where the first sum is a fourth-degree polynomial in time (including a constant) that is allowed to differ across markets, and $\overline{\ln p}_{mt}$ and $\overline{x}_{mt}$ are the average log price and average characteristics in market m and year t. The average terms capture the effect of the attractiveness of the inside goods in competition with the outside good. This equation is estimated using a modified GMM-IV procedure, minimizing:

$$O(\varsigma) = G(\varsigma)'WG(\varsigma) + \varsigma'\mathbf{\Lambda}\varsigma$$

where $G(\varsigma) = Z'\xi(\varsigma)$ is the vector of moments, Z is a matrix of the exogenous variables in equa-

[38] It can be noted that only the aggregate expenditure on the outside good is important in this model, meaning that no assumptions on p_{0mt} is necessary.

tion (14) including price in Norway for the corresponding ATC and year, and W is the weighting matrix for the moments. The quadratic form in the parameter vector ς is a regularization term, with $\mathbf{\Lambda}$ as a diagonal matrix of regularization parameters. We allow for two different regularization parameters, one for parameters controlling the market-time polynomials, and one for the parameters controlling the effect of average characteristics. We add regularization due to the large number of parameters in equation (14), and calibrate the regularization parameters to minimize RMSE using cross-validation.

C.3 Aggregation of Unbranded Drugs for Demand Estimation

Since the data on sales from IQVIA only contain sales for unbranded generics at the aggregate level, we use the average expenditure share $s_{ut} \equiv \frac{1}{N_{ut}} \sum_{j \in \mathcal{J}_u} s_{jt}$ for this segment as the observed expenditure share, implicitly utilizing only the information in the choice of the segment for estimation, and not of particular products within the segment. Note that the expenditure share of a product $j \in \mathcal{J}_g$ (omitting the subscript m for market) can be written as:

$$s_{jt} = \frac{e^{\frac{\delta_{jt}}{1-\sigma}}}{e^{I_{gt}}} \frac{e^{(1-\sigma)I_{gt}}}{e^{I_t}}$$

where $I_{gt} \equiv \ln \sum_{j \in \mathcal{J}_g} e^{\frac{\delta_{jt}}{1-\sigma}}$ and $I_t \equiv \ln \sum_g e^{(1-\sigma)I_{gt}}$. Note that δ is here defined as the average utility *including* price, as opposed to the exposition in C.1. Define $\delta_{jt} \equiv \delta_{gt} + v_{jt}$, which decomposes δ_{jt} into a group average and deviation from this average, where the first term is:

$$\delta_{gt} = v_{gt} - \alpha \ln p_{gt} + x'_{gt}\beta + \lambda_t + \xi_{gt}$$

where any variable with subscript gt, say z_{gt}, is the average of z_{jt} across products in group g, and the second term is:

$$v_{jt} = \tilde{v}^g_{jt} - \alpha \widetilde{\ln p}^g_{jt} + \tilde{x}'_{gt}\beta + \tilde{\xi}^g_{jt}$$

where $\tilde{z}^g_{jt} = z_{jt} - z_{gt}$ measure deviation of the value of variable z_{jt} for a firm relative to the group average. We can then write:

$$I_{gt} = \frac{\delta_{gt}}{1-\sigma} + \ln N_{gt} + \ln \left(\frac{1}{N_{gt}} \sum_{j \in \mathcal{J}_g} e^{\frac{v_{jt}}{1-\sigma}} \right)$$

where N_{gt} is the number of products in group g at t. Note that the last term is zero if $v_{jt} = 0 \quad \forall j \in \mathcal{J}_g$, and that it is positive in any other case. We can then write:

$$\begin{aligned}
\ln N_{ut} s_{ut} &= \ln \sum_{j \in \mathcal{J}_u} s_{jt} \\
&= (1-\sigma) I_{ut} - I_t \\
&= \delta_{ut} + (1-\sigma) \ln N_{ut} + (1-\sigma) \ln \left(\frac{1}{N_{ut}} \sum_{j \in \mathcal{J}_u} e^{\frac{v_{jt}}{1-\sigma}} \right) - I_t
\end{aligned}$$

This expression leads to the following estimating equation for the unbranded segment:

$$\ln s_{ut} - \ln s_{0t} = v_{ut} - \alpha \ln p_{ut} + x'_{ut}\beta + \lambda_t - \sigma \ln N_{ut} + \nu_{ut},$$

with the term $-\sigma \ln N_{ut}$ reflecting the average conditional market share $1/N_{ut}$ among unbranded generics, and the composite error term

$$\nu_{ut} = \xi_{ut} + (1-\sigma) \ln \left(\frac{1}{N_{ut}} \sum_{j \in \mathcal{J}_u} e^{\frac{v_{jt}}{1-\sigma}} \right).$$

C.4 Compensating Variation

Denote $u_{ij} \equiv u(y_i, \delta_j, \xi_j, \nu_{ij})$, where the dependence on market m and t is omitted for notational convenience, and $\nu_{ij} = \zeta_i^k + (1-\sigma)\epsilon_{ij}$. Consider a policy that changes the initial $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_J)$ to $\boldsymbol{\delta}'$ while ξ, y and the distribution of ν stays fixed. The compensating variation for individual i is defined as CV_i such that:

$$\max_j u(y_i, \delta_j, \xi_j, \nu_{ij}) = \max_{j'} u(y_i - CV_i, \delta'_{j'}, \xi_{j'}, \nu_{ij'})$$

such that the income reduction in the post-policy situation necessary to equate maximum utility before and after the policy change. With the preferences given above, (log of) income is additively separable, and we can write:

$$\begin{aligned}
\alpha\varphi^{-1}[\ln y_i - \ln(y_i - CV_i)] &= \max_{j'}\{\delta'_{j'} + \nu_{ij'}\} - \max_j\{\delta_j + \nu_{ij}\} \equiv \Delta_i \\
CV_i &= y_i(1 - e^{-\varphi\alpha^{-1}\Delta_i})
\end{aligned}$$

The expected compensating variation in the market can be obtained by integrating over y and ν. Assuming that y and ν are independently distributed and thus $F_{y,\nu}(y,\nu) = F_y(y)F_\nu(\nu)$, we

get:

$$\mathrm{E}[CV] = \int y \, \mathrm{d}F_y(y) \int (1 - e^{-\varphi \alpha^{-1} \Delta_i}) \, \mathrm{d}F_\nu(\nu) = \bar{y} \int (1 - e^{-\varphi \alpha^{-1} \Delta_i}) \, \mathrm{d}F_\nu(\nu),$$

where $\bar{y}$ is the average income and the integral over the random utility components can be approximated using the procedure suggested by McFadden (1996), which can be simplified for this case as follows:

1. Draw a sequence of vectors ν^t for $t = 1, \ldots, T$, which empirical distribution approximates F_ν as $T \to \infty$.

2. For each draw, compute the maximum indirect utility for the baseline and counterfactual policy scenarios to obtain:

$$\Delta(\nu^t) \equiv \max_j \{\delta_j + \nu_j^t\} - \max_{j'} \{\delta'_{j'} + \nu_{j'}^t\}$$

3. Approximate $\mathrm{E}[CV]$ as:

$$\overline{CV} = \bar{y} \frac{1}{T} \sum_{t=1}^{T} (1 - e^{-\varphi \alpha^{-1} \Delta(\nu^t)})$$

C.5 Model Simulation Details

To implement our counterfactual analysis in Section 7, we simulate equilibrium outcomes using our estimated model. We start from a dataset that includes all potential entrants to a market, their segments, and product attributes; along with estimates of demand, marginal costs, entry costs, certification costs, the distribution of drug quality and the scale of profit shocks. For each market in this dataset, we proceed as follows:

1. To start simulation s, take N_{MS} draws of potential market structures, from the distribution of market structures.

2. Take a draw of profit shocks $(\tilde{\varepsilon}_{0j}^{(s)}, \tilde{\varepsilon}_{1j}^{(s)})$ for each potential entrant. Take a draw of quality $\tilde{\psi}_j^{(s)}$ for each potential entrant.

3. For each market structure simulated in step 1, solve for optimal pricing and compute variable profits for all entrants using our demand and marginal cost estimates. Compute expected variable profits for each potential entrant across all draws of potential market structures.

4. Using expected variable profits from step 3, estimates of fixed costs—and certification costs when considering environments with quality regulation—, draws of drug quality

$\tilde{\psi}_j^{(s)}$ from step 2, and an initial guess for entry probabilities as inputs, solve for a fixed point in entry probabilities.

5. Using entry probabilities from step 4 and draws of profit shocks $(\tilde{\varepsilon}_{0j}^{(s)}, \tilde{\varepsilon}_{1j}^{(s)})$ from step 2, compute actual market structure, solve for optimal price and compute demand given that market structure.

6. Repeat this procedure from 1 to 5 for $s = 1, \ldots, N_S$ runs, so we can then average across simulations to smooth the realizations of shocks. The results in the paper average across $N_S = 50$ runs of this procedure.

To obtain the results in the paper, we use this procedure to generate a baseline market structure over which to develop our counterfactual analysis. We proceed as follows:

1. We start from an environment without quality regulation, to simulate a baseline market structure. This market structure essentially accounts for the fact that when the regulation is imposed, there is already a particular market structure in place, and that is generated by the same underlying parameters for demand, marginal costs, quality, and entry costs.

2. Taking the baseline market structure as given and as known to all firms, we simulate market outcomes again under different environments that range from no regulation to particular forms of regulation. Depending on the counterfactual, we adjust different parameters of the model for each simulation. The results we report in the paper correspond to equilibrium outcomes from this second simulation.

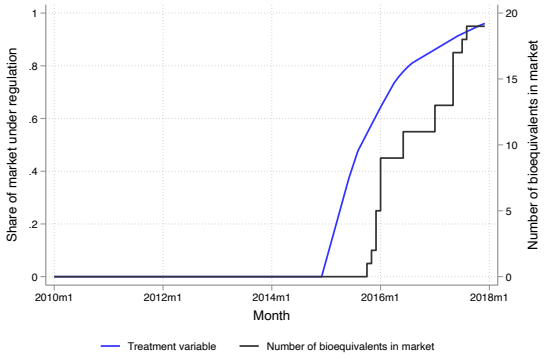
Figure A.1: Labeling of bioequivalent drugs

(a) Instructions for bioequivalent drugs labeling

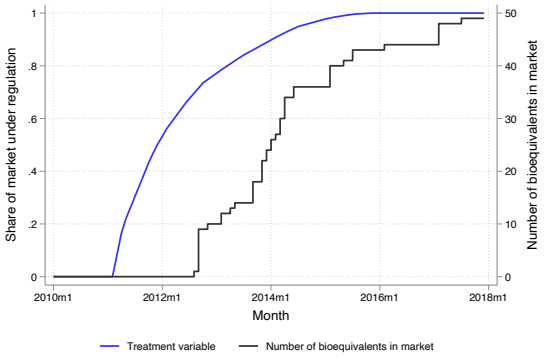
(b) Examples of labeled bioequivalent drugs

Notes: These figures display both instructions and examples of required labeling of bioequivalent drugs. The objective of this labeling was to highlight drugs with bioequivalence approval.

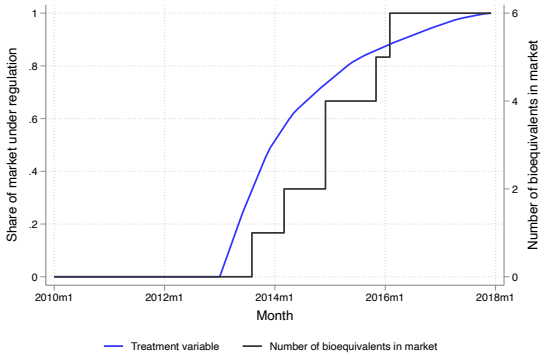
Figure A.2: Policy variation induced by bioequivalence requirements



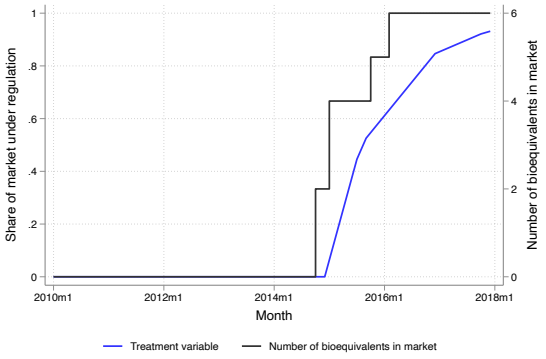
(a) Aripiprazole



(b) Atorvastatin



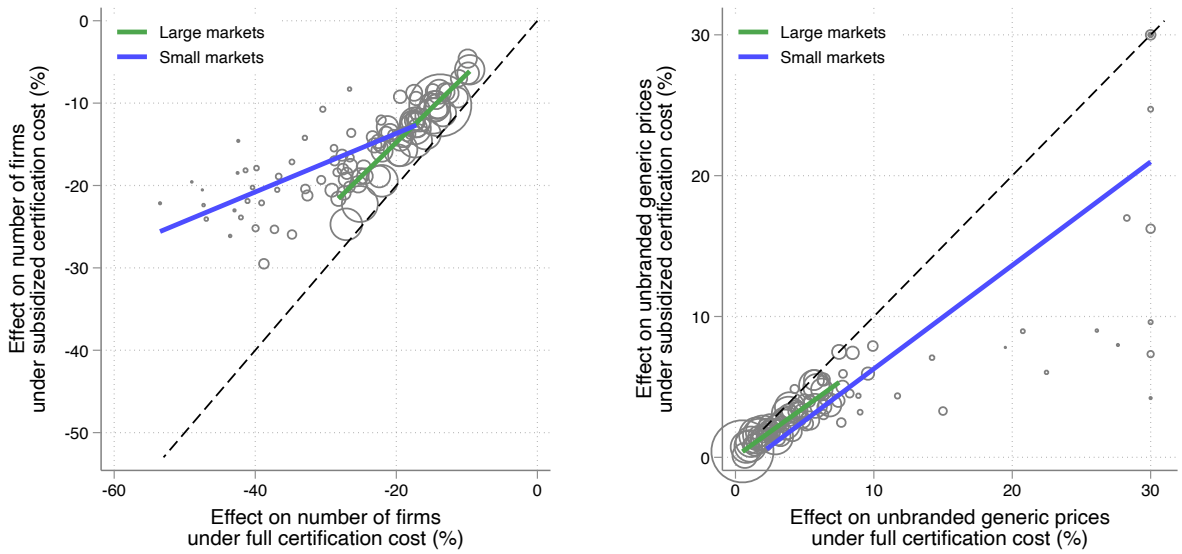
(c) Citalopram



(d) Deflazacort

Notes: Each figure displays the values of the treatment variable and the number of bioequivalent drugs in a different market. This version of the treatment variable uses the first deadline as the relevant date. The instrument is displayed in blue and takes a value of 0 before the first decree, and then increases as renewal dates of drugs in the molecule approach. The number of bioequivalent drugs in the molecule is displayed in gray. These four examples are plotted along all other markets in our sample in Figure 3-c.

Figure A.3: Heterogeneity in simulated effects and the role of certification costs

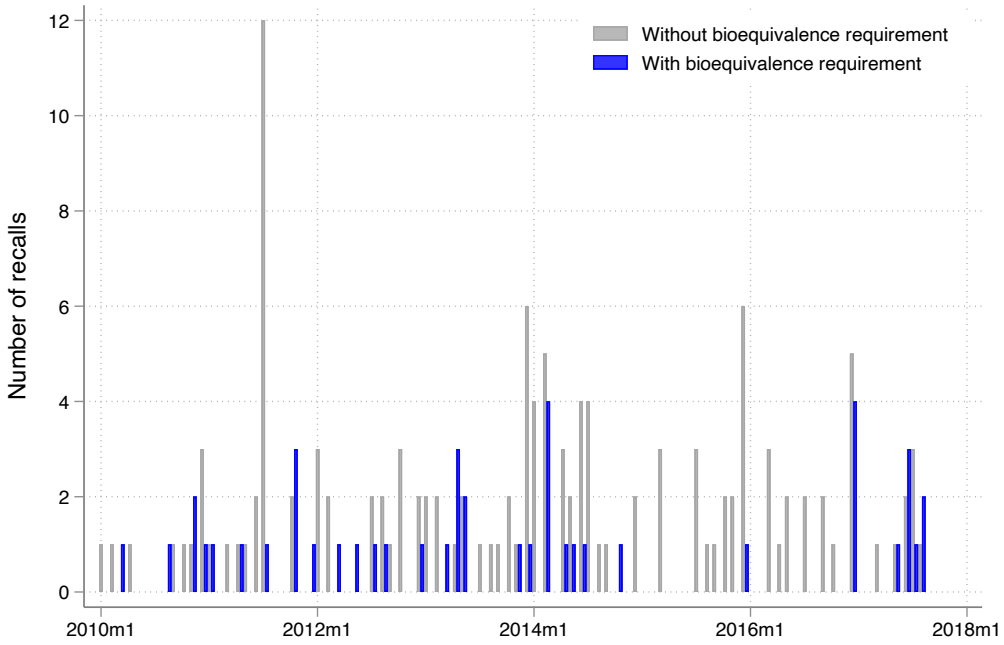


(a) Number of firms

(b) Price of unbranded generics

Notes: These figures compare simulated effects of quality regulation across environments where firms must pay the full certification cost (x-axis), and where the certification cost is fully subsidized (y-axis). Each circle is a market, and the size of each circle indicates market size. The green (blue) line is a linear fit for large (small) markets, defined around the median market size. The dashed line is a 45-degree line. Panel (a) displays results for the effect of quality regulation on the number of firms in the market, Panel (b) displays results for the effect of quality regulation on the average price of unbranded generics in the market.

Figure A.4: Number of recalls per month



Notes: The figure shows the number of product recalls over time split into markets with bioequivalence requirements and markets without bioequivalence requirements.

Figure A.5: Consumer survey elicitation of perceived quality and price

Notes: This figure displays the sheet surveyors provided consumers in our survey sample. This sheet displays the 4 drugs we used as an example to elicit perceived quality and price differences. While observing this sheet, surveyors asked consumers first to assign a score in a 1-7 scale to each drug regarding their quality, and then to estimate the price of each drug given that the innovator had a price of $50,000 CLP ($77.5 USD).

Table A.1: Effects of quality regulation on number of drugs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dep. var.: $\sinh^{-1}$(Number of drugs)							
	All	Innovator	Branded generics			Unbranded generics		
			All	BE	Non-BE	All	BE	Non-BE
A - Average effects								
Regulation	-0.25***	-0.06	-0.26***	0.60***	-0.41***	-0.35***	0.76***	-0.52***
	(0.05)	(0.04)	(0.05)	(0.22)	(0.08)	(0.11)	(0.16)	(0.11)
R^2	0.94	0.93	0.96	0.71	0.94	0.88	0.65	0.88
B - Heterogeneity by market size								
Regulation × Low revenue	-0.39***	-0.17**	-0.37***	0.09	-0.39***	-0.57***	0.32**	-0.53***
	(0.07)	(0.08)	(0.08)	(0.23)	(0.12)	(0.13)	(0.16)	(0.14)
Regulation × High revenue	-0.14***	0.03	-0.18***	0.97***	-0.43***	-0.20	1.07***	-0.51***
	(0.05)	(0.04)	(0.06)	(0.25)	(0.09)	(0.12)	(0.20)	(0.11)
R^2	0.94	0.93	0.96	0.73	0.94	0.89	0.67	0.88
Observations	11,040	11,040	11,040	11,040	11,040	11,040	11,040	11,040
Market FE	Y	Y	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y	Y	Y

Notes: Each column in this Table is a regression of the inverse hyperbolic sine of the number of drugs in a segment on the policy roll-out variable constructed using the first decree deadline. Panel B provides results by baseline revenue. Markets are classified as having a low or high revenue according to the total revenue in the market in 2010 relative to the median across markets in that year. Clustered standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table A.2: Effects of quality regulation on proxies of drug quality

	(1)	(2)	(3)	(4)
	Drug adverse effects			Drug
	Admissions	Hospital days	Surgeries	recalls
A - Average effects				
Regulation	-0.023	-0.120	-0.000	-0.002
	(0.023)	(0.112)	(0.000)	(0.002)
R^2	0.849	0.869	0.142	0.170
B - Heterogeneity by market size				
Regulation × Low revenue	-0.072	-0.235	-0.000	-0.000
	(0.071)	(0.224)	(0.000)	(0.002)
Regulation × High revenue	0.022	-0.016	-0.000	-0.003
	(0.024)	(0.014)	(0.000)	(0.003)
R^2	0.850	0.871	0.142	0.170
Pre-regulation average	0.073	0.132	0.000	0.000
Observations	568	568	568	867
Market FE	Y	Y	Y	Y
Month FE	Y	Y	Y	Y

Notes: Each column in this table is an outcome related to drug quality on the policy roll-out variable constructed using the first decree deadline, as in equation (11). Outcomes are constructed as the ratio of the variable of interest over drug sales measured in thousands of daily doses. Columns 1-3 are related to adverse health effects, whereas column 4 is related to drug recalls, and includes all recalls. Panel B provides results by baseline revenue. Markets are classified as having a low or high revenue according to the total revenue in the market in 2010 relative to the median across markets in that year. Clustered standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table A.3: Summary statistics from consumer survey data

Variable	N	Mean	SD	p10	p50	p90
Perceived quality of innovator drug (1-7)	361	6.32	1.01	5.00	7.00	7.00
Perceived quality of bioequivalent branded drug (1-7)	378	5.69	1.31	4.00	6.00	7.00
Perceived quality of bioequivalent unbranded drug (1-7)	386	5.63	1.28	4.00	6.00	7.00
Perceived quality of non-bioequivalent unbranded drug (1-7)	381	4.68	1.65	3.00	5.00	7.00
Perceived price of bioequivalent branded drug (CLP 1,000s)	398	25.37	14.13	6.00	25.00	45.00
Perceived price of bioequivalent unbranded drug (CLP 1,000s)	401	15.69	10.98	3.00	15.00	30.00
Perceived price of non-bioequivalent unbranded drug (CLP 1,000s)	399	12.60	9.97	2.00	10.00	25.00
Recognizes bioequivalent drug label	401	0.84	0.37	0.00	1.00	1.00
Understanding about bioequivalence (1-5)	401	2.91	1.47	1.00	3.00	5.00
=1 if physicians specify brand in prescriptions	299	0.65	0.48	0.00	1.00	1.00
=1 if always purchases physician recommendation	310	0.15	0.36	0.00	0.00	1.00
=1 if sometimes deviate from physician recommendation	310	0.52	0.50	0.00	1.00	1.00
=1 if always chooses cheapest available drug	310	0.34	0.47	0.00	0.00	1.00
Purchases innovator drugs	338	0.41	0.49	0.00	0.00	1.00
Purchases bioequivalent branded drugs	338	0.20	0.40	0.00	0.00	1.00
Purchases bioequivalent unbranded drugs	338	0.28	0.45	0.00	0.00	1.00
Purchases non-bioequivalent unbranded drugs	338	0.11	0.31	0.00	0.00	1.00
Chronic illness by household member	401	0.58	0.49	0.00	1.00	1.00
Atorvastatin consumption by household member	401	0.34	0.48	0.00	0.00	1.00

Notes: This table displays summary statistics from our consumer survey. The total number of surveys is $N = 401$. Whenever the number of observations is smaller, is due to the consumer not answering the question, except for the case of questions regarding physicians' prescription behavior, which were added to the survey with a lag and are therefore not available for around a fourth of the sample.