

Demand Responsiveness and International Pharmaceutical Pricing

Federico Nutarelli^{a,*}

Duygu Buyukyazici^b

Massimo Riccaboni^{a,c}

^a IMT School for Advanced Studies Lucca, Italy

^b London School of Economics and Political Science, London, United Kingdom

^c IUSS Pavia, Pavia, Italy

Abstract

We examine whether international pharmaceutical prices are higher where demand is less price responsive. Using data from 63 countries over 2008–2020, we adapt double machine learning to estimate own- and rival-price elasticities that vary across countries, medicines, and years. We establish conditions under which a rich price-response specification approximates responses from identified logit-based demand models and characterize identification limits of unrestricted structural specifications. Estimated responsiveness varies substantially across markets and is not consistently captured by GDP per capita. Robustness checks against elasticities from structural models are presented. Higher branded-drug prices are associated with less responsive demand, consistent with inverse-elasticity ordering, while, as expected in a perfect competitive environment, generic products show no similarly stable relationship. A historical exercise motivated by U.S. Most-Favored-Nation pricing illustrates why these findings matter for policy. On the same sample, the median U.S.-to-reference price ratio is 6.95 using the lowest eligible foreign price, compared with 1.62 using a foreign benchmark that accounts for U.S. demand responsiveness and country characteristics. Our results show how directly estimated elasticities can inform international price benchmarks beyond comparisons based on income alone.

*Corresponding author: federico.nutarelli@imtlucca.it, IMT School for Advanced Studies Lucca, Lucca, Italy.

Acknowledgements: This research was supported in part by MSD’s Project on Price Transparency. The empirical analysis uses data obtained under license from the IQVIA MIDAS™ database. We would like to thank these sponsors for making the research possible. The conclusions and views expressed herein are not necessarily those of MSD or IQVIA. We would like to thank Pierre Dubois, Angélique Acquatella, Martin Huber, and Chuqing Jin for fruitful discussions. We also thank Victor Chernozhukov and his team, who developed the first implementation of the DML algorithm, for their helpful replies since the early versions of the EconML package.

1 Introduction

Pharmaceutical innovation inherently creates a classic second-best pricing problem. Developing a medicine entails substantial R&D costs that are incurred before production, cannot be readily attributed to the patients or countries that ultimately use it, and therefore require prices above marginal cost for their recovery (Danzon, 1997; Danzon, 2018). This raises a fundamental question: how should the common cost of pharmaceutical innovation be allocated across national markets? The answer is not straightforward, because pharmaceutical pricing must reconcile two competing objectives: generating sufficient revenues to sustain investment in pharmaceutical innovation and the development of breakthrough medicines, while preserving access to medicines, particularly for low-income populations. In theory, differential pricing can help reconcile these objectives by allowing prices to vary across markets according to differences in demand conditions and ability to pay. Pharmaceutical markets have therefore become one of the most prominent settings in which international price discrimination is both practiced and debated (Maleug and Schwartz, 1994; Danzon, 1997; Jack and Lanjouw, 2005; Towse et al., 2015; Towse et al., 2018; Danzon, 2019).

However, the success of international price discrimination to reconcile these two objectives depends not only on allowing prices to vary across countries, but also on how the resulting price–cost margins are distributed across markets. The same proportional increase in price may have little effect on medicine use in one market but induce a substantial contraction in another. An economically relevant allocation of common R&D costs must therefore account for differences in the price responsiveness of demand across markets. The Ramsey–Boiteux inverse-elasticity principle provides a canonical benchmark for this allocation problem. Conditional on a given revenue requirement, proportional price–cost margins should be higher in markets with less elastic demand and lower in markets with more elastic demand, thereby raising the required revenue while limiting the aggregate contraction in consumption (Ramsey, 1927; Baumol and Bradford, 1970). The pharmaceutical-pricing literature has applied this logic to the international recovery of common R&D costs (Danzon, 1997; Barros and Martinez-Giralt, 2008). Empirically, however, the inverse-elasticity principle has often been assessed indirectly, by examining whether pharmaceutical prices increase with country income and treating income as a proxy for differences in demand responsiveness (Yadav, 2010b). Such an approach leaves the central object of the inverse-elasticity rule unobserved: the responsiveness of medicine use to price.

The international allocation of pharmaceutical R&D costs and the role of demand responsiveness in that allocation has acquired renewed policy relevance in recent years. In May 2025, Executive Order 14297 in the U.S. directed the Administration to pursue Most-Favored-Nation (MFN) pricing for prescription drugs. The subsequent proposal would benchmark certain branded medicines without generic or biosimilar competition to the lowest price observed in an OECD country whose GDP per capita PPP is at least 60 percent of the U.S. level (Executive Office of the President, 2025; HHS, 2025). The proposal is motivated both by the large gap between U.S. and foreign pharmaceutical prices and by the claim that U.S. patients bear a disproportionate share of the global cost of pharmaceutical innovation. Recent estimates, for example, place U.S. branded-drug prices at more than three times those in comparison countries even after accounting for estimated rebates (Parasrampuria and Murphy, 2024).

The MFN policy framework makes the choice of an international pricing benchmark central. GDP per capita PPP is a natural measure of economic comparability, but it does not necessarily capture the demand responsiveness that matters for differential pricing. Countries with similar income levels may differ substantially in insurance coverage, reimbursement, procurement, prescribing behaviour, substitution rules, bargaining institutions, and product availability, all of which can shape how medicine use responds to price. As a result, similarly wealthy countries need not have similarly elastic demand for the same medicine, while countries at different income levels may display comparable price responsiveness. The same issue arises

in the policy’s burden-sharing argument. To say that one country contributes disproportionately to the cost of pharmaceutical innovation requires a benchmark for how those common costs should be allocated across markets. The lowest foreign price establishes that the U.S. price is higher, but not whether the entire difference represents an excessive contribution. Under an inverse-elasticity benchmark, a lower price abroad may partly reflect more price-responsive demand and the larger reduction in medicine use that a higher price would induce. Conversely, if two markets have similar demand elasticities, differences in price responsiveness cannot account for a large price gap. Price comparisons and income differences alone, therefore, do not reveal how much of observed international price dispersion is related to heterogeneity in demand responsiveness.

Building on these theoretical and policy motivations, this paper examines whether international pharmaceutical price variation is associated with heterogeneity in demand responsiveness. Using IQVIA MIDAS data covering 47,589 products, 171 molecules, and 63 countries from 2008 to 2020, we first estimate, to our knowledge, the most comprehensive set of comparable country–molecule–year own- and rival-price elasticities of product-level demand across national pharmaceutical markets, and document how demand responsiveness varies across countries, molecules, and over time. Second, we examine whether global price dispersion is systematically aligned with this heterogeneity—in particular, whether raw and residualized prices are higher in markets where demand is less price responsive, as implied by the inverse-elasticity benchmark. Finally, for the branded segment targeted by the MFN proposal, we compare the lowest eligible foreign price with an elasticity-aligned U.S. price predicted from the relationship between molecule prices, demand responsiveness, and observable country characteristics in non-U.S. markets. Together, these exercises provide a demand-based benchmark for interpreting international pharmaceutical price differences, rather than relying on income alone as a proxy for price responsiveness.

Delivering such a demand-based international benchmark requires an empirical framework that combines comparability across markets with sufficient flexibility to accommodate heterogeneous demand responses. Our methodological approach is therefore informed by three complementary strands of literature: the pharmaceutical differential-pricing literature, which identifies price responsiveness as the relevant object for allocating common R&D costs (Danzon, 1997; Barros and Martinez-Giralt, 2008); the empirical industrial-organization (IO) literature on differentiated-product demand, which provides structural methods for recovering own- and cross-price responses (Coscelli, 2000; Crawford and Shum, 2005; Dubois and Laisio, 2018); and the recent literature on high-dimensional semiparametric estimation (Semenova et al., 2023), which allows these responses to vary flexibly with a large set of observed characteristics¹.

Existing approaches in the pharmaceutical differential-pricing literature and empirical IO face a trade-off between breadth and structural detail. Cross-country studies of pharmaceutical pricing achieve broad coverage largely by restricting the dimensions of heterogeneity or by relying on proxies such as income. Structural demand models of empirical IO, by contrast, typically recover richer substitution patterns but require the researcher to specify markets, outside options, utility indices, and the relevant dimensions of preference heterogeneity. These modeling choices are particularly consequential in a global pharmaceutical panel, where observed demand may reflect different combinations of decisions by patients, physicians, pharmacists, insurers, purchasers, and regulators across countries and therapeutic markets. Estimating separate structural systems country by country would not eliminate this specification problem and would make international comparability depend on a collection of country-specific modeling choices.

We address this tension using Double Machine Learning (DML) to estimate heterogeneous price responses within a common but flexible specification. Our methodological contribution is twofold. First, we establish (sufficient) conditions under which a DML with a sufficiently rich price-response dictionary can approximate the price-response component of canonical structural demand models used in empirical

¹The reader is referred to Appendix A for a schematic literature review.

IO—namely, a simple logit with an extension to random-coefficients logit (RCL) model—and, under additional regularity conditions, its implied elasticity function. This result provides a first—to the best of our knowledge—formal link between the semiparametric price-response object estimated by DML and the elasticity patterns generated by economically disciplined structural demand models. In the empirical part—after harmonizing our set-up with that in Dubois et al. (2022)—we compared RCL estimates with DML estimates. The average DML own-price elasticities are of the same order of magnitude as the structural demand benchmarks, although DML reveals greater variation across therapeutic classes and product types. Second, we show that an unrestricted RCL specification need not identify a unique elasticity vector when the linear mean-utility block becomes high-dimensional relative to the available observations or instruments. This boundary result is particularly relevant when the dimensions governing pharmaceutical demand are uncertain *ex ante*. Rather than imposing a narrow utility specification from the outset, DML allows the researcher to begin with a broad set of candidate controls and sources of heterogeneity and, under approximate sparsity, recover a parsimonious representation of the price-response function. We therefore use DML as our main estimator of international pharmaceutical price responsiveness and retain RCL as the structural benchmark where a harmonized comparison is feasible.

Our empirical findings show what directly estimated demand responses add to international comparisons based on prices and income alone. First, we find substantial heterogeneity in price responsiveness across countries and molecules. Country-level mean own-price elasticities are negative in all 63 countries and range from approximately -0.73 to -0.26 . We next examine what accounts for this heterogeneity. Restricting the comparison to molecules observed across a common set of countries changes the country ranking and leaves substantial dispersion. Differences in which molecules are marketed, therefore, affect unconditional country comparisons, but do not explain all of the cross-country variation. Molecule rankings are considerably more stable under the same restriction, indicating that molecule-level heterogeneity is not driven primarily by differences in country coverage. Country income does not provide a satisfactory alternative explanation. Indeed, the relationship between GDP per capita and estimated demand responsiveness is modest and varies across estimators and product-composition samples. GDP therefore does not consistently rank countries by the demand elasticity relevant to differential pricing ²

We then ask whether observed prices are ordered by demand responsiveness in the direction suggested by the inverse-elasticity benchmark ³. We found that the price–elasticity relationship is concentrated in branded markets. Across the alternative estimators, price measures, and country and molecule comparisons, higher branded-drug prices are consistently associated with less responsive demand, with a stronger pattern for residualized prices. Generic products do not exhibit a similarly stable relationship. This contrast is consistent with inverse-elasticity ordering being more informative where firms have greater discretion over price–cost margins. Generic prices may instead be more directly constrained by entry, procurement, reference pricing, regulation, and competition (Vogelsang, 2003; Armstrong and Vickers, 2018).

Finally, we use these estimates to interpret the historical price compression implied by an international reference-price rule. For molecule–periods without locally observed generic competition, the pseudo-MFN exercise applies the policy’s OECD and GDP eligibility criteria and selects the lowest eligible foreign price. We compare that minimum with an elasticity-aligned U.S. price: the price predicted for the United States

²As an alternative-identification check, Appendix C compares a pooled DML estimate with a pooled specification that instruments domestic prices using foreign-price variation. The two approaches produce an own-price response with the same sign and a broadly similar magnitude.

³Since the theoretical benchmark concerns price–cost margins – and marginal costs are not observed – we examine whether higher-price country–year environments and higher-priced molecule–years are associated with less responsive demand, using both observed prices and the component of prices not predicted by the demand-stage controls. An auxiliary U.S. structural demand and pricing exercise in Appendix K.2 provides supporting evidence that these controls absorb a large share of model-implied marginal-cost variation.

by the relationship between prices, demand responsiveness, and observed characteristics in eligible foreign markets.

Across 116 retained U.S. molecule–period cells, the observed U.S. price exceeds the pseudo-MFN minimum in 96.6 percent of cases, but exceeds the elasticity-aligned fitted price in only 55.2 percent. The fitted price is above the eligible-country minimum in every cell. Thus, the historical U.S.–pseudo-MFN gap reflects both the position of the U.S. price relative to the foreign price–responsiveness relationship and the additional compression created by selecting the lowest eligible foreign price.

The remainder of the paper is organized as follows. Section 2 presents the heterogeneous price-response model, estimation procedure, and identification assumptions. Section 2.2 develops the relationship between DML and logit-based structural demand models, states the approximation and rank results, and reports the simulation evidence. Section 3 describes the data and how DML adapts to our setting. Section 4 presents the elasticity estimates and the comparison with the structural benchmark. Section 5 examines the relationship between estimated price responsiveness and pharmaceutical prices and applies the resulting elasticity-aligned benchmark in a historical pseudo-MFN exercise. Section 6 concludes. The appendices contain the proofs, simulation design, structural-benchmark details, additional validity checks, marginal-cost diagnostics, and the continuous-choice Ramsey framework.

2 Estimating Heterogeneous Pharmaceutical Demand

Differential pharmaceutical pricing has been proposed to reconcile access to medicines with the recovery of global R&D costs (Danzon and Towse, 2003). However, differential pricing is based on heterogeneous demand responsiveness. A major difficulty in assessing price responsiveness or whether its inverse-elasticity implication holds across many products and countries is due to the fact that demand factors may vary with a high-dimensional set of product, therapeutic, country, time, and institutional characteristics whose relevance is unknown *ex ante*. Existing reduced-form cross-country studies, therefore, often rely on a small set of products and ATCs (Berndt et al., 2007; Chintagunta and Desiraju, 2005) or use income-based proxies for price elasticity (Maskus, 2001; Sherer and Watal, 2002; Lichtenberg, 2010). The empirical relevance of the latter approximation, however, remains difficult to establish with several studies finding only weak or inconsistent relationships between prices and income (Waning et al., 2009; Morel et al., 2011; Petrou and Vondoros, 2016). On the theoretical side, GDP per capita is not a sufficient proxy for pharmaceutical price responsiveness. In standard consumer theory, the uncompensated price effect is decomposed into an income effect and a substitution effect (the well-known Slutsky decomposition). This is the main message of Appendix J. GDP per capita may proxy the former—purchasing power or ability to pay, but it does not capture the compensated substitution response or other determinants of pharmaceutical demand such as insurance coverage, reimbursement rules, public procurement, prescribing behavior, pharmacist substitution, reference-pricing systems, product availability, and the entry of generic competitors (Danzon, 2018). Countries with similar incomes may, therefore, exhibit very different price elasticities, while institutional arrangements may generate relatively elastic demand even in high-income markets.

Structural applications of pharmaceutical pricing instead typically examine narrower settings and recover elasticities from logit or RCL demand systems⁴. These models provide a coherent structural substitution

⁴The need to restrict the analysis to small samples is often due, among other challenges, to the choice of the market share of the outside option. A typical example is found in our setting and in Dubois and Lasio (2018) and Dubois et al. (2022), where the market is defined at such a fine level that demographic or population-level descriptors are unavailable. One possible approach, which we also

system, but require the researcher to specify the relevant market, the outside option, the utility index, and the distribution of random coefficients.⁵ These modeling choices are especially important in pharmaceutical markets. Observed pharmaceutical demand, indeed, may reflect patient learning, physician prescribing, pharmacist substitution, insurance coverage, and reimbursement rules (Coscelli, 2000; Crawford and Shum, 2005; Dubois and Lasio, 2018). The relative importance of these margins may vary across countries, therapeutic classes, and sales channels. With aggregate multi-country sales data, however, one cannot easily identify the preferences and decision margins of the different actors or select, without additional structural assumptions, a single utility specification that is appropriate across all these settings.⁶

The multi-country dimension makes this specification problem especially important. Defining, as an example, markets at the country–ATC–4–quarter level allows choice sets and mean utilities to vary across markets, but does not by itself determine which product and institutional characteristics generate substitution or price sensitivity. Country and market fixed effects can absorb differences in average mean utility, but do not by themselves generate country-specific price responsiveness. An RCL model could accommodate such heterogeneity through country-specific parameters and interactions, but the relevant dimensions must be selected *ex ante* and the resulting specification may become high-dimensional. In this environment, an unrestricted high-dimensional RCL specification can also encounter an identification problem.

Based on the aforementioned motivation, the empirical strategy presented in this section combines a flexible semiparametric approach to estimating heterogeneous demand with the structural demand framework established in empirical IO. We first adapt the many-product DML framework (Semenova et al., 2023) to estimate heterogeneous own- and rival-price responses across countries, molecules, and time (Section 2.1). We then relate this approach formally to the logit and RCL models commonly used in empirical IO (Section 2.2). In particular, we establish conditions under which a sufficiently rich DML price-response dictionary can approximate the price-response component, and under additional regularity conditions the elasticity function, implied by a fixed and identified logit or RCL demand model. We also characterize a high-dimensional boundary of the unrestricted RCL specification, showing why identification can become problematic when the set of potential demand characteristics is large relative to the available observations or instruments (Section 2.2.1). To the best of our knowledge, we provide the first formal bridge between the DML and RCL approaches, showing a novel dictionary approximation result for fixed logit-based demand functions in Proposition 2.1. These results theoretically motivate using DML as the primary estimator for the global analysis, while retaining RCL as an economically disciplined structural benchmark.

2.1 A heterogeneous price-response model

The paper asks whether international pharmaceutical prices are higher where demand is less price responsive, following the inverse-elasticity principle rooted in Ramsey (1927) and applied in the pharmaceutical literature as a benchmark for recovering common R&D costs through international price differentiation while preserving access to medicines (see previous sections and, among others, (Danzon and Towse, 2003; Danzon, 2018)). To answer this question, we estimate how own- and rival-price elasticities differ across coun-

follow in the present work, is that proposed by Huang and Rojas (2013); however, it relies on strong assumptions.

⁵Nonparametric discrete-choice approaches relax distributional assumptions, establishing conditions for point identification (Berry and Haile, 2014) or bounds on demand counterfactuals (Tebaldi et al., 2023). Lieber et al. (2026) develop more scalable methods for these bounds and assess sensitivity to distributional assumptions. Our DML analysis estimates own-price and rival-price-index responses without specifying a distribution of consumer preferences, while maintaining sequential conditional exogeneity and restrictions on response heterogeneity.

⁶For example, Dubois and Lasio (2018) interpret their aggregate RCL demand model as a reduced-form representation of the interaction among patients, prescribers, and pharmacists. Similarly, Dubois et al. (2022) restrict attention to hospital purchases, where the relevant purchasing decision is more clearly defined.

tries, molecules, and years. Capturing these differences in price responsiveness requires estimating many coefficients in the demand equation. It is easy to imagine that – when we also account for product characteristics, market conditions, and past prices and sales – together, these price-response coefficients and the extensive set of controls create a high-dimensional estimation problem (Belloni et al., 2014; Chernozhukov et al., 2022). These dimensionality concerns are enhanced by considerations on how these characteristics affect demand, since a given change in a characteristic may have different effects depending on its starting value, or on the country or product concerned, thus potentially increasing the terms to estimate and introducing nonlinearities. This is well established in demand estimation, where flexible specifications allow nonlinearities in demand, such as nonlinear expenditure effects and interactions between expenditure and household characteristics (Blundell et al., 2007; Lewbel and Pendakur, 2009). As a more focused example in the pharmaceutical context, Einav et al. (2015) show that drug purchases respond nonlinearly to insurance spending thresholds that change patients’ out-of-pocket prices.⁷ We believe that, in our context, this motivates a semiparametric and quasi-linear demand model in which (log) sales depend linearly on (log) own and rival prices – with coefficients that vary across countries, molecules, and years – and other observed determinants of demand enter through flexible functions that allow selected nonlinear transformations and interactions.

Given the above observations, recent developments in high-dimensional semiparametric econometrics and DML provide a natural framework for this type of problem, enabling the estimation of target parameters in the presence of rich, potentially data-adaptive nuisance functions (Chernozhukov et al., 2018). Within this literature, the many-product framework of Semenova et al. (2023) is particularly suited to our setting because it accommodates many price-response coefficients in panel data while accounting for persistent differences across products and dependence over time. In what follows, we adapt this framework to international pharmaceutical demand.

Let y_{igt} be log standard units and p_{igt} log price for product i , containing molecule g , in country c and year-quarter t . If $\mathcal{C}_{gct}$ is the set of products containing g in that country and quarter, define the leave- i -out rival-price index (p_{igt}^-) as the logarithm of the arithmetic mean of rival prices in levels (when at least one rival is observed), i.e.

$$p_{igt}^- = \log \left[\frac{1}{|\mathcal{C}_{gct}| - 1} \sum_{\substack{j \in \mathcal{C}_{gct} \\ j \neq i}} \exp(p_{jgt}) \right].$$

Let $T_{igt} = (p_{igt}, p_{igt}^-)'$, where the superscript “-” identifies the rival-price index, and let K_{igt} collect the variables whose interactions with prices describe how elasticities differ across observations. In the baseline specification, K_{igt} contains a constant and country, molecule, and calendar-year indicators, so each elasticity combines these components additively⁸. K_{igt} can be more general and also include interactions of the latter. Each cell’s elasticity is therefore the sum of a common component and country-, molecule-, and year-specific components, i.e. all products containing the same molecule in a given country and calendar year share the same own- and rival-price elasticities, including across quarters⁹. This specification keeps the

⁷This illustrates how nonlinear incentives can shape pharmaceutical demand, although our aggregate specification does not explicitly model patient-level insurance contracts.

⁸Additivity implies, for example, that the elasticity difference between two countries is the same across molecules in a given year. This allows observations from different cells to inform common components, reducing the number of coefficients to estimate. Country-molecule interactions would allow these differences to vary across medicines, but estimating them precisely requires sufficient residual price variation within the relevant groups, which is not always guaranteed. Of course, the trade-off is that with additivity we may miss relevant heterogeneity. Section 4.1 explores richer specifications of elasticity heterogeneity.

⁹We estimate these components using individual product-quarter observations. Observations on different products help estimate

number of coefficients manageable and allows observations from multiple cells to contribute to estimating each component. Furthermore, since including molecule indicators in K_{igct} may conceal differences in price responsiveness between products containing the same molecule, we present, in Section 4.1 a separate benchmark in which elasticities can differ between branded and generic products (containing the same molecule) within ATC-4 classes.

Interacting own and rival prices with K_{igct} gives our empirical price-response dictionary, i.e. the $2d_K$ -vector of technical treatments

$$D_{igct} = \begin{pmatrix} p_{igct} K_{igct} \\ \bar{p}_{igct} K_{igct} \end{pmatrix}, \quad b_0 = \begin{pmatrix} b_0^{\text{own}} \\ b_0^{\text{rival}} \end{pmatrix}.$$

The quasi-linear demand equation is

$$y_{igct} = D'_{igct} b_0 + g_0(\mathcal{W}_{igct}) + u_{igct}, \quad (1)$$

where $\mathcal{W}_{igct}$ collects the observed characteristics and history defined below, and $g_0(\mathcal{W}_{igct})$ allows demand to depend flexibly on these variables in addition to own and rival prices. Under this specification, the own- and rival-price elasticities for each country–molecule–year cell are given by the following expressions, where $a = a(t)$ denotes the calendar year containing quarter t :

$$\epsilon_{gca(t)}^{\text{own}} = K'_{igct} b_0^{\text{own}}, \quad \epsilon_{gca(t)}^{\text{rival}} = K'_{igct} b_0^{\text{rival}}. \quad (2)$$

From now on, we refrain from explicitly stating the dependence of a on time in order not to further complicate the notation¹⁰. These elasticities measure the percentage response of a product’s quarterly quantity sold to a 1% increase in its own price or the rival-price index, respectively, holding the other price variable and the controls fixed. For a given country and molecule, the model allows these responses to change from one calendar year to the next. Here, $\epsilon_{gca}^{\text{rival}}$ is the elasticity with respect to the average price of same-molecule rivals. Each dimension of K_{igct} has a distinct economic role. Country terms allow price responsiveness to differ across national environments characterized by differences in insurance coverage, reimbursement arrangements, procurement systems, prescribing and substitution rules, and bargaining institutions. Molecule terms allow responsiveness to differ across medicines with different clinical uses, therapeutic alternatives, and branded–generic competitive conditions (Chaudhuri et al., 2006; Dubois and Lasio, 2018; Gugler et al., 2026). Considering heterogeneity at the molecule level, moreover, reflects the view that products containing the same active ingredient face their closest substitution margin within the molecule, while operating within a common country–year institutional environment. Finally, year terms allow the demand response to a given percentage price change to vary from one calendar year to the next. Together, these terms allow us to compare price responsiveness for the same molecule across countries and for different molecules within a country, while allowing responses to change over time.

Identifying the elasticities in Equation (2) requires separating the effects of own- and rival-price changes on sales. These prices, indeed, may move together even after accounting for the controls in $\mathcal{W}_{igct}$, which constitutes a major problem in the empirical reduced form literature (Chaudhuri et al., 2006; Yeung et al., 2018). If both prices rise, for example, the change in sales may reflect both changes. Including only the own price

the same coefficients, providing more information for each estimate. A richer specification could allow elasticities to differ across individual products, but would require sufficient price variation to estimate the additional coefficients. An agreement with the data provider, moreover, explicitly forbids reporting elasticities at the product level for privacy reasons.

¹⁰As a clarification on the notation, although K_{igct} is recorded for each product–quarter observation, its time indicators refer to the calendar year $a = a(t)$ containing quarter t . Its value, therefore, depends only on the country, molecule and year, which explains why the resulting elasticities are indexed by $g, c, a(t)$

could therefore mistakenly attribute the effect of rival prices to the own price coefficient. Accordingly, we include both prices and estimate their effects jointly, so that the own-price elasticity measures the response to own price holding the rival-price index fixed. This serves to account for substitution between products, as emphasized in the pharmaceutical-demand literature (Chaudhuri et al., 2006; Yeung et al., 2018). Distinguishing the two effects also requires sufficient variation in each price beyond that explained by the other price and the controls, both of which are standard assumptions in demand estimation.

A further difficulty for identification is that prices themselves respond to demand conditions. An unobserved increase in demand, for example, may raise both prices and sales, obscuring the effect of a price increase on quantity demanded. Thus, to interpret their coefficients as demand responses, we assume that the unobserved demand shock has mean zero conditional on current own and rival prices, T_{igct} , and the observed state and history, $\mathcal{W}_{igct}$. This is the sequential conditional exogeneity assumption stated below:

$$E[u_{igct} \mid T_{igct}, \mathcal{W}_{igct}] = 0. \quad (3)$$

In other words, after accounting for the observed characteristics and history, higher (lower) current own and rival prices must not systematically coincide with unobserved increases (decreases) in demand. This is the same type of identifying restriction used by Semenova et al. (2023). Section 3.2.1 presents a complementary pooled IV exercise that allows residual price endogeneity explicitly under instrument validity and relevance assumptions.

Under the identifying assumption in Equation (3), we use DML to estimate the elasticities from the variation in log quantity and both log price variables remaining after subtracting their components predicted by $\mathcal{W}_{igct}$. DML allows these prediction functions (the nuisance functions) to involve many candidate controls and flexible functional forms. This is useful when many characteristics may affect prices and demand, but we do not know in advance which are most useful for prediction or how their effects vary. Within the chosen set of candidate controls, the data can guide these choices. Using the same conditioning set for all three variables, we define the corresponding conditional mean functions:

$$\ell_0(\mathcal{W}) = E[y \mid \mathcal{W}], \quad m_0(\mathcal{W}) = E[T \mid \mathcal{W}], \quad m_0(\mathcal{W}) = \begin{pmatrix} m_0^p(\mathcal{W}) \\ m_0^{p^-}(\mathcal{W}) \end{pmatrix}.$$

More specifically, each prediction combines a component that is constant over time for each product–molecule–country unit (i, g, c) with a component that depends on the time-varying controls, denoted by X_{igct} . The constant components are the fixed effects a_{igc}^y , a_{igc}^p , and $a_{igc}^{p^-}$. We allow them to differ across the equations predicting log quantity, log own price, and the rival-price index, respectively, and assume that these fixed effects enter additively, so the nuisance functions take the form

$$\ell_0(\mathcal{W}_{igct}) = a_{igc}^y + h_0^y(X_{igct}), \quad m_0^p(\mathcal{W}_{igct}) = a_{igc}^p + h_0^p(X_{igct}),$$

and

$$m_0^{p^-}(\mathcal{W}_{igct}) = a_{igc}^{p^-} + h_0^{p^-}(X_{igct}).$$

The fixed effects a_{igc}^y , a_{igc}^p , and $a_{igc}^{p^-}$ capture persistent differences across product–molecule–country units in log quantity, log own price, and the rival-price index, respectively ¹¹

Define the population residuals $y_{igct}^\circ = y_{igct} - \ell_0(\mathcal{W}_{igct})$, $T_{igct}^\circ = T_{igct} - m_0(\mathcal{W}_{igct}) = \begin{pmatrix} p_{igct}^\circ \\ p_{igct}^{\circ,-} \end{pmatrix}$, with $p_{igct}^\circ =$

¹¹Notice that they differ from the terms in K_{igct} .

$p_{igct} - m_0^p(\mathcal{W}_{igct})$, $p_{igct}^{-,\circ} = p_{igct}^- - m_0^{p^-}(\mathcal{W}_{igct})$, and where the superscript $\circ$ denotes residualized quantities. Notice now that K_{igct} consists of a constant and country, molecule, and calendar-year indicators, which are also included in $\mathcal{W}_{igct}$, conditioning on $\mathcal{W}_{igct}$ fixes the values of all the indicators in K_{igct} , i.e., formally, K_{igct} is measurable with respect to $\mathcal{W}_{igct}$. We can, therefore, remove the predictable component of each price first and then multiply the remaining price variation by K_{igct} . This gives the same result as removing the predictable component of each original price interaction¹². We therefore define the residualized treatment vector V_{igct} by removing the component of D_{igct} predicted by $\mathcal{W}_{igct}$, which, as explained above, is equivalent to multiplying the own- and rival-price residuals by K_{igct} :

$$V_{igct} := D_{igct} - E[D_{igct} | \mathcal{W}_{igct}] = \begin{pmatrix} K_{igct} p_{igct}^\circ \\ K_{igct} p_{igct}^{-,\circ} \end{pmatrix}.$$

Take the conditional expectation of Equation (1) given $\mathcal{W}_{igct}$. Under the exogeneity assumption, $E[u_{igct} | \mathcal{W}_{igct}] = 0$, so

$$\begin{aligned} \ell_0(\mathcal{W}_{igct}) &= E[y_{igct} | \mathcal{W}_{igct}] \\ &= E[D_{igct} | \mathcal{W}_{igct}]' b_0 + g_0(\mathcal{W}_{igct}) + \underbrace{E[u_{igct} | \mathcal{W}_{igct}]}_{=0 \text{ under exogeneity}}, \end{aligned}$$

which, as mentioned, we model as $a_{igc}^y + h_0^y(X_{igct})$ and where $g_0(\mathcal{W}_{igct})$ remains unchanged because it is already determined by the conditioning variables. Thus, subtracting this expression from the original demand equation removes $g_0(\mathcal{W}_{igct})$. The outcome and price interactions are replaced by their residuals, but the coefficients b_0 remain unchanged. Moreover, because V_{igct} is constructed from T_{igct} and $\mathcal{W}_{igct}$, the same exogeneity assumption implies $E[V_{igct} u_{igct}] = 0$. The residualized demand equation and its associated moment condition are therefore

$$y_{igct}^\circ = V_{igct}' b_0 + u_{igct}, \quad E[V_{igct} (y_{igct}^\circ - V_{igct}' b_0)] = 0, \quad (4)$$

which is the estimating equation. DML, therefore, estimates the nuisance functions $\ell_0(\mathcal{W})$ and $m_0(\mathcal{W})$ to remove the predictable components of log sales and prices. The price-response coefficients are then estimated from the residualized equation. This follows the partialling-out logic of the Frisch–Waugh–Lovell theorem, while allowing flexible, cross-fitted prediction models and a high-dimensional second-stage estimator. Since the nuisance functions are estimated, they are subject to estimation error. DML is designed to limit how these errors affect the estimated elasticities. Under the identification and regularity conditions, the leading effects of small nuisance-estimation errors cancel in the population moment condition (Neyman orthogonality). The remaining effects involve products of prediction errors, so they shrink faster as the predictions improve. Cross-fitting provides an additional safeguard, i.e. each observation’s residuals are calculated using models trained on other observations. This reduces the risk that fitting noise in the training data distorts the estimated price responses (Chernozhukov et al., 2018; Semenova et al., 2023). Together, these features support flexible prediction models for the nuisance parameters, allowing residualization of high-dimensional covariates while preserving the identification conditions for elasticity. Inference is discussed in Appendix B. Finally, in panels with many products but relatively short histories, each product’s fixed effect is estimated from limited information, and errors in these estimates can bias the estimated elasticities. As Bonhomme

¹²In other words, because K_{igct} is fixed when conditioning on $\mathcal{W}_{igct}$, it can be taken outside the conditional expectation. Hence, $K_{igct} p_{igct} - E[K_{igct} p_{igct} | \mathcal{W}_{igct}] = K_{igct} (p_{igct} - E[p_{igct} | \mathcal{W}_{igct}]) = K_{igct} p_{igct}^\circ$, and $K_{igct} p_{igct}^- - E[K_{igct} p_{igct}^- | \mathcal{W}_{igct}] = K_{igct} (p_{igct}^- - E[p_{igct}^- | \mathcal{W}_{igct}]) = K_{igct} p_{igct}^{-,\circ}$. Thus, residualizing each price interaction gives the same result as multiplying the corresponding price residual by K_{igct} .

et al. (2026) show, first-order Neyman orthogonality alone does not generally eliminate this incidental-parameter problem. The framework of Semenova et al. (2023) addresses this difficulty through additional structure. In their setup, observed characteristics account for persistent differences across products, reducing the need to estimate each product’s effect entirely separately. The remaining product-specific deviations are assumed to be mostly small, with relatively few substantial exceptions, and are estimated using regularization. Valid inference still requires sufficiently accurate first-stage predictions, including the contributions of these unit effects. We rely on the same structural assumptions of Semenova et al. (2023).

2.2 DML and Logit-Based Structural Demand

We next relate the heterogeneous price-response object from the previous section to the logit-based structural demand models commonly used in empirical IO to assess own and cross-elasticities. The comparison serves two purposes. First, it shows conditions under which the price-response component – and, under stronger regularity conditions, the elasticity function – generated by a fixed and identified logit or RCL models can be approximated using a sufficiently rich technical dictionary of the same type as D_{iget} in the DML. Second, it clarifies some interesting trade-offs between our DML specification and RCL, both in the assumptions they require and in how much detail they provide about substitution between medicines.

For the sake of aligning with the terminology of RCL, let $m = (c, g, t)$ denote a market, where c is country, g is a molecule, and t is quarter. Let $j \in \mathcal{J}_m$ index products and ι index consumers. Let p_{jm} denote log price and let p_m collect the within-market vector of log prices. Let x_{jm} collect the observed characteristics of product j , and let X_m^{obs} collect observed market conditions other than prices, including the characteristics of all available products, $(x_{\ell m})_{\ell \in \mathcal{J}_m}$. For this structural comparison, we also include potential market size and unobserved demand factors, such that the structural state (without prices) is $X_m = (X_m^{\text{obs}}, M_m, \xi_m)$, where M_m is potential market size and $\xi_m = (\xi_{\ell m})_{\ell \in \mathcal{J}_m}$ collects unobserved factors affecting demand for each product in that market. Let H_ι collect consumer demographics and let v_ι represent unobserved differences in consumer tastes. Briefly, the utility of consumer ι for product j in market m is represented by

$$u_{\iota jm} = \delta_{jm}(x_{jm}, p_m, \xi_{jm}; \theta_1) + \mu(x_{jm}, p_m, H_\iota; v_\iota; \theta_2) + \varepsilon_{\iota jm},$$

where the conventional RCL parameterization writes $\theta = (\theta'_1, \theta'_2)'$, with $\theta_1 = (\alpha, \beta')'$ collecting the linear mean-utility coefficients and θ_2 stacking the free elements of (Π, Σ) , which govern random-coefficient heterogeneity (Nevo, 2000). In the linear mean-utility block used below, $\delta_{jm} = x'_{jm}\beta - \alpha p_{jm} + \xi_{jm}$. For brevity, write $\mu_{\iota jm} = \mu(x_{jm}, p_m, H_\iota; v_\iota; \theta_2)$. It can be shown (Nevo, 2000) that the market share of product j is

$$s_{jm}(p_m, X_m; \theta) = \int \frac{\exp(\delta_{jm} + \mu_{\iota jm})}{1 + \sum_{\ell \in \mathcal{J}_m} \exp(\delta_{\ell m} + \mu_{\iota \ell m})} dF(H_\iota, v_\iota).$$

Here $\varepsilon_{\iota jm}$ is an idiosyncratic Type-I extreme-value utility shock. Quantities satisfy $q_{jm} = M_m s_{jm}(p_m, X_m; \theta)$, so the log-quantity outcome is

$$Y_{jm} := \log q_{jm} = \log M_m + \log s_{jm}(p_m, X_m; \theta). \quad (5)$$

Thus, the model implies the conditional demand function

$$m_j(X_m, p_m) := E[Y_{jm} | X_m, p_m],$$

and the RCL elasticities are

$$\varepsilon_{jk,m}^{\text{RCL}} = \frac{\partial \log s_{jm}(p_m, X_m; \theta)}{\partial \log p_{km}} = \frac{\partial \log s_{jm}(p_m, X_m; \theta)}{\partial p_{km}},$$

where $P_{km} = \exp(p_{km})$ is the price level. The two derivatives coincide because $p_{km} = \log P_{km}$.¹³

While RCL derives this dependence on prices from the choice probabilities above, our DML demand equation represents it through price terms and interactions. The question is how closely such a representation can approximate the price-dependent part of the structural demand function. For the result below, let $D_j(X_m, p_m)$ denote a general collection of price terms, including nonlinear transformations and interactions with the non-price variables in X_m . The price interactions used in Section 2.1 are a particular version of this general dictionary. For a fixed logit-based structural model, we can write

$$m_j^{\text{LB}}(X_m, p_m) = \log M_m + \log s_{jm}^{\text{LB}}(p_m, X_m; \theta) = a_j(X_m) + D_j(X_m, p_m)' \gamma_D + \tau_{jm}, \quad (6)$$

where LB denotes either simple logit or RCL, and $a_j(X_m)$ collects baseline terms that do not depend on current prices, including $\log M_m$. The remainder τ_{jm} is defined as whatever the approximation $a_j(X_m) + D_j(X_m, p_m)' \gamma_D$ does not capture, so the equality holds by construction.

The proposition below establishes how accurately this representation can reproduce structural demand. Under its assumptions, a finite collection of polynomial terms and suitable coefficients can be chosen to make $|\tau_{jm}|$ smaller than any prescribed positive tolerance throughout the specified range of prices and market conditions. These terms include powers and interactions, and greater accuracy may require more terms. Thus, a specification that is linear in its coefficients can still approximate the nonlinear dependence on prices generated by a structural demand model.

Proposition 2.1 (Dictionary approximation of logit-based demand). *Let $(X_m, p_m) \in \mathcal{S}_{\text{state}} \subset \mathbb{R}^{d_{\text{state}}}$ denote the relevant market state, where d_{state} is the dimension of that state, and suppose that $\mathcal{S}_{\text{state}}$ is compact. Let*

$$m_j^{\text{LB}}(X_m, p_m) = \log M_m + \log s_{jm}^{\text{LB}}(p_m, X_m; \theta)$$

denote the logit-based implied conditional log-quantity function. For brevity, $(\cdot)$ denotes the previously specified market-variable arguments of a function, with the model parameters held fixed. Assume that $s_{jm}^{\text{LB}}(\cdot)$ is continuous and bounded away from zero on $\mathcal{S}_{\text{state}}$. Define $a_j(X_m)$ to collect baseline demand components, including $\log M_m$ and other non-price demand shifters, and let

$$\tau_j(X_m, p_m) := m_j^{\text{LB}}(X_m, p_m) - a_j(X_m)$$

denote the remaining price-response component. If $\tau_j(\cdot)$ is continuous on $\mathcal{S}_{\text{state}}$, then, for every $\varepsilon > 0$, there exists a finite technical dictionary $D_j(X_m, p_m)$ and a coefficient vector γ_D such that

$$\sup_{(X_m, p_m) \in \mathcal{S}_{\text{state}}} |\tau_j(X_m, p_m) - D_j(X_m, p_m)' \gamma_D| < \varepsilon.$$

Equivalently,

$$m_j^{\text{LB}}(X_m, p_m) = a_j(X_m) + D_j(X_m, p_m)' \gamma_D + \tau_{jm}, \quad \sup_{(X_m, p_m) \in \mathcal{S}_{\text{state}}} |\tau_{jm}| < \varepsilon.$$

The proof is reported in Appendix E. The same argument is valid for both simple-logit and RCL shares, as shown in Proposition E.1 for the latter. A simulation exercise is reported in Appendix G.1.

In this sense, a fixed logit-based model of the type specified in Equation (6) provides a useful benchmark because its implied log quantity can be separated into a baseline component $a_j(X_m)$ and a price-dependent component $\tau_j(X_m, p_m)$. The baseline includes the market-size term $\log M_m$, product and market

¹³When differentiating with respect to the price level instead, the chain rule gives $\frac{\partial \log s_{jm}}{\partial p_{km}} = P_{km} \frac{\partial \log s_{jm}}{\partial P_{km}} = \frac{P_{km}}{s_{jm}} \frac{\partial s_{jm}}{\partial P_{km}}$. Holding potential market size fixed also makes share and quantity elasticities equal.

fixed effects, and other determinants that do not depend on current prices. Proposition 2.1 shows that, when the remaining component is continuous, it can be uniformly approximated on the specified compact domain by $D_j(X_m, p_m)' \gamma_D$. These terms allow nonlinear price effects while remaining linear in their coefficients. For elasticity comparisons, however, matching log quantities is not enough because, trivially, two demand curves can be close together but have different slopes. As a consequence, for the elasticity comparison, Appendix E, therefore, requires that increasingly rich dictionary approximations can reproduce both log demand and its log-price derivatives¹⁴. Under the latter assumption – since holding X_m fixed, the baseline $a_j(X_m)$ contributes no price derivative – the RCL own-price elasticity can be approximated by $\varepsilon_{jj,m}^{\text{RCL}} \approx \frac{\partial}{\partial p_{jm}} [D_j(X_m, p_m)' \gamma_D]$. In the empirical specification of Section 2.1, we specialize the general technical dictionary to an estimable dictionary based on the observed market state. Let

$$D_j^{\text{emp}}(X_m^{\text{obs}}, p_m) = \begin{pmatrix} p_{jm} K_j(X_m^{\text{obs}}) \\ p_{jm}^- K_j(X_m^{\text{obs}}) \end{pmatrix}, \quad b_0 = \begin{pmatrix} b_0^{\text{own}} \\ b_0^{\text{rival}} \end{pmatrix},$$

where $K_j(X_m^{\text{obs}})$ contains the observed country, molecule, and calendar-year terms. Setting $j = i$ and $m = (c, g, t)$, it equals K_{igct} , with p_{jm}^- being the leave- j -out rival-price index defined above. The fitted price-response component is therefore

$$D_j^{\text{emp}}(X_m^{\text{obs}}, p_m)' b_0 = p_{jm} K_j(X_m^{\text{obs}})' b_0^{\text{own}} + p_{jm}^- K_j(X_m^{\text{obs}})' b_0^{\text{rival}}.$$

Because p_{jm}^- excludes product j and $K_j(X_m^{\text{obs}})$ does not depend on contemporaneous prices, differentiating with respect to the focal product's log price gives the DML-implied own-price elasticity

$$\varepsilon_{j,m}^{\text{DML,own}} = \frac{\partial}{\partial p_{jm}} [D_j^{\text{emp}}(X_m^{\text{obs}}, p_m)' b_0] = K_j(X_m^{\text{obs}})' b_0^{\text{own}}.$$

Analogously, the response to the scalar rival-price index is

$$\varepsilon_{j,m}^{\text{DML,rival}} = \frac{\partial}{\partial p_{jm}^-} [D_j^{\text{emp}}(X_m^{\text{obs}}, p_m)' b_0] = K_j(X_m^{\text{obs}})' b_0^{\text{rival}}.$$

To summarize, the preceding analysis shows how a sufficiently rich dictionary can approximate logit-based demand and, under the additional derivative conditions, its elasticities. As mentioned, our empirical specification uses a simpler dictionary to limit the number of elasticity coefficients estimated from the available price variation, providing sufficient variation to distinguish the effects of additional price interactions. Formally proving which dictionary would make our estimates identical to the ones of RCL goes beyond the scope of the present paper. The practical question is whether the different restrictions imposed by DML and RCL affect the own-price elasticities relevant to our international pricing analysis. Section 4.1 examines this question by comparing DML and RCL estimates. For comparable observations and elasticity summaries, agreement may indicate that similar price responses emerge under different demand specifications, while disagreement would reveal sensitivity to those choices.

The next subsection examines when adding many characteristics to an unrestricted RCL model prevents its parameters from being uniquely determined by the sample moments. This helps explain our choice to accommodate rich demand controls through DML while focusing on the estimation of price-response coefficients.

¹⁴The importance of similar assumptions for controlling derivative approximation when estimating elasticities is also emphasized by Elbadawi et al. (1983).

2.2.1 A high-dimensional boundary result for RCL

Pharmaceutical demand reflects drug characteristics, treatment experience, and insurance arrangements (Crawford and Shum, 2005; Einav et al., 2015). When modeling demand across many markets, as in our case, however, the appropriate set of observed characteristics and interactions, is generally unknown in advance, motivating consideration of a broad set of candidate terms (Chernozhukov et al., 2022). Including many such terms in RCL mean utility introduces additional structural coefficients, potentially exceeding the number of product–market observations or instruments available to distinguish them. Intuitively, for each candidate value of the nonlinear parameters θ_2 , the BLP share inversion recovers mean utilities, and GMM uses the instruments to estimate the linear coefficients $\theta_1 = (\alpha, \beta')'$ (Nevo, 2000). Ambiguities may arise when increasing one coefficient may be compensated by changes in others, leaving every sample moment unchanged, such that GMM assigns the same objective value to these different parameter combinations¹⁵. In such a case, if one minimizes the objective, the others do as well. The estimation, therefore, yields multiple equally good solutions. The reader is referred to Appendix F for a more extensive discussion. The following proposition states the sample rank conditions under which this ambiguity arises in the unrestricted model with high-dimensional sparse covariates. Recall from above that by definition of mean utility, $\xi_{jm}(\theta_1, \theta_2) = \delta_{jm}(\theta_2) - (x'_{jm}\beta - \alpha p_{jm})$. For convenience, we write these equations together for the N product–market observations. Let X contain the k_x characteristics in x_{jm} , with one row per observation, and let p_N collect log prices. Defining $X_1 = (-p_N, X)$ and recalling that $\theta_1 = (\alpha, \beta')'$, we obtain $\xi_N(\theta_1, \theta_2) = \delta_N(\theta_2) - X_1\theta_1$, $\theta_1 \in \mathbb{R}^{k_x+1}$, with ξ_N collecting the demand components left unexplained by the observed characteristics and prices at the candidate parameter values.

Proposition 2.2 (Rank failure in the unrestricted BLP/RCL GMM problem). *Let*

$$\xi_N(\theta_1, \theta_2) = \delta_N(\theta_2) - X_1\theta_1, \quad \theta_1 \in \mathbb{R}^{k_x+1},$$

where k_x is the number of mean-utility characteristics and $X_1 = (-p_N, X)$ is the $N \times (k_x + 1)$ matrix of log prices and those characteristics. Define

$$g_N(\theta_1, \theta_2) = \frac{1}{N}Z'\xi_N(\theta_1, \theta_2), \quad Q_N(\theta_1, \theta_2) = g_N(\theta_1, \theta_2)'W_Ng_N(\theta_1, \theta_2),$$

where Z is the $N \times L_Z$ instrument matrix, L_Z is the number of instruments, and W_N is positive definite. Consider candidate values of θ_2 for which share inversion is defined.

1. For any fixed θ_2 , if $\text{rank}(Z'X_1) < k_x + 1$, the unrestricted linear parameter vector θ_1 is not uniquely determined by the sample moments. This rank deficiency necessarily occurs when $k_x + 1 > N$ or $k_x + 1 > L_Z$.
2. Under the rank deficiency in part 1, the full sample moment Jacobian with respect to $\theta = (\theta_1', \theta_2')'$, wherever it exists, cannot have full column rank.
3. If $\text{rank}(X_1) = N$, then, for every candidate θ_2 , there exists a θ_1 satisfying $X_1\theta_1 = \delta_N(\theta_2)$. Hence $\xi_N(\theta_1, \theta_2) = 0$ and $Q_N(\theta_1, \theta_2) = 0$. The profiled sample objective is therefore zero for every candidate θ_2 , and the sample demand moments cannot uniquely determine the joint structural parameter vector.

¹⁵A two-coefficient example illustrates the ambiguity that can arise when a larger model contains more coefficients than the sample moments can distinguish. For example, suppose X_1 has two linearly independent columns, but there is only one instrument, with $N^{-1}Z'X_1 = (1, 1)$. Writing $\theta_1 = (\theta_{11}, \theta_{12})'$, the sample moment is $g_N(\theta_1, \theta_2) = N^{-1}Z'\delta_N(\theta_2) - (\theta_{11} + \theta_{12})$. For fixed θ_2 , this moment depends only on the sum of the two coefficients. Increasing one coefficient and decreasing the other by the same amount therefore leaves the moment unchanged. For instance, $\theta_1 = (1, 2)'$ and $\theta_1 = (2, 1)'$ both have sum 3, producing identical moments and GMM objective values despite different fitted components $X_1\theta_1$.

The proof is reported in Appendix F where we also propose a possible way to mitigate such an issue in logit-based models, by passing through the regularization of the BLP/RCL GMM problem itself ¹⁶. Importantly for our scope is the fact that parameter nonuniqueness due to rank failures in RCL also affects elasticities when the indistinguishable parameter combinations imply different responses to a price change. Simulation 2 in Appendix G.2 illustrates this case: holding θ_2 fixed, changes in the mean price coefficient α are offset by changes in the characteristic coefficients. The recovered mean utilities and sample moments remain unchanged, but the predicted response to price changes differs. A similar concern is discussed by Crawford and Shum (2005, Section 4.1) for the Italian anti-ulcer market. In their work, they treat drug prices as constant over time and note that, with linear utility, assigning a drug a higher benefit can compensate for assuming that consumers are more sensitive to its price. These combinations explain the observed purchases equally well ¹⁷. If the price increased while the drug’s benefit remained unchanged, however, the more price-sensitive specification would predict a larger fall in demand.¹⁸

Our DML specification handles high-dimensional sparsity in demand determinants by exploiting regularization and the adoption of nuisance functions (Chernozhukov et al., 2022). As seen, observed characteristics and history enter $\mathcal{W}_{igct}$, including product characteristics that could also enter RCL mean utility. Their contribution to demand at given prices is represented through $g_0(\mathcal{W}_{igct})$ and any associated differences in price responsiveness must be captured by the price interactions with K_{igct} . The idea is that – under approximate sparsity – a relatively small set of terms can adequately describe the conditional means $\ell_0(\mathcal{W}_{igct})$ and $m_0(\mathcal{W}_{igct})$, even when many candidate controls are available and their relevance is unknown beforehand. Regularization helps estimate these functions without requiring every coefficient in an unrestricted representation to be separately identified, making the presence of different coefficient combinations harmless for this step if they produce the same predictions. Orthogonality then reduces sensitivity to errors in the estimated conditional means under the stated conditions (Semenova et al., 2023).

2.2.2 Implications for the empirical model

The two results in this section inform our use of DML for the main elasticity analysis. Proposition 2.1 shows that a sufficiently rich technical dictionary can approximate the price-dependent component of log quantity generated by a fixed, identified logit or RCL model under the stated assumptions. With the additional smoothness and derivative-approximation conditions in Appendix E, this representation can also approximate the model’s implied elasticities. Notice that a richer dictionary can reduce approximation error by including additional interactions or nonlinear terms, but estimating these additional coefficients require sufficient residual price variation and may reduce precision. We therefore use additive country, molecule, and year components to share information across observations and limit the number of coefficients to estimate. Proposition 2.1 suggests that this setting may not only restrict the heterogeneity we can capture, but also affect how closely our elasticity estimates align with those obtained from an RCL model. Whether these differences in specification translate into different elasticity estimates is examined in Section 4.1.

¹⁶This requires assumptions on structural sparsity, valid and sufficiently strong instruments, an appropriate regularization scheme, and restricted identification of the GMM moments (Fan and Liao, 2014; Sawada and Kawaguchi, 2020). Fully developing and validating such an estimator is beyond the scope of this paper and is left for future research. Appendix F.1 describes an illustrative regularized profiling approach.

¹⁷To see this, consider an anti-ulcer medicine observed at the same price throughout the sample. Its purchases could reflect substantial benefits from symptom relief that compensate for strong price sensitivity. Alternatively, patients might attach less value to those benefits but also respond less to price. With suitable parameter values, both explanations predict the same purchases at the observed price.

¹⁸Their implemented model addresses this difficulty through nonlinear utility, using variation in patients’ experience and perceived uncertainty.

On the other hand, pharmaceutical demand differs across countries and therapeutic categories (Chaudhuri et al., 2006) which motivates the use of a broad set of candidate controls in demand estimation, allowing uncertainty about their predictive importance (Belloni et al., 2014; Chernozhukov et al., 2022). However, Proposition 2.2 has shown that including them as unrestricted RCL coefficients can exceed what the sample moments can distinguish.

These considerations, together, motivate the use of DML as the main estimator to accommodate rich controls through regularized prediction while imposing a manageable structure on elasticity heterogeneity, and RCL as a benchmark for robustness checks.

3 Empirical Specification

3.1 Data and Sample

We use IQVIA MIDAS data to follow quarterly pharmaceutical sales across 171 molecules and 63 countries from 2008Q4 through 2020Q3. The panel comprises 1,968,845 observations over these 48 consecutive quarters. Each observation records sales of a molecule under a given local and international product label, supplied by a given corporation in a particular country and quarter. The data also identify the active molecule, the product’s ATC-4 therapeutic classification, along with other covariates. To track products over time, we combine product labels with corporation identifiers. The 47,589 distinct labels yield 63,878 product-tracking keys, which we follow separately within each country and molecule. This defines the 79,565 country–molecule–product (CMP) units in our panel.

For each observed product–country–quarter cell, IQVIA reports ex-manufacturer sales in constant-exchange U.S. dollars and quantities sold in standard units. Let S_{igct} denote sales, Q_{igct} standard units and P_{igct} ex-manufacturer prices for product-tracking key i , molecule g , country c , and quarter t . We construct $p_{igct} = \log P_{igct}$, $y_{igct} = \log Q_{igct}$.

Products are not observed in every country and quarter, so the panel is unbalanced. We retain their observed sales episodes without requiring continuous presence throughout the study period, which would select only long-lived products. Table 1 summarizes the data coverage and the additional requirements for the baseline DML sample, described in Section 2.

Table 1. Data coverage and baseline DML sample

Panel A: Prepared panel	
Observations	1,968,845
Sample period	2008Q4–2020Q3
Quarters	48
IQVIA geographic markets	63
Molecules	171
Distinct local product labels	47,589
Corporation–local-product tracking keys	63,878
Country–molecule–product panel units	79,565
ATC-4 classes	285
Panel B: Baseline DML sample	
Domestic rival-price support	1,853,974
Focal foreign-price-predictor support	579,111
Domestic-rival and focal-foreign support	509,787
Complete $t - 1, \dots, t - 4$ histories and estimation variables	429,675
Estimation period	2009Q1–2020Q3
Calendar quarters represented	44
CMP units in the estimation sample	17,117
Product-tracking keys in the estimation sample	5,540

Notes: Panel A describes the prepared panel. A product-tracking key combines the corporation and local product label recorded by IQVIA; combining this key with the country and molecule defines a country–molecule–product (CMP) panel unit. An observation has domestic-rival-price support when at least one other product containing the same molecule is observed in the same country and quarter, allowing the leave-own-out domestic rival-price index to be constructed. It has exact foreign-key support when the focal product’s corporation–local-product tracking key is observed in at least one other country in the same quarter, allowing its leave-country-out foreign price to be constructed. This foreign-key condition applies only to the focal product; it does not require every domestic rival to be observed abroad. These two support counts are calculated separately, while “both support requirements” reports their intersection. The baseline DML sample additionally requires non-missing log price and log quantity in each of the four preceding quarters, $t - 1, \dots, t - 4$, and complete values for all other estimation variables. The 44 quarters reported in Panel B are the calendar quarters represented in the baseline DML sample, not the number of lags used.

The demand estimates use a subset of the descriptive panel. To construct the price predictors, we require another product containing the same molecule to be observed in the same country and quarter, and the focal product’s corporation and local product label to appear in another country in that quarter. We also require price and quantity observations for the four immediately preceding quarters and complete values for the other estimation variables. These requirements, together with the cut of 2008 from the panel (to avoid incomplete years) leave 429,675 observations. We do not observe any evident sample selection after the mentioned operations.

For the MFN application in Section 5.2, we supplement IQVIA with OECD membership dates and GDP per capita in current PPP-adjusted U.S. dollars to identify eligible comparator countries. We also use the IMF producer-price index (2010 = 100) and OECD health expenditure per person in PPP-adjusted U.S. dollars at constant 2020 prices to account for differences in general price conditions and health-system resources.¹⁹

¹⁹OECD membership dates are taken from <https://www.oecd.org/en/about/legal.html>. GDP per capita is from

We match these annual variables by country and calendar year over 2009–2020. After applying the eligibility and support requirements in Section 5.2, the historical comparison contains 116 U.S. molecule–period observations covering 17 molecules.

3.2 DML Identification and Estimation

We apply the model of Section 2.1 to the IQVIA panel. With 63 countries, 171 molecules, and 12 calendar years, the constant and the indicators in K_{igct} give $d_K = 244$ terms. As explained, we allow elasticities to vary across countries, molecules, and years, via the indicators in K_{igct} with the molecule dimension reflecting the close substitution between products containing the same active ingredient (Dubois and Lasio, 2018; Dubois et al., 2022)²⁰. The interactions of K_{igct} with own price and the rival-price index therefore produce 244 coefficients for each response, or $d = 488$ in total.

Following Section 2.1, we first estimate the nuisance functions $\ell_0(\mathcal{W}_{igct})$, $m_0^p(\mathcal{W}_{igct})$, and $m_0^{p^-}(\mathcal{W}_{igct})$, which predict log quantity, log own price, and the rival-price index, respectively. Subtracting their predictions gives the residualized quantities $\hat{y}_{igct}^\circ$, $\hat{p}_{igct}^\circ$, and $\hat{p}_{igct}^{\circ,-}$. We then interact the two price residuals with K_{igct} to construct $\hat{V}_{igct}$ and estimate the price-response coefficients from the residualized demand equation.

The information set used for the *predictions* is $\mathcal{W}_{igct}$ combines panel identity, country, molecule, ATC-3 class, product type, calendar time and other covariates with four quarterly lags of own price and standard-unit sales and the focal product’s leave-country-out price. The latter is calculated from observations with the same corporation–local-product tracking key in other countries during the same quarter. Within the prediction models, country and molecule interactions allow the relationships with past prices and sales to differ across markets and medicines via interactions of lagged prices and quantities with country and molecule indicators. Quarter-of-year indicators capture seasonal patterns, while common and country-specific linear and quadratic trends capture changes over time.

The possibility of using a rich set of predictors used in the nuisance functions above is accommodated through DML’s extension of the partialling-out logic of linear regression. As shown, DML enables us to separate traditional regression models into two stages, the first of which removes the predicted components of quantities and prices before estimating the price responses (Chernozhukov et al., 2018; Semenova et al., 2023). Since, crucially, predictive accuracy – rather than causal interpretation – guides the first-stage fits, we are able to account for many covariates through flexible prediction models, without requiring a causal interpretation of their individual coefficients.

The second stage, instead, concerns the causal interpretation of the elasticities, and requires identification guarantees like the sequential exogeneity in Equation (3).

In practice, the DML procedure starts by predicting log quantity, log own price, and the rival-price index using separate elastic-net models. We then divide calendar time into contiguous blocks shared across panel units and hold out each block in turn. The training sample excludes both the target block and its immediate neighbors, reducing dependence between training and held-out observations. This is the neighbors-left-out cross-fitting procedure of Semenova et al. (2023). Within each training sample, cross-validation compares predictive performance on internal validation folds to select the elastic-net tuning parameters. We then fit

OECD dataflow OECD.SDD.NAD:DSD_NAMAIN10@DF_TABLE1_EXPENDITURE_HCPC(2.0); the producer-price index is from IMF.STA:PPI(3.0.0); and health expenditure is from OECD dataflow OECD.ELS.HD:DSD_SHA@DF_SHA. The IMF extract lacks producer-price observations for Canada and Japan in 2019–2020, which we do not impute. The archived OECD health-expenditure file does not retain its precise retrieval date or dataflow vintage.

²⁰The baseline preserves differences between molecules that broader therapeutic classes may combine, while Section 4.1 provides a complementary benchmark using ATC-4 classes.

the selected models on that training sample and predict the observations in the target block. Subtracting these predictions from observed quantities and prices gives the cross-fitted residuals. The target observations therefore enter neither model fitting nor tuning, limiting overfitting bias in the subsequent estimation of price responses.

Once the nuisance functions have been computed via prediction, we use them to remove the components of quantities and prices explained by the controls. Namely, for observation $r = 1, \dots, n$ – with r serving as shorthand for (i, g, c, t) – the resulting residuals are

$$\begin{aligned}\hat{y}_r^\circ &= y_r - \hat{\ell}^{(-k(r))}(\mathcal{W}_r), \\ \hat{p}_r^\circ &= p_r - \hat{m}^{p, (-k(r))}(\mathcal{W}_r), \\ \hat{p}_r^{\circ, -} &= p_r^- - \hat{m}^{p^-, (-k(r))}(\mathcal{W}_r),\end{aligned}$$

where $k(r)$ identifies the observation’s time block, and $(-k(r))$ indicates that this block and its neighbors are excluded from training.

In the second stage, we estimate how residual quantity varies with residual own and rival prices, retaining the country, molecule, and year structure specified in K_r . We estimate the residualized demand equation by LASSO, obtaining $\hat{b}^L$. These estimates are subsequently adjusted to reduce the bias introduced by shrinkage as described below.

For an observation in cell (g, c, a) , the fitted second-stage equation can therefore be written as

$$\hat{y}_r^\circ = \hat{\alpha}^L + \hat{\epsilon}_{gca}^{\text{own}, L} \hat{p}_r^\circ + \hat{\epsilon}_{gca}^{\text{rival}, L} \hat{p}_r^{\circ, -} + \hat{u}_r^L,$$

where $\hat{\alpha}^L$ is an unpenalized intercept, $\hat{u}_r^L$ is the fitted regression residual, and $\hat{\epsilon}_{gca}^{\text{own}, L} = K_r' \hat{b}^{\text{own}, L}$, $\hat{\epsilon}_{gca}^{\text{rival}, L} = K_r' \hat{b}^{\text{rival}, L}$, and the latter equation is the sample version of Equation (4).

Regularization plays different roles in the two estimation stages. First-stage elastic net regularizes the coefficients used to predict quantities and prices from the controls. Second-stage LASSO instead penalizes the 488 price-response coefficients in b , whose estimates determine the elasticities. The other controls do not enter this regression as separate covariates since they have already been accounted for through the predictions used to construct the residuals.

By shrinking coefficients toward zero, LASSO can stabilize estimation when many components contribute little to the price-response surface, as assumed under approximate sparsity. However, this shrinkage can also bias the estimated elasticities. Since the elasticities in the second stage are our parameters of interest under the maintained identifying assumption, predictive performance alone is insufficient, and causal guarantees are needed. Accordingly, the next paragraph explains how we reduce shrinkage bias, while Appendix B provides details on inference with regularization. In particular, to reduce shrinkage bias, we apply a ridge-adjusted one-step correction, which we call adjusted debiased LASSO (ADL). The correction revises the coefficients using the remaining association between the price regressors and quantity variation unexplained by the second-stage fit, accounting for correlations among regressors. Ridge regularization limits the amplification of noise in this adjustment. Appendix B explains the procedure and its relationship to CLIME, used by Semenova et al. (2023). Valid normal-based confidence intervals require the remaining bias and nuisance-estimation errors to be negligible relative to sampling uncertainty, alongside appropriate dependence conditions and consistent variance estimation.

ADL is our principal regularized specification. We also report the LASSO estimates and an unpenalized OLS benchmark that retains all 488 price-response coefficients.

3.2.1 Alternative identification with foreign-price shifters

We examine sensitivity to the sequential conditional exogeneity assumption underlying our baseline DML specification (Semenova et al., 2023) by using foreign prices as instruments for domestic own and rival prices. The own-price instrument is the logarithm of the arithmetic mean price in levels of the same molecule–corporation–local-product key in other countries during the same quarter (Dubois et al., 2021). For the rival-price instrument, we replace each domestic rival’s price with that product’s average foreign price in levels, then construct the log rival-price index using the same aggregation rule as for domestic prices. As usual, for correct identification – conditional on the controls – the instruments must be unrelated to unobserved domestic demand shocks, and affect domestic sales only through domestic prices. Additionally, they should jointly provide variation that distinguishes the two domestic price effects.

For this comparison, we reestimate DML as a pooled partially linear regression (PLR) and compare it with 2SLS that instruments both prices jointly as described in Appendix C. Both methods use the same use individual product–quarter observation, controls $\mathcal{W}^{IV}$, and cross-fitting folds. The comparison controls exclude both foreign-price instruments, whereas the baseline conditioning set $\mathcal{W}$ includes the focal product’s foreign price. In this comparison, we estimate just two elasticities: one for own price and one for the rival-price index, each assumed to be the same across countries, molecules, and years. Allowing these elasticities to differ in the IV case would, indeed, require enough instrument-predicted price variation to estimate those differences as well (Sanderson and Windmeijer, 2016). Since in our IV sample, even a country-group comparison leaves the own-price responsiveness differences imprecisely estimated ²¹, we retain two common price responses for this sensitivity exercise.

The pooled own-price coefficients are -0.537 under DML and -0.587 under IV (Appendix Table C.2) and point estimates remain robust with additional fixed effects. The two approaches therefore give a similar picture of own-price responsiveness. As a note of caution we highlight that the IV confidence sets remain wide when we turn to Anderson–Rubin confidence sets for weak instruments (Anderson and Rubin, 1949).

4 Results: Estimated Price Elasticities of Demand

In this section, we present the own-price and rival-price index elasticities estimated using DML with ADL in the second stage. Section 4.1 provides benchmark estimates with RCL.

Figure 1 shows that the vast majority of the reported country–molecule cells have negative own-price elasticities, with their magnitude varying considerably. Demand is more price responsive for some molecules than others within the same country, and the same molecule can face markedly different responses across national markets.

²¹Results available upon request.

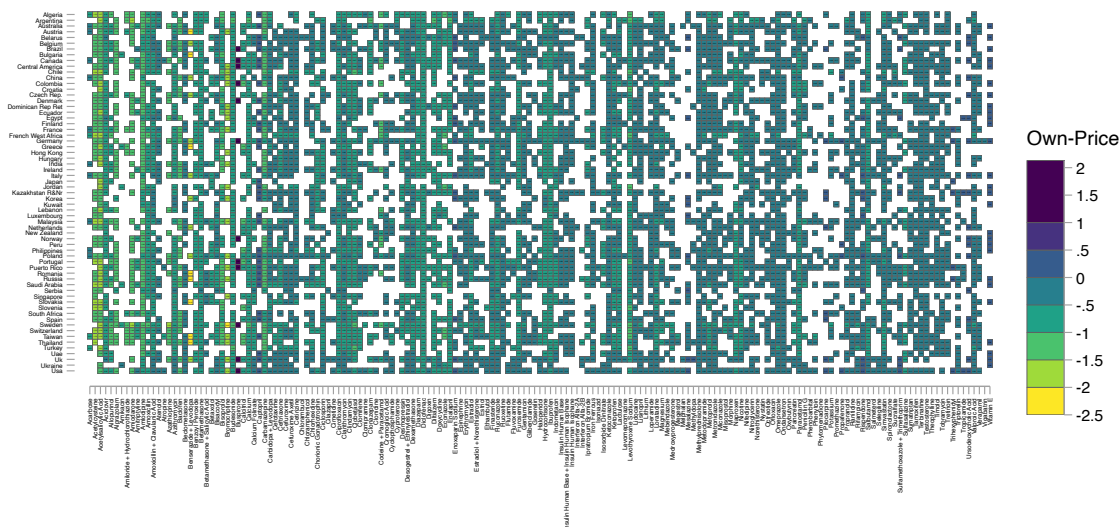


Figure 1. Own-price elasticities by country and molecule. Each cell reports the mean baseline ADL own-price elasticity across the statistically significant annual estimates available for a country–molecule pair during 2009–2020. Annual estimates whose 95 percent confidence interval includes zero are excluded before aggregation. More negative values indicate greater estimated price responsiveness.

When we focus on the country dimension, relatively strong responsiveness is observed not only in lower-income markets such as Jordan but also in high-income markets such as Sweden, providing an initial indication that pharmaceutical demand responsiveness is not simply ordered by national income as suggested by the empirical reduced form literature. Conversely, countries such as the USA and the UK contain many relatively less responsive country–molecule cells.

When we move to the molecule dimension, benserazide–levodopa, bromocriptine, acetylsalicylic acid, acetylcysteine, and amiodarone, display comparatively large negative elasticities across a number of markets, while other molecules exhibit substantially weaker responses. Importantly, however, the same molecule can still display markedly different elasticities across countries. The figure therefore suggests that there is no obvious pattern whereby a country or a molecule can simply be characterized as having uniformly elastic or inelastic pharmaceutical demand. Thus, demand responsiveness appears to be a joint property of the medicine and the national market in which it is sold²². This pattern supports allowing price responsiveness to vary jointly along the country and molecule dimensions rather than imposing either a common elasticity across medicines or a country-specific response that is constant across products.

Accordingly, our estimation does not remove institutional features that affect the responsiveness of quantity to price, since such slope heterogeneity is parameterized by the country, molecule, and year terms in K_{iget} and is part of the elasticity being estimated. Our estimates, therefore, describe how demand responds to prices under the institutional environments prevailing in each country and period. These include the set of rules determining how insurance affects what patients pay and whether substitution rules direct them toward cheaper alternatives. Cross-country differences in estimated responsiveness may consequently reflect

²²Patient-level studies also document differences in price responsiveness across drugs and therapeutic classes, associated with clinical context and available substitutes (Goldman et al., 2004; Landsman et al., 2005; Einav et al., 2018).

differences in reimbursement, procurement, substitution rules, insurance coverage, or other institutional arrangements, although the contribution of any individual mechanism cannot be separately identified. Evidence from Austria illustrates the relevance of these institutional differences. Gugler et al. (2026) estimate less price-responsive demand in the regulated, reimbursed segment than in the unregulated segment of the market for platelet aggregation inhibitors. Sweden also provides a documented illustration of this issue: despite its high income, it displays relatively strong estimated price responsiveness. Under its reference-pricing and substitution rules, patients who decline a cheaper interchangeable medicine generally pay the price difference giving them an incentive to switch products (Granlund and Rudholm, 2011; Granlund and Bergman, 2018). This mechanism may help explain the pattern observed in our estimates.

The estimates in Figure 1 provide a snapshot of the heterogeneous nature of pharmaceutical demand. However, the unbalanced pattern of molecule availability across countries cautions against interpreting the figure as a ranking of national demand responsiveness, since observed cross-country differences may partly reflect differences in the composition of medicines available in each market. We examine this possibility explicitly using a common-molecule sample. Figure 2 compares country-level average elasticities from the baseline DML sample with those obtained from a common basket of 24 molecules that remained on the market for at least 3 years. Countries with fewer than 50 products are excluded from the estimates.

The comparison indicates that both molecule composition and country-specific differences in responsiveness contribute to the observed cross-country heterogeneity. In the baseline DML sample, the UK, the USA, and Luxembourg lie toward the less price-responsive end of the distribution, whereas Argentina, Chile, Jordan, and Algeria lie toward the more responsive end. Restricting the sample to the common-molecule basket, however, substantially reshuffles this ordering, and among countries represented in both samples, the Spearman rank correlation is approximately 0.4. Differences in molecule coverage, therefore, account for an important part of the unconditional country ordering. At the same time, substantial dispersion in country elasticities remains within the common-molecule sample, indicating that differences in product composition alone do not account for the observed cross-country heterogeneity²³.

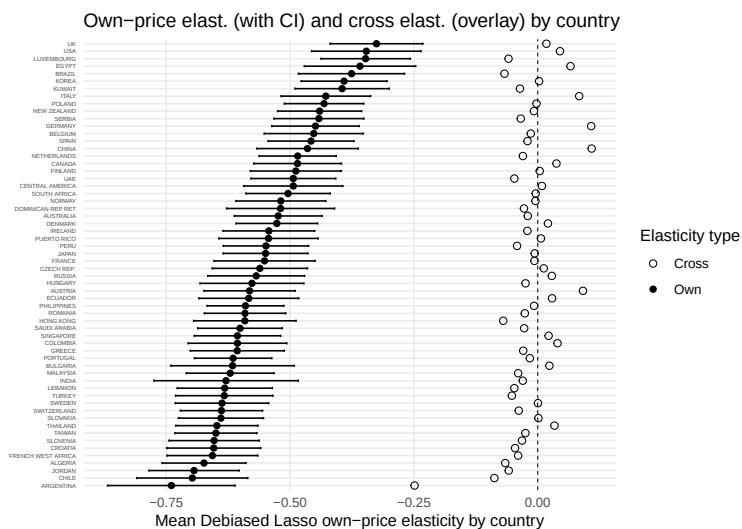
Looking at the same comparison from the molecule side reveals a considerably more stable pattern. Appendix Figure H.2, in particular, shows that most molecule-level mean own-price elasticities are negative and inelastic, although a lower tail extends below 1. Unlike the country ranking, however, the relative ordering of molecules is largely preserved when country coverage is made more comparable. Among the 24 molecules in the common-support sample, the Spearman rank correlation between the full and restricted samples is approximately 0.87. This stability suggests that molecule-level heterogeneity in price responsiveness is not primarily driven by differences in the countries in which particular medicines are observed. This pattern is consistent with the sensitivity exercise in Einav et al. (2018), Appendix C.5, which finds that drug elasticity estimates remain strongly correlated after restricting the insurance-plan sample. In our setting, the stable ranking could also reflect persistent differences in substitution possibilities in the sense that medicines that are easier to replace may remain relatively more price responsive across countries (Dubois and Lasio, 2018). Our comparison does not, however, establish this mechanism. Comparable evidence is, in general, scarce in the literature.

The descriptive comparisons in this section show that price responsiveness differs across medicines and countries, and that country averages are sensitive to which medicines are included. To assess whether prices are higher where demand is less price responsive, we therefore need to account for these compositional differences: comparisons of national averages alone may partly reflect which medicines are observed, rather

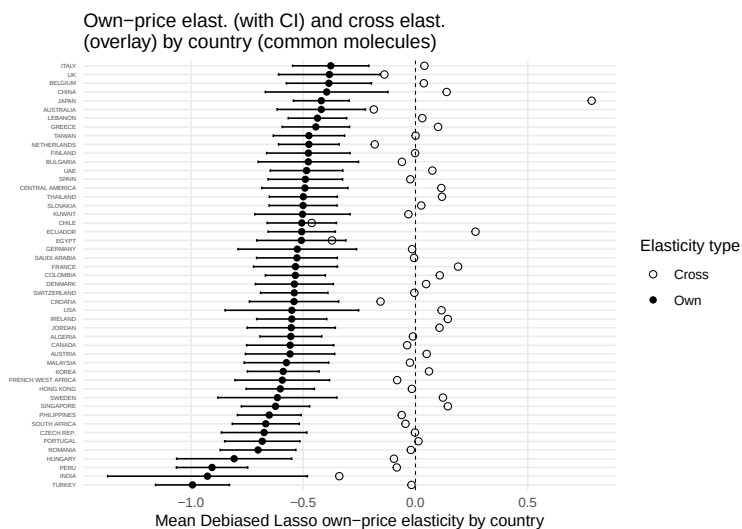
²³Because imposing common support also changes the estimation sample, this exercise should be interpreted as a sensitivity analysis rather than as a formal decomposition of composition and country effects.

than differences in responsiveness across national markets.

Accordingly, to address these composition concerns, Section 5 uses the country–molecule–year elasticity estimates to examine the price–responsiveness relationship while accounting for systematic differences that may be associated with both prices and demand responsiveness. First, we ask whether demand is less price-responsive in countries with higher average prices in a given year, accounting for average differences across molecules and common year effects. Second, we ask whether higher-priced molecules face less responsive demand, accounting for average differences across countries and common year effects.



(a) Baseline DML sample.



(b) Common-molecule restricted sample.

Figure 2. Price elasticities by country. Panel (a) reports country-level estimates from the baseline unbalanced DML sample. Panel (b) restricts the sample to molecules observed across all countries and present in the market for at least three years. Hollow circles denote country-level mean own-price elasticities from the baseline ADL specification, and horizontal bars report 95 percent confidence intervals based on clustered standard errors. The vertical dashed line marks zero. Countries are ordered by their point estimates, with more negative values indicating greater price responsiveness. Appendix H reports the corresponding molecule-level results.

4.1 Structural RCL Benchmark

In this section, we compare the own-price elasticities estimated by DML with those obtained from a structural RCL model as a diagnostic exercise. The RCL model is estimated using GMM-IV, with price instruments constructed as in Dubois et al. (2022). To strengthen comparability, we also estimate an RCL model using our data and adopt the estimates of Dubois et al. (2022) as the main external benchmark. Their study is particularly suitable for comparison because it examines pharmaceutical demand using closely related product-level data on prices, quantities, and sales. Products are classified at the ATC-4 level, quantities are measured in standard units, and sales are reported in US dollars at fixed exchange rates²⁴. Their demand model defines markets as country–ATC-4–quarter cells and estimates substitution among products within the same ATC-4 class. We therefore adopt the same market definition in the benchmarking exercise.

To make the DML and RCL estimates as comparable as possible, we harmonize the level at which demand responses are measured, as detailed in Appendix I. Table 2 compares our DML and RCL own-price elasticities with those reported by Dubois et al. (2022). Here, we call the structural specification of the latter "low-dimensional" since its number of unrestricted coefficients does not exceed either the number of observations or the number of instrument moments, avoiding the rank failure studied in Proposition 2.2. In this benchmarking exercise, the orthogonal OLS estimates are generally closer to the reported RCL elasticities. This is not surprising, since the comparison is deliberately constructed as a relatively low-dimensional exercise rather than as the high-dimensional sparse setting that motivates our baseline estimator. More specifically, the reason is that the heterogeneity structure is selected *ex ante* to match the level at which Dubois et al. (2022) report their elasticities, namely ATC-4 $\times$ branded/generic cells. Orthogonal OLS retains the full set of these pre-specified cell-level responses and estimates a separate elasticity for each supported cell. In this sense, it is closer in spirit to the dense elasticity structure generated by the RCL benchmark.

While OLS estimates all the *pre-specified* cell responses in Dubois et al. (2022) without shrinkage, the richer ADL specification allows additional heterogeneity within these cells via K . A possibly intuitive way to see how this would create structural difficulties to RCL is to figure that adding variables to K introduces further price interactions, allowing responsiveness to depend on more product and market characteristics. In RCL, analogous interactions between own log price and these characteristics could enter mean utility, adding coefficients to the linear parameter block θ_1 (Nevo, 2000; Dubois et al., 2022). If the instrument moments cannot distinguish these additional coefficients, different coefficient combinations produce identical moments, preventing unique unrestricted estimation as described in Proposition 2.2.

In our context, the ADL estimator imposes an additional sparsity discipline on the residualized heterogeneous price-response coefficients. This regularization is valuable when the heterogeneity dictionary is high-dimensional and the underlying price-response function is approximately sparse, but it need not reproduce the dense elasticity structure implied by a relatively low-dimensional RCL specification. In the present comparison, if many ATC-4 $\times$ branded/generic cells exhibit nonzero price responses, regularization may shrink some cell-specific coefficients toward zero, even after the one-step adjustment, yielding elasticities that are less negative than those obtained from orthogonal OLS or the RCL benchmark.

We therefore report two ADL specifications. In the low-dimensional (LD) ADL case, the heterogeneity dictionary $K^{\text{US,LD}}$ contains only the pre-specified ATC-4 $\times$ branded/generic reporting-cell terms used to harmonize the comparison with Dubois et al. (2022). In the high-dimensional (HD) ADL case, the heterogeneity dictionary $K^{\text{US,HD}}$ contains the ATC-4 $\times$ branded/generic reporting-cell indicators together with additional molecule, corporation, time, age, lagged, institutional, and market-state effect modifiers.

²⁴We restrict the external comparison to the U.S. elasticities reported by Dubois et al. (2022). Although their study also reports estimates for Canada, a substantial share of the Canadian elasticities is missing. An alternative benchmark would be the ATC-4 class A02B estimates for France, the UK, and the US reported by Dubois and Lasio (2018). We did not perform such a comparison due to lack of data with respect to the latter specification.

Given the two ADL specifications mentioned, the comparison in Table 2 illustrates an important trade-off: regularization is useful when the relevant dimensions of heterogeneity are numerous and only approximately sparse, but it maybe unnecessarily restrictive when the economically relevant heterogeneity structure is low-dimensional and largely prespecified. For this reason, the following empirical models report results using both orthogonal OLS and the ADL estimator, including in Table 3.

Table 2. Own-price elasticities by ATC-4 class and product type: DML estimates versus Dubois et al. (2022) for common ATC-4s

		US				
ATC-4 Class		Orthogonal OLS	ADL (low-dim)	ADL (high-dim)	RCL (our)	RCL (Dubois et al., 2022)
A10H0	Branded	-1.604	-0.959	-0.899	-1.472	-1.494
	Generic	–	–	–	–	-1.315
B3A1	Branded	-1.155	-0.850	-0.163	-1.548	-1.563
	Generic	–	–	–	–	-1.177
C2A1	Branded	-0.793	-0.005	-0.293	-1.563	-1.543
	Generic	-0.875	-0.214	-0.511	-1.300	-1.272
C7A0	Branded	-2.286	-1.239	-1.946	-1.596	-1.522
	Generic	-1.677	-1.312	-1.137	-1.576	-1.475
C8A0	Branded	-2.163	-0.847	-0.743	-1.595	-1.520
	Generic	-1.141	-0.905	-0.887	-1.577	-1.406
L1A0	Branded	-0.870	-0.013	-0.611	-1.410	-1.512
	Generic	-0.624	-1.112	-0.169	-1.136	-1.410
L2A2	Branded	-1.155	-0.008	-1.206	-1.589	-1.455
	Generic	-0.611	-0.421	-0.585	-1.113	-0.288
L4X0	Branded	-2.925	-0.786	-1.048	-1.539	-1.453
	Generic	-1.123	-1.250	-0.767	-1.461	-1.445
M1A1	Branded	-1.417	-0.474	-1.531	-1.552	-1.555
	Generic	-0.558	-0.569	-0.812	-1.478	-1.447
N1B1	Branded	-1.929	-1.007	-1.181	-1.545	-1.488
	Generic	-0.619	-0.425	-0.058	-1.506	-1.351
N2A0	Branded	-1.292	-1.308	-0.659	-1.518	-1.531
	Generic	-1.402	-1.187	-0.927	-1.508	-1.455
N2B0	Branded	-1.3393	-1.123	-1.544	-1.582	-1.556
	Generic	-1.184	-1.122	-1.201	-1.590	-1.455
N3A0	Branded	-2.486	-1.313	-0.873	-1.586	-1.541
	Generic	-1.081	-0.966	-0.998	-1.576	-1.468
N5A1	Branded	-3.927	-3.141	-1.096	-1.587	-1.432
	Generic	-1.208	-0.855	-0.814	-1.531	-1.334
N5A9	Branded	-0.162	-0.023	-0.341	-1.603	-1.568
	Generic	-1.129	-0.846	-0.653	-1.418	-1.413
N5C0	Branded	-2.545	-1.701	-1.266	-1.602	-1.575
	Generic	-1.086	-1.035	-1.079	-1.549	-1.419
N6A4	Branded	-2.397	-2.058	-0.163	-1.601	-1.396
	Generic	-0.745	-1.241	-0.715	-1.543	-1.359
N6A9	Branded	-0.645	-0.727	-0.573	-1.602	-1.541
	Generic	-0.947	-0.693	-1.129	-1.455	-1.476
US avg. (present estimates)	Overall	-1.006	-0.754	-1.112	-1.506	[-1.2, -0.3] (U.S. literature range)

Note: A dash denotes an unavailable estimate. The U.S. literature range is assembled from several sources, including Manning et al. (1987), Danzon and Chao (2000a), Gilman and Kautter (2008), Chandra et al. (2014), Brot-Goldberg et al. (2017), Ellis et al. (2017), Einav et al. (2018), and Yeung et al. (2018).

Table 3 shows how – although the high-dimensional ADL estimates preserve the predominantly negative pattern of own-price responses – they only partly reproduce the RCL elasticities reported by Dubois et al. (2022) exactly. Two conceptually distinct factors may account for these differences²⁵. First, the estimates rely on different datasets and empirical supports. Dubois et al. (2022) focus on hospital purchases, aggregate products at the molecule–corporation–market level, and combine generic products within molecules. Our data cover a different period, use a different product aggregation, and do not separately identify hospital and retail sales. These differences affect the distributions of prices and quantities, the composition of choice sets, and the joint distribution of the characteristics entering the estimators. These differences are particularly relevant for LASSO, which uses the data to select terms and shrink their coefficients. Changing the sample can therefore change which price interactions are retained and how strongly they are shrunk, especially when their separate relationships with sales are difficult to distinguish. Differences in sales channels may provide one possible interpretation of this point. Relative to a hospital-only sample, a sample combining hospital and retail sales, as the one we use, may generate smaller elasticities in absolute value if retail utilization is less directly responsive to ex-manufacturer price variation because demand is mediated by insurance coverage, prescribing decisions, brand loyalty, or clinical inertia (Goldman et al., 2007; Duggan and Morton, 2010; Einav et al., 2015). Despite being a potential channel, this mechanism cannot be directly evaluated in our empirical setting since our data do not identify the sales channel.

Second, the estimators impose different forms of pooling and structure across products and markets.

Our RCL model is estimated on the same mixed-channel data used for the DML specifications, yet it produces elasticities that are relatively close to those reported by Dubois et al. (2022). This suggests that – under closely aligned structural specification – RCL can generate similar elasticity patterns despite differences in sample composition. In constructing our RCL benchmark, we deliberately align the market definition, utility specification, fixed effects, choice-set instruments, and market-size calibration with Dubois et al. (2022). The model assumes a common distribution of price sensitivity across products and quarters within each country, allowing observations from different medicines and periods to inform the same preference parameters. Together with product characteristics and available alternatives, these parameters determine how demand for each medicine responds to its own price and competitors’ prices.

In the remainder of this section, we briefly overview the existing evidence on the overall (average) pharmaceutical demand responsiveness in the US (the last "US avg." line of Table 3), to provide a benchmark for our estimates. However, much of this literature estimates patient-level responses to changes in out-of-pocket drug prices, often exploiting variation in insurance copayments or cost-sharing arrangements. These estimates, therefore, capture a different object from the market-level elasticity with respect to ex-manufacturer prices studied here and should be interpreted primarily as a broad benchmark for magnitude rather than as directly comparable estimates.

The RAND Health Insurance Experiment provides one of the earliest and best-known estimates in this literature. Using randomized variation in cost sharing, Manning et al. (1987) report an elasticity of approximately 0.2 for total health-care utilization, including prescription drugs. Subsequent studies exploiting changes in insurance cost sharing generally find relatively inelastic responses of pharmaceutical utilization, with estimates ranging from approximately 0.33 to 0.12 (Gilman and Kautter, 2008; Chandra et al., 2014; Brot-Goldberg et al., 2017). Considerable heterogeneity is also observed across therapeutic categories. Goldman et al. (2004), for example, estimate elasticities with respect to drug days supplied of 0.45

²⁵A further factor may also be the specification difference, concerning how elasticities depend on own price. Both DML benchmark specifications allow elasticities to vary with observed characteristics, but hold them constant with respect to own price conditional on those characteristics. In RCL, purchasing probabilities respond nonlinearly to utility, so elasticities can change with own price even when mean utility is linear in log price. Adding further price–characteristic interactions to our DML dictionary does not capture this source of variation.

for NSAIDs, 0.44 for antihistamines, 0.34 for antihyperlipidemics, 0.32 for antiasthmatics, 0.26 for antihypertensives and antidepressants, and 0.25 for antidiabetics. Similarly, Landsman et al. (2005) find smaller responses for drugs treating largely asymptomatic conditions, such as ACE inhibitors and statins (0.10 to 0.16), and larger responses for drugs used for symptomatic conditions, including COX-2 inhibitors, NSAIDs, and SSRIs (0.24 to 0.60).

A related concern is that changes in insurance cost sharing may simultaneously affect the effective prices of multiple substitute or complementary drugs. Yeung et al. (2018) emphasize that, in such settings, estimating an own-price response without accounting for accompanying changes in the prices of related medicines may confound own- and cross-price effects. Using a value-based formulary design that explicitly incorporates substitution and complementarity across drugs, they report an overall pharmaceutical price elasticity of approximately -0.16. Other studies exploit instrumental-variable strategies to isolate exogenous price variation and obtain estimates in the range of approximately -0.14 to -0.4 (Ellis et al., 2017; Einav et al., 2018).

More closely related to our empirical setting, Danzon and Chao (2000a) use comprehensive IMS retail-pharmacy data for seven countries and estimate two-stage least-squares quantity equations at the molecule level. They report a U.S. own-price elasticity of approximately -1.29, substantially larger in absolute value than most patient-level cost-sharing estimates. This difference is important for interpretation: their elasticity is based on manufacturer-level price variation and market-level substitution across pharmaceutical products rather than on changes in patient out-of-pocket prices. It therefore provides a more directly relevant benchmark for the demand responses estimated in our analysis.

5 Evidence on the Inverse Elasticity Rule Pricing

The previous sections recover the heterogeneous demand responses needed to examine whether international pharmaceutical prices are systematically related to price responsiveness. The relevant theoretical benchmark for this task is the discussed inverse-elasticity principle: conditional on the financing or revenue requirement, larger price–cost margins can be sustained in markets with less elastic demand, whereas more price-responsive markets should bear smaller margins. Appendix J develops this logic in a continuous-choice framework that is close to the canonical Ramsey and pharmaceutical differential-pricing formulations. Namely, it extends current setups to a non-homothetic CES model with interdependent demands and price distortions due to insurance or reimbursement wedges. In this framework, each consumer group allocates a given drug budget between branded and generic medicines, choosing how much of each to consume. The latter set-up is particularly useful in that it provides two results that are particularly relevant for the empirical analysis below and relatively easy to implement and test. First, income is not, in general, a sufficient statistic for price responsiveness. Even for a common underlying substitution parameter, the Marshallian price elasticity – i.e. the response to a price change holding the group’s drug budget fixed – depends not only on substitution possibilities, but also on expenditure shares and on how pharmaceutical demand responds to total drug expenditure. GDP per capita may therefore correlate with some determinants of demand, but it does not provide a direct mapping to the elasticity relevant for differential pricing.

Second, Appendix J formalizes the inverse-elasticity benchmark and clarifies its welfare implications. In this aggregate Ramsey model, the optimal proportional markup satisfies

$$\frac{P_{R,c} - MC}{P_{R,c}} = \frac{\mu_c}{\eta_c},$$

where η_c denotes the magnitude of the aggregate demand elasticity in country c , and μ_c measures the tightness of its financing constraint. Intuitively, for a given financing pressure, the rule assigns larger proportional

markups to less responsive markets, where price increases cause smaller percentage reductions in consumption. Allowing financing pressure to differ across countries implies that even markets with similar elasticities can have different optimal markups²⁶. When financing pressure is common across countries ($\mu_c = \mu$), proportional markups vary inversely with the magnitude of aggregate demand elasticity.

Given that neither marginal costs nor the corresponding μ_c are observed in our panel, Appendix K establishes conditions under which the price coefficients provide evidence about markups despite unobserved differences in marginal costs and will be discussed through the Section whereby appropriate.

Section 5.1 relates estimated demand responsiveness to both observed log prices and log prices residualized using the demand-stage controls, and examines whether GDP per capita summarizes the estimated differences in price responsiveness.

Section 5.2 illustrates the policy relevance of the findings in Section 5.1 by using the estimated international price–responsiveness relationship to construct an elasticity-aligned US price for the branded segment relevant to the MFN proposal described in the Introduction.

5.1 Assessing the presence of the inverse-elasticity rule in global markets

We examine here whether international pharmaceutical prices display the ordering suggested by the inverse-elasticity benchmark, i.e. whether higher prices are associated with less price-responsive demand. A direct test of the pricing rule would relate price–cost margins to demand elasticities, conditional on the relevant financing requirement. However, as underlined above, marginal costs are not observed in our global panel and recovering them would require a harmonized structural supply model with additional assumptions about firm conduct across different institutional settings. In order to mitigate the latter concern, we studied the relationship between estimated demand responsiveness and not only observed log prices but also prices residualized using the demand-stage controls. This solution is motivated by the findings of Appendix K, which establish conditions under which these comparisons are informative about price–cost ratios despite unobserved differences in marginal costs. Specifically, after the same adjustments, the price coefficient’s sign reflects the markup association if the remaining log-cost component is uncorrelated with estimated elasticities. Auxiliary structural diagnostics support the use of residualized prices. As an example, observable characteristics predict substantial model-implied cost variation and attenuate its association with elasticities. More details are provided below and in Appendix K.

More formally, we relate the estimated own-price elasticities to average prices in two complementary fixed-effects regressions. Let $\hat{\epsilon}_{gca}^{\text{own},e}$ denote the elasticity for molecule g , country c , and calendar year a , obtained using final-stage estimator $e \in \{\text{ADL}, \text{OLS}\}$, where OLS denotes orthogonal OLS. We estimate

$$\hat{\epsilon}_{gca}^{\text{own},e} = \beta_C^e \bar{p}_{ca} + \alpha_g^e + \lambda_{C,a}^e + \zeta_{gca}^{C,e}, \quad (7)$$

$$\hat{\epsilon}_{gca}^{\text{own},e} = \beta_M^e \bar{p}_{ga} + \delta_c^e + \lambda_{M,a}^e + \zeta_{gca}^{M,e}. \quad (8)$$

Here, $\bar{p}_{ca}$ is the average log product price in country c and year a , while $\bar{p}_{ga}$ is the corresponding average for molecule g and year a . The labels C and M distinguish the country-price and molecule-price regressions. The terms α_g^e and δ_c^e are molecule and country fixed effects, respectively, and $\lambda_{C,a}^e$ and $\lambda_{M,a}^e$ are regression-specific year effects, and the ζ terms are projection residuals. Additional controls enter through the price predictions used for residualization.

The first regression asks whether higher-price country–year environments have less responsive demand, accounting for molecule and year effects. The second asks whether higher-priced molecule–years have less

²⁶Appendix J further establishes when these financing differences reverse the ranking of welfare losses as a share of surplus, showing why elasticities alone cannot determine that ranking.

responsive demand, accounting for country and year effects. In other words, Equation (7) compares responsiveness across countries within molecules, while Equation (8) compares responsiveness across molecules within countries. The coefficients β_C^e and β_M^e measure how much estimated own-price elasticities differ with average log prices, after accounting for the fixed effects. For example, a difference of 0.1 in average log prices corresponds to a fitted elasticity difference of $0.1\beta_C^e$ or $0.1\beta_M^e$, respectively. A positive coefficient therefore means that higher prices tend to accompany less negative elasticities, i.e. that demand responds less to a given percentage change in price.

We repeat both regressions using $\hat{\varepsilon}_{gca}^{\text{rival},e}$ as the dependent variable to examine the corresponding relationships for rival-price responsiveness. Furthermore, as mentioned, we also repeat the exercise by replacing observed average log prices with averages of residualized log prices. Using the notation of Section 2, we reuse the estimated own-price residuals $\hat{p}_{igct}^\circ$ from the demand-estimation stage. Remember that each residual is the difference between observed log price and its prediction from $\mathcal{W}_{igct}$, obtained from a model fitted and tuned without that observation. Moreover, we average these residuals within country–year and molecule–year groups, denoting the resulting measures by $\hat{p}_{ca}^\circ$ and $\hat{p}_{ga}^\circ$. These replace $\bar{p}_{ca}$ and $\bar{p}_{ga}$ in the respectively in Equations (7) and (8), retaining the same fixed effects. The comparison in the latter configuration therefore asks whether the price–elasticity association persists after subtracting the price component predicted by the observed characteristics and fixed effects used in demand estimation.

Because the elasticity estimates for different molecule–country–year combinations vary in precision, we give greater weight to those with smaller estimated variances. Specifically, we use inverse-variance weights based on the ADL elasticities, retaining the same weights in the orthogonal OLS projections. Neyman orthogonality and cross-fitting, moreover, make nuisance-estimation errors asymptotically second order in each elasticity estimate (Chernozhukov et al., 2018), while the final-stage inference accounts for sampling uncertainty and the ridge adjustment attenuates LASSO shrinkage (Chernozhukov et al., 2022). We cluster standard errors by country in Equation (7) and by molecule in Equation (8). This allows observations within the same country or molecule to be affected by common shocks, including across years. The reported standard errors do not, however, fully incorporate uncertainty from the preceding demand-estimation stage. Elasticities across cells share estimated coefficients, and the residualized prices also depend on estimated predictions. Accounting for all these sources of uncertainty would require a joint variance calculation or a suitably designed bootstrap.

The results are presented in Table 3. Panel 1 of Table 3 reports the above described relationships for branded and generic medicines together and for each group separately. It presents country and molecule projections using observed and residualized prices, with columns distinguishing own- and rival-price responsiveness and the estimator used to obtain the elasticities. The panel also reports associations with national income, discussed below. For own-price responsiveness, the clearest pattern emerges among branded medicines for which higher prices consistently accompany less negative elasticities, indicating weaker demand responses. With residualized prices, the (across) country coefficients are 0.063 under ADL and 0.129 under orthogonal OLS. Correspondingly, the (across) molecule regressions show that higher-priced medicines also tend to face less responsive demand, after accounting for country and year effects. This association is similar across the two estimators, with coefficients of 0.249 under ADL and 0.231 under orthogonal OLS. All four coefficients are statistically significant. The association is also positive using observed prices, although its magnitude and precision vary.

Putting it differently, we find that among branded medicines, higher prices are associated with less responsive demand. While there is no consensus in the literature, a possible source of this association is variation in how medicines are purchased and reimbursed. Insurers that can steer purchases toward competing medicines make manufacturers’ sales more sensitive to their prices and can use that threat to obtain lower prices. Dug-

gan and Morton (2010) provide evidence for this mechanism ²⁷.

We next assess whether this price ordering also appears when branded and generic products are considered together. In the combined sample in Panel 1, higher residualized prices are associated with less responsive demand across both countries and molecules, with statistically significant coefficients under both estimators. The country association is weak when observed prices are used. Across molecules, the positive association appears with either price measure. These results extend the finding to the combined sample. Panel 2 provides a complementary comparison by grouping molecules according to whether their observed products are predominantly branded or generic. Notice that both product types are retained within each selected molecule, including generic products that compete with branded ones²⁸. The residualized-price coefficients are positive for both groups across the two estimators and both projections, but are estimated more precisely for predominantly branded molecules.

A separate concern throughout the regressions are composition effects, i.e. the fact that countries differ in which medicines are observed. Country price averages may therefore reflect differences in molecule coverage as well as differences in prices. Panel 3 examines this possibility using molecules observed across all 63 countries and present for at least three years, subject to the country sample-size restriction. This makes molecule coverage more comparable across countries, but concentrates the analysis on widely marketed medicines that may differ systematically from those excluded. The exercise therefore deliberately assesses whether the association persists within this common set, without eliminating selection associated with product availability.

The residualized-price coefficients in the across-country comparison remain positive and statistically significant: 0.054 under ADL and 0.142 under orthogonal OLS. The across-molecule coefficients are also positive, although larger and less precisely estimated. The country association therefore persists when molecule coverage is made more comparable. The less consistent findings for generic products could partly reflect competitive constraints on pricing. If competition keeps generic prices close to marginal costs, differences in demand responsiveness may translate into smaller differences in margins, while cost differences remain an important source of price variation. The comparisons do not directly test this mechanism, and more generally, interpreting the price–elasticity association in terms of markups requires examining whether marginal costs also vary with demand responsiveness.

Accordingly, Appendix K clarifies when the price–elasticity association can be interpreted as evidence about markups. The idea is that higher prices may reflect higher costs or larger markups. If the cost differences left after our adjustments are unrelated to estimated demand responsiveness, they cannot explain the positive price–elasticity association. In that case, the association indicates that markups relative to costs tend to be larger where demand is less price responsive. The relevant question is therefore whether the controls can remove cost differences that would otherwise contribute to the price–elasticity association. In the auxiliary U.S. sample, we first recover model-implied marginal costs by combining RCL demand estimates with Bertrand pricing under observed multiproduct ownership. Using these recovered costs, we then assess how much cost variation observable characteristics can explain. The set of diagnostic controls predicts log marginal costs with an out-of-sample R^2 of 0.815, suggesting that much of this variation can be accounted for through the controls. More directly relevant to the markup interpretation, the association between log costs and the sign-reversed ADL own-price elasticity weakens after both variables are adjusted for these controls with the estimated slope changing from approximately -0.688 to 0.128 , with a standard error of

²⁷More specifically, Duggan and Morton (2010) shows that Medicare Part D generated relative price declines for branded medicines where insurers had scope to redirect purchases toward therapeutic alternatives. Differences in such purchasing arrangements could therefore contribute to the price–elasticity ordering we observe.

²⁸A molecule is predominantly branded (generic) if branded (generic) products account for at least half its observations, pooled across countries, in more than half its observed years and more often than the other type.

0.112 for the latter. Thus, the adjustment captures substantial cost variation and reduces the estimated cost-elasticity association in this auxiliary sample ²⁹.

Alongside these own-price results, we examine whether responses to competitors' prices also vary with price levels. The rival-price-index elasticity measures how demand responds to the prices of other products containing the same molecule in the same country and period. Positive values are consistent with buyers switching toward a product when its rivals become more expensive. In Table 3, higher prices are associated with more positive rival-price elasticities in several across-country projections for the pooled and branded samples, while the evidence across molecules is less consistent. Differences in national reimbursement and dispensing rules may contribute to this variation, since they influence both prices and switching between competing medicines (Duggan and Morton, 2010; Dubois and Lasio, 2018).

Finally, we ask whether national income provides a useful summary of the estimated differences in demand responsiveness, given its frequent use as a proxy for demand conditions in international pharmaceutical-pricing research and policy. We estimate two versions of the country specification, retaining molecule and year fixed effects. The first replaces average log prices ($\bar{p}_{ca}$) with log GDP per capita at purchasing power parity (PPP), asking whether richer country-years tend to have less price-responsive demand. The second compares the pre-specified high-income markets (HICs) listed in the table notes with the remaining markets. These comparisons use our baseline, uncompensated (Marshallian) elasticities. For the pooled own-price comparisons, we also report an exploratory Slutsky-style adjustment toward compensated (Hicksian) elasticities. Intuitively, a price increase can reduce purchases both by making a medicine relatively more expensive and by reducing the purchasing power of a given drug budget. The Hicksian elasticity isolates the substitution response by allowing expenditure to adjust enough to keep the consumer group's utility unchanged. The Hicksian comparison therefore asks whether the relationship between income and demand responsiveness changes after adjusting for the purchasing-power effect of a price change. The estimated income coefficients change very little after the adjustment, suggesting that our conclusions about income are not sensitive to this correction. In the pooled sample, higher income remains associated with less negative own-price elasticities, but this pattern is not consistent across all product groups. The adjustment therefore does not make GDP a reliable summary of demand responsiveness. Nor does it isolate an income mechanism, since GDP may also reflect differences in insurance, reimbursement, and purchasing arrangements.

To provide some numbers, in the pooled sample, the log GDP coefficients are 0.025 under ADL and 0.036 under orthogonal OLS. A doubling of GDP per capita is therefore associated with own-price elasticities that are approximately 0.017 and 0.025 points less negative, respectively. The listed high-income markets have elasticities that are, on average, 0.070 points less negative under ADL and 0.063 points less negative under orthogonal OLS, conditional on molecule and year effects. Applying the Slutsky-style adjustment changes the ADL log GDP coefficient from 0.0251 to 0.0256 and the high-income-group coefficient from 0.0704 to 0.0716.

²⁹These findings provide indirect support for interpreting the adjusted-price results as evidence about markup ordering. Extending that interpretation to the international panel remains conditional on the corresponding cost restriction, which the U.S. diagnostic does not establish.

Table 3. Demand responsiveness and international pharmaceutical prices: Evidence on inverse-elasticity ordering

		ADL (own)		Orthogonal OLS (own)		ADL (rival)		Orthogonal OLS (rival)	
		Country	Molecule	Country	Molecule	Country	Molecule	Country	Molecule
Panel 1									
Branded + Generics	Original price	0.0012 (0.0102)	0.0405*** (0.0128)	0.0043 (0.0261)	0.0350*** (0.0120)	0.0303* (0.0151)	-0.0005 (0.0087)	0.1226* (0.0684)	-0.0108 (0.0133)
	Partialled out price	0.0621*** (0.0121)	0.3038*** (0.0780)	0.1166*** (0.0118)	0.2428** (0.0934)	0.0252*** (0.0081)	0.0174 (0.0931)	0.4544*** (0.0234)	0.0336 (0.1265)
	<i>GDP PPP:</i>								
	Log GDP (PPP) Marshallian	0.0251* (0.0143)		0.0364** (0.0167)		0.0036 (0.0073)		0.0425 (0.0404)	
	Log GDP (PPP) Hicksian	0.0256* (0.0141)		0.0357** (0.0161)					
	Listed HICs vs other markets Marshallian	0.0704*** (0.0212)		0.0630** (0.0247)		0.0302*** (0.0108)		0.0611 (0.0391)	
	Listed HICs vs other markets Hicksian	0.0716*** (0.0210)		0.0634*** (0.0233)					
Generics only	Original price	-0.0043 (0.0201)	0.0186 (0.0214)	-0.0168 (0.0285)	0.0234 (0.0315)	0.0531*** (0.0147)	-0.0176 (0.0171)	0.0673*** (0.0194)	-0.0121 (0.0297)
	Partialled out price	-0.0107 (0.0574)	0.1487** (0.0663)	0.1552*** (0.0435)	0.1276 (0.0931)	-0.0522 (0.0504)	0.1413* (0.0741)	-0.1946** (0.0922)	0.2172* (0.1185)
	<i>GDP PPP:</i>								
	Log GDP (PPP)	-0.0160 (0.0424)		-0.0276 (0.0529)		0.0075 (0.0349)		-0.0543 (0.0403)	
	Listed HICs vs other markets	0.0209 (0.0455)		-0.0991 (0.1391)		-0.0214 (0.0261)		-0.0210 (0.0209)	
Branded only	Original price	0.0173*** (0.0006)	0.0374*** (0.0106)	0.0423* (0.0212)	0.0302** (0.0129)	0.0424* (0.0216)	0.0073 (0.0126)	0.1040** (0.0457)	0.0045 (0.0156)
	Partialled out price	0.0628*** (0.0205)	0.2489*** (0.0721)	0.1291*** (0.0162)	0.2308*** (0.0861)	0.0253** (0.0115)	0.0716 (0.0786)	0.3405*** (0.0187)	0.0924 (0.0986)
	<i>GDP PPP:</i>								
	Log GDP (PPP)	0.0109 (0.0115)		0.0237* (0.0140)		0.0128 (0.0097)		0.0486 (0.0330)	
	Listed HICs vs other markets	0.0251 (0.0162)		0.0342* (0.0202)		0.0391** (0.0170)		0.0762* (0.0398)	
Panel 2									
Products belonging to molecules with mostly generics (mainly generics competition)	Original price	-0.1827* (0.0999)	0.0854* (0.0403)	-0.1660 (0.1193)	0.0877* (0.0414)	0.0816 (0.0725)	-0.0118* (0.0063)	0.0905 (0.0815)	-0.0289* (0.0140)
	Partialled out price	0.0946 (0.1199)	0.4586** (0.2051)	0.2508 (0.1886)	0.4314 (0.2473)	0.1727 (0.1360)	-0.0606 (0.0468)	0.1110 (0.1462)	-0.3692 (0.2301)
	<i>GDP PPP:</i>								
	Log GDP (PPP)	0.1489 (0.1266)		0.0309 (0.1218)		-0.0837 (0.1065)		-0.0384 (0.1160)	
	Listed HICs vs other markets	-0.0025 (0.1727)		0.0859 (0.1237)		-0.0540 (0.1576)		-0.0569 (0.1555)	
Products belonging to molecules with mostly branded products (mainly branded composition)	Original price	0.0032 (0.0198)	0.0432*** (0.0109)	0.0056 (0.0302)	0.0390*** (0.0133)	0.0454*** (0.0167)	-0.0028 (0.0084)	0.1311** (0.0579)	-0.0182 (0.0140)
	Partialled out price	0.0447** (0.0174)	0.2990*** (0.0768)	0.1685*** (0.0242)	0.2941*** (0.0941)	0.0281 (0.0187)	0.0072 (0.0514)	0.4070*** (0.0249)	0.1396 (0.1400)
	<i>GDP PPP:</i>								
	Log GDP (PPP)	-0.0096 (0.0203)		0.0002 (0.0273)		0.0007 (0.0261)		0.0187 (0.0407)	
	Listed HICs vs other markets	0.0145 (0.0235)		0.0205 (0.0310)		-0.0003 (0.0270)		0.0282 (0.0384)	
Panel 3									
Products belonging to molecules common across countries for at least 3 years (no country composition effects)	Original price	-0.0591 (0.0375)	0.1267** (0.0578)	-0.0474 (0.0500)	0.1399 (0.0850)	0.0836*** (0.0307)	0.0832 (0.0549)	0.1388*** (0.0366)	0.1795 (0.1220)
	Partialled out price	0.0535** (0.0248)	0.8977** (0.3698)	0.1421*** (0.0285)	0.9848 (0.5919)	0.0215 (0.0279)	0.6201 (0.3907)	0.1996*** (0.0216)	1.3241 (0.8625)
	<i>GDP PPP:</i>								
	Log GDP (PPP)	0.0318 (0.0409)		0.0161 (0.0201)		0.0005 (0.0298)		0.1250 (0.0900)	
	Listed HICs vs other markets	0.0733 (0.0542)		0.0539* (0.0278)		0.0866** (0.0431)		0.2172** (0.0956)	

Clustered standard errors are reported in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Entries are coefficients from regressions of estimated molecule–country–year own- and rival-price elasticities on prices or income measures. “ADL” and “orthogonal OLS” identify the elasticity estimator. “Country” projections include molecule and year fixed effects; “Molecule” projections include country and year fixed effects. Prices are averaged by country–year or molecule–year, respectively, using observed log prices or their residuals from demand-stage predictions. Income rows replace prices with log GDP per capita (PPP) or the listed-HIC indicator in the country specification. Elasticities are uncompensated (Marshallian) unless labelled “Hicksian”; these rows report exploratory Slutsky-style adjustments for the pooled own-price income comparisons.

“Listed HICs” denotes the set of markets assigned a high-income indicator in this exercise: Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Japan, Kuwait, Luxembourg, the Netherlands, New Zealand, Norway, Poland, Korea, Singapore, Slovakia, Spain, Sweden, Switzerland, the United Arab Emirates, the United Kingdom, Ukraine, and the United States. “Other markets” denotes the complement in the estimation sample.

Panel 2 classifies molecules as mostly generic (branded) when that type accounts for at least half their annual observations, pooled across countries, in a majority of observed years and more often than the other type. Both product types are retained.

In Panel 3 estimates, the analysis is restricted to molecules that are common across all 63 countries and remained on the market for at least three years. As for Figures 2b and H.2b. Countries with fewer than 50 products remaining after this restriction were dropped.

5.2 A Descriptive policy exercise: MFN example

One of the main findings of the preceding analysis is that higher branded-drug prices are associated with less responsive demand, while national income does not consistently summarize this responsiveness. These findings have practical policy implications for international reference pricing, like the fact that countries selected for their similar income levels need not have similar demand conditions.

We illustrate this through a historical exercise motivated by the U.S. MFN pricing targets announced in May 2025. The Department of Health and Human Services (HHS) called on manufacturers to align U.S. prices for branded medicines without generic or biosimilar competition with the lowest price among OECD countries whose GDP per capita PPP reached at least 60 percent of the U.S. level. The selected foreign price would thus serve as a target for lowering U.S. prices. However, selecting countries by income, as shown, does not ensure that their demand responds to prices in the same way as U.S. demand. Our estimates allow us to examine how the measured U.S. price premium changes when the foreign benchmark also accounts for demand responsiveness and observed country characteristics.

For each year, we apply the 60-percent income threshold by comparing each OECD country’s annual PPP GDP per capita with the U.S. value for that same year. For each molecule–period, we call the lowest price among eligible foreign countries the historical pseudo-MFN price. We compare observed U.S. prices (i) with this minimum – which measures the gap to the historical MFN reference – (ii) with the eligible countries’ mean price – which shows whether the U.S. premium extends beyond comparison with the cheapest market – and (iii) with an elasticity-aligned (EA) benchmark – which uses the fitted foreign relationship to predict a price at U.S. demand responsiveness and country characteristics. Intuitively, the EA comparison assesses how U.S. prices differ from a foreign benchmark that accounts for these observed differences across markets.

The price panel covers 30 potential non-U.S. OECD comparators. For each molecule–period, we require positive U.S. sales and standard units in every included quarter and exclude the observation if any product is flagged as generic. Foreign comparators must satisfy the same restrictions, have OECD membership in every included quarter, and meet the annual income threshold. Twenty-three foreign countries qualify in at least one comparison. Let $\mathcal{C}_{ga}^{\text{pMFN}}$ denote the set of eligible foreign comparator countries. For each eligible set, the pseudo-MFN price and the equally weighted foreign mean are

$$\bar{P}_{ga}^{\text{pMFN}} = \min_{c \in \mathcal{C}_{ga}^{\text{pMFN}}} \bar{P}_{gca}, \quad \bar{P}_{ga}^{\text{COMP}} = \frac{1}{|\mathcal{C}_{ga}^{\text{pMFN}}|} \sum_{c \in \mathcal{C}_{ga}^{\text{pMFN}}} \bar{P}_{gca},$$

where $\bar{P}_{gca}$ is a unit-weighted molecule price, with weights proportional to the number of standard units sold, giving greater weight to higher-volume observations. We construct EA by first estimating how prices in eligible foreign markets vary with demand responsiveness and country characteristics. We then apply this fitted relationship to U.S. elasticity estimates and country characteristics to obtain a reference price for each molecule–period³⁰. Observed U.S. prices are then compared with the resulting reference. Let $\hat{\chi}_{gca}^{\text{ADL}} = -\hat{\epsilon}_{gca}^{\text{own,ADL}} > 0$ denote the magnitude of the estimated own-price response. Using the eligible foreign observations, we estimate

$$\log \bar{P}_{gca} = \alpha_{ga} + W'_{ca} \delta + \psi \left(\frac{1}{\hat{\chi}_{gca}^{\text{ADL}}} \right) + \zeta_{gca}^{\text{EA}}.$$

The molecule–period fixed effects α_{ga} capture conditions specific to each molecule and period that are shared across countries. The country controls W_{ca} contain log PPP GDP per capita, log IMF producer price index, and log PPP-adjusted health expenditure per capita.

³⁰The benchmark regression uses only foreign observations. Unlike the preceding regressions, which examine the association between prices and elasticities, as explained, this regression is used to construct a price benchmark.

Motivated by the association between higher prices and less responsive demand observed among branded medicines, we constrain ψ to be weakly increasing. This incorporates the inverse-elasticity ordering into the EA benchmark so that holding the other regressors fixed, less responsive demand predicts the same or a higher price. The elasticity inputs are the country–molecule–year product-demand responses estimated by the paper’s joint all-product model, including those for the United States. U.S. observations are excluded only from the EA price regression, which estimates the foreign benchmark relationship. We obtain the U.S. EA benchmark by inserting the previously estimated U.S. elasticity and observed country characteristics into this fitted relationship, $\bar{P}_{g,US,a}^{EA} = \exp\left\{\hat{\alpha}_{ga} + W'_{US,a}\hat{\delta} + \hat{\psi}\left(\frac{1}{\hat{\lambda}_{g,US,a}^{ADL}}\right)\right\}$. The idea is to convert the fitted log price into price units and provide a reference for assessing how far the observed U.S. price lies from that relationship.

The U.S.–pseudo-MFN gap can then be written as

$$\log \frac{\bar{P}_{g,US,a}}{\bar{P}_{ga}^{pMFN}} = \log \frac{\bar{P}_{g,US,a}}{\bar{P}_{g,US,a}^{EA}} + \log \frac{\bar{P}_{g,US,a}^{EA}}{\bar{P}_{ga}^{pMFN}},$$

where the first term compares the observed U.S. price with EA, and the second compares EA with the lowest eligible foreign price. While following by simple algebraic manipulation³¹, the mentioned decomposition, helps us better understanding whether a large U.S. premium over the foreign minimum also corresponds to a large deviation from the fitted foreign relationship at U.S. characteristics. A U.S. price can be close to that benchmark even when it is far above the foreign minimum. Neither component, however, measures how much of the gap is explained by elasticities alone, since EA also incorporates country characteristics and molecule–period effects.

Table 4. Historical pseudo-MFN and elasticity-aligned price comparisons

Price comparison	Median ratio	Observations with ratio > 1	Share (%)
U.S. price / pseudo-MFN price	6.95	112	96.6
U.S. price / EA benchmark	1.62	64	55.2
EA benchmark / pseudo-MFN price	4.63	116	100.0
U.S. price / comparator mean	1.81	78	67.2

Notes: All comparisons use the same 116 U.S. molecule–period observations covering 17 molecules, each receiving equal weight. The pseudo-MFN price is the lowest eligible foreign price. The comparator mean equally weights eligible foreign countries. EA is the fitted foreign benchmark evaluated at U.S. elasticity estimates and country characteristics. Prices are constant-exchange ex-manufacturer sales per standard unit. Median ratios are calculated separately and need not satisfy the multiplicative identity that holds for each observation. Estimation uncertainty is not propagated through these comparisons.

Table 4 reports comparisons for 116 U.S. molecule–period observations covering 17 molecules, with each observation receiving equal weight. The choice of foreign reference substantially changes the comparison. U.S. prices exceed the foreign minimum in 96.6 percent of observations, with a median ratio of 6.95. Against the foreign mean, the median ratio falls to 1.81, and the share of U.S. prices above the reference falls to 67.2 percent.

³¹It follows by inserting the EA benchmark as an intermediate reference between the U.S. price and the foreign minimum. In logs, the U.S.–EA difference and the EA–minimum difference add up exactly to the original U.S.–minimum gap, because the two log EA terms cancel.

The EA comparison also gives a different picture from the foreign minimum. The median U.S.-to-EA ratio is 1.62, with U.S. prices above EA in 55.2 percent of observations. EA exceeds the foreign minimum in every comparison, with a median ratio of 4.63. Thus, a large premium over the lowest foreign price need not imply a similarly large deviation from the fitted foreign relationship evaluated at U.S. characteristics.

We believe that the exercise offers an important lesson from a policy perspective. Taken together with the preceding results, it suggests that selecting countries by PPP GDP per capita should not be taken as evidence that their demand is similarly price responsive. Policymakers concerned with demand comparability can therefore complement the income criterion with directly estimated elasticities when assessing foreign comparators.

Furthermore, for policy evaluation, the choice of reference price also matters substantially. In our historical sample, the foreign minimum produces a much larger measured U.S. premium than either the foreign mean or EA. Reporting these comparisons together helps evaluators assess whether a large premium principally reflects comparison with the cheapest eligible market or also persists relative to average foreign prices and a benchmark accounting for U.S. demand responsiveness and country characteristics.

Several features limit the scope of this comparison. Matching is by molecule and period, so U.S. and foreign prices can average different products or presentations. The prices are ex-manufacturer unit values, rather than observed net transaction prices or patient payments. Finally, U.S. health expenditure per capita exceeds the range observed in the foreign estimation sample. The EA benchmark therefore assumes that the fitted relationship between log prices and log health expenditure continues to apply at the higher U.S. spending levels.

6 Conclusions

The international pricing of pharmaceuticals presents a fundamental economic challenge arising from the need to reconcile access to medicines with the recovery of the substantial fixed costs associated with pharmaceutical innovation (Danzon and Towse, 2003). Within this context, differential pricing has long been proposed as a second-best solution to recover common R&D costs while limiting the welfare losses associated with higher prices (Danzon and Towse, 2003; Danzon, 2018). The underlying inverse-elasticity principle suggests that, conditional on a given revenue requirement, markets with less price-responsive demand should bear higher proportional price–cost margins. Nevertheless, despite its theoretical importance and longstanding policy relevance, empirical evidence on the extent to which international pharmaceutical prices follow this principle remains limited, partly due to the methodological difficulties associated with estimating comparable demand elasticities across heterogeneous institutional and therapeutic environments.

Structural demand models commonly employed in pharmaceutical demand estimation (Nevo, 2000) require assumptions regarding market definitions, consumer preferences, and substitution patterns that may not be equally appropriate across countries and therapeutic settings. Moreover, there is no theoretical reinsurance – to the best of our knowledge – on whether accommodating extensive heterogeneity within a common structural framework introduces identification difficulties when the number of parameters is large relative to the number of available observations or instruments. These challenges have contributed to the reliance on simplified measures of demand responsiveness in parts of the international differential-pricing literature, particularly GDP per capita as a proxy for price elasticity (Danzon and Towse, 2003; Danzon and Furukawa, 2003). Although income-based differential pricing has an economic rationale grounded in differences in purchasing power and ability to pay, income and price elasticity are conceptually distinct. Demand responsiveness also depends on substitution possibilities, insurance coverage, reimbursement arrangements, and other institutional characteristics that may vary substantially across countries with similar income levels.

Consequently, GDP per capita does not necessarily provide a reliable approximation of the demand elasticity relevant to the inverse-elasticity principle and it leaves the actual heterogeneity in price responsiveness largely unmeasured. The present study addresses these limitations by developing a comprehensive empirical framework to estimate internationally comparable pharmaceutical demand elasticities while accommodating heterogeneity across countries, medicines, and time. Building on these estimates, we examine whether observed international pharmaceutical price differences are systematically associated with demand responsiveness in a manner consistent with the inverse-elasticity benchmark.

Drawing on a large-scale pharmaceutical dataset covering 47,589 products, 171 molecules, and 63 countries over the period 2008–2020, we employ a Double Machine Learning (DML) framework to estimate heterogeneous own- and rival-price elasticities of pharmaceutical demand. By adapting the high-dimensional dynamic-panel framework of Semenova et al. (2023), we account for a broad set of observed demand determinants while allowing price responsiveness to vary across countries, molecules, and time. We adopt regularized prediction that accommodates a broad set of demand controls without requiring a separate consumer-choice model for each country. In doing so, the present study contributes to the literature on pharmaceutical demand and differential pricing by moving beyond income-based approximations of demand responsiveness and providing directly estimated, internationally comparable elasticities. From a methodological perspective, we further contribute to the empirical industrial organisation literature by establishing conditions under which a sufficiently rich DML price-response representation can approximate the price-dependent component and implied elasticities of fixed, identified logit and random-coefficients logit models. We also demonstrate the identification difficulties that may arise when structural demand models become high-dimensional relative to the available observations or instruments. Together, these results provide a theoretical reason for employing flexible semiparametric approaches as a complementary empirical strategy to established structural demand models particularly in settings characterized by substantial heterogeneity and uncertain model specification (Chernozhukov et al., 2018; Chernozhukov et al., 2022).

The empirical findings reveal substantial heterogeneity in pharmaceutical demand responsiveness across countries and medicines. Country-level mean own-price elasticities range from approximately -0.73 to -0.26, indicating that pharmaceutical demand is relatively price inelastic on average, although the degree of responsiveness differs considerably across national markets. Importantly, these differences cannot be adequately captured by GDP per capita, whose relationship with estimated demand responsiveness is modest and sensitive to the estimation method and product composition. This finding challenges the common practice of employing national income as a proxy for pharmaceutical demand elasticity in international pricing comparisons.

A second important finding concerns the relationship between demand responsiveness and pharmaceutical prices. Our results robustly reveal that higher branded-drug prices are systematically associated with less price-responsive demand, in line with the ordering implied by the inverse-elasticity principle. Conversely, generic medicines do not exhibit an equally stable relationship. One possible interpretation of this heterogeneity is that branded and generic pharmaceutical markets operate under different competitive and regulatory conditions. In particular, manufacturers of branded medicines may have greater discretion in setting prices, whereas generic prices are more directly constrained by competition, procurement systems, and regulatory arrangements.

These findings convey important implications for the ongoing debate on international reference pricing and the distribution of pharmaceutical innovation costs across countries. This debate has gained renewed policy relevance following the United States’ 2025 MFN pricing initiative, which proposed benchmarking certain branded pharmaceutical prices against the lowest prices observed in economically comparable OECD countries. Although such policies aim to reduce pharmaceutical expenditures and address concerns about unequal international contributions to innovation costs, their reliance on GDP per capita as a criterion for se-

lecting comparator countries does not necessarily ensure comparability in demand conditions. Our historical pseudo-MFN exercise further illustrates this point. Our estimates show that accounting for demand responsiveness and observable country characteristics yields substantially different international price comparisons than those based on the lowest eligible foreign price. From a policy perspective, these results emphasise the importance of adopting a more nuanced approach to international pharmaceutical price comparisons. In particular, the selection of reference countries based exclusively on income thresholds may overlook economically relevant differences in demand responsiveness, potentially leading to benchmarks that are not representative of the domestic pharmaceutical market. Incorporating directly estimated demand elasticities alongside income and other country characteristics could therefore provide additional information for the design and evaluation of international reference-pricing policies.

As any other, the present study is subject to several limitations that should be acknowledged. First, despite the flexibility offered by DML, the causal interpretation of the estimated demand elasticities relies on the maintained assumption of sequential conditional exogeneity. Although our specification accounts for extensive observed product and market characteristics, as well as historical prices and sales, unobserved demand shocks correlated with contemporaneous prices may still introduce endogeneity. The complementary instrumental-variable exercise provides supporting evidence, but does not eliminate this concern in every country–molecule market. Second, our empirical framework estimates own-price responses and responses to an aggregate rival-price index rather than a complete system of product-specific substitution patterns. Consequently, it does not directly recover the full demand structure and supply-side parameters required to estimate pharmaceutical markups or simulate the welfare consequences of alternative pricing policies. Moreover, while the country, molecule, and year components accommodate substantial heterogeneity, the baseline additive specification may not fully capture more complex interactions between national institutional arrangements and therapeutic characteristics. Third, our international comparisons rely on observed ex-manufacturer unit values, which may differ from net transaction prices due to confidential rebates, discounts, and reimbursement agreements. These differences, together with variation in product availability and presentation across markets, may affect international price comparisons.

Appendix

Contents of the Appendix

A	Selected Related Literature and Positioning of the Paper	2
B	Inference in DML: the Ridge Correction and Its Relation to CLIME	9
B.1	CLIME and the One-Step Bias Correction	10
B.2	The Ridge Adjustment and Its Implications	12
C	Pooled DML and IV Estimates of Price Elasticities	12
D	Limitations of DML	16
E	Existence of a Dictionary Approximation for Logit-Based Demand	17
F	High-Dimensional RCL: Rank Failure and Regularization	21
F.1	An illustrative regularized structural route	24
G	Simulations	26
G.1	Simulation results: Proposition 2.1	26
G.2	Simulation results: Proposition 2.2	28
H	Further validity checks on estimated elasticities	30
H.1	Permutation placebo exercise	30
H.2	Price elasticities by molecule	32
H.3	Additional elasticity-related descriptives	35
I	Details of the Structural RCL Benchmark	38
J	The inverse elasticity rule with continuous choice	39
J.1	Intensive margin	40
J.2	GDP (per capita) as a measure of elasticity	43
J.3	Efficiency of Ramsey pricing across countries	44
K	Interpreting Price–Elasticity Associations with Heterogeneous Marginal Costs	47
K.1	Prices, costs, and the inverse-elasticity benchmark	48
K.2	Price residualization and structural cost diagnostics	50

A Selected Related Literature and Positioning of the Paper

Table A.1 shows that the paper lies at the intersection of several literatures.

Relative to international price–income comparisons, we argue that income-based proxies may inadequately capture pharmaceutical demand responsiveness, which also reflects substitution possibilities and institutional arrangements.

Our paper also draws from the discrete choice models literature. While structural pharmaceutical demand models derive price elasticities from an explicit model of consumer choice, we estimate how product sales respond to own prices and average rival prices, using DML to flexibly account for a broad set of researcher-specified controls. As highlighted in the main text, DML allows the researcher to start from a broad set of candidate demand controls without determining in advance their relative importance for predicting prices and sales. This provides a common framework for comparing demand responsiveness across many products, therapeutic markets, and countries without specifying a complete consumer-choice model for each setting. The trade-off is that we estimate these targeted price responses without recovering the full product-by-product substitution system or the supply-side structure needed to simulate the market-wide effects of pricing policies.

Empirically, relative to these two strands of literature, moreover, we estimate the elasticities of products in 63 countries rather than restricting to a small subset of molecules or countries as generally done. Some examples follow. The closest cross-country evidence using manufacturer prices finds negative and heterogeneous molecule-level elasticities across seven countries (Danzon and Chao, 2000a). Dubois et al. (2021) performs product-level 2SLS estimates for 40 essential molecules in seven low- and middle-income settings, finding elasticities that are close to one in absolute value and vary across procurement channels. Structural estimates for the narrowly defined French anti-ulcer market are considerably more elastic (Dubois and Lasio, 2018), whereas U.S. studies of patient copayments find inelastic responses that vary across drug classes (Gatwood et al., 2014; Yeung et al., 2018). More recently, Gugler et al. (2026) estimate a nested-logit demand model for platelet aggregation inhibitors in Austria, finding mean own-price elasticities of -1.74 in the regulated, reimbursed segment and -2.34 in the unregulated segment. Table A.1 provides further details.

The paper also contributes methodologically by formally relating DML-based elasticity estimation to the RCL models widely used in structural demand analysis. We establish conditions under which a sufficiently rich DML dictionary can approximate the price responses generated by a fixed, identified RCL model and, under additional regularity conditions, its implied elasticities. We also characterize how an unrestricted RCL specification with too many demand parameters relative to the available observations or instruments can fail to uniquely identify price coefficients, leaving different elasticities consistent with the same demand moments. Accordingly, Table A.1 positions our contribution as a formal link between orthogonal-learning methods and BLP/RCL demand estimation, highlighting how, in a high-dimensional sparse setting, GMM identification adopted in RCL fails to affect RCL elasticities as described in the main text.

Finally, we connect our elasticity analysis to the quantity-based framework commonly used in Ramsey pricing, which asks how costs can be recovered while limiting reductions in medicine consumption. Continuous-choice models make explicit how the quantities consumed adjust to prices and expenditure. Namely, the first model in Appendix J formally shows why income alone need not predict price responsiveness when spending patterns and insurance coverage differ. A separate Ramsey model shows that the welfare consequences of pricing depend on both demand elasticity and countries’ financing constraints. These results help explain why we estimate elasticities directly rather than infer them from GDP, and why those estimates alone cannot determine which countries bear greater welfare losses.

Table A.1. Selected related literature and relation to the present paper

Literature and studies	Setting and question	Approach and findings	Relation to this paper
<i>A. Differential pricing and international price comparisons</i>			
Ramsey pricing and pharmaceutical R&D (Ramsey, 1927; Baumol and Bradford, 1970; Danzon, 1997; Danzon and Towse, 2003)	How prices can finance a revenue requirement or common costs, including pharmaceutical R&D, while limiting losses from reduced consumption.	Develop the Ramsey benchmark and its pharmaceutical applications. With independent demands across markets, proportional price–cost margins are inversely related to demand elasticities, for a given financing requirement.	We estimate demand responsiveness and examine whether higher prices are associated with less elastic demand.
Insurance and international cost sharing (Barros and Martinez-Giralt, 2008)	How differences in insurance coverage affect the allocation of pharmaceutical R&D costs across countries.	Incorporates insurance into Ramsey pricing. Optimal prices depend on insurance arrangements and financing distortions as well as demand elasticities; the simple inverse-elasticity rule requires modification.	Appendix J uses continuous quantities to connect demand responsiveness to this cost-recovery problem. Its branded–generic model explains why income alone need not rank markets by elasticity when spending patterns and insurance differ. A separate Ramsey model shows why welfare comparisons also depend on financing constraints.
Differential-pricing policy and market segmentation (Maskus, 2001; Yadav, 2010a)	International access to medicines and the conditions under which prices can differ across markets.	Review the economic and policy conditions for international price differentiation, including market segmentation and parallel imports.	We provide estimates of demand responsiveness across countries and medicines to inform these pricing questions. These estimates allow us to assess whether higher prices occur where medicine use is less sensitive to price, as suggested by the inverse-elasticity benchmark for allocating R&D costs across markets.

Continued on next page

Table A.1 — Continued

Literature and studies	Setting and question	Approach and findings	Relation to this paper
International prices, income, and affordability (Sherer and Watal, 2002; Danzon and Furukawa, 2003; Lichtenberg, 2011; Morel et al., 2011; Petrou and Vondros, 2016)	How pharmaceutical prices and affordability differ across countries with different income levels.	Compare prices or price indexes across countries, sometimes adjusting for income. Findings on the price–income relationship differ across samples and methods. Such comparisons do not identify product-level own- and cross-price elasticities.	We estimate quantity responses to price and then assess how well GDP per capita and income groups summarize differences in demand responsiveness. This allows us to examine whether richer countries actually have less price-sensitive demand and, indirectly, empirically assess whether GDP per-capita are valid proxies for demand elasticity.
Elasticities inferred from procurement markups (Srivastava and McGuire, 2014)	Government procurement of 48 branded medicines in 16 countries, using price data without corresponding quantities.	Uses generic prices as proxies for marginal costs and an inverse-elasticity pricing formula to back out demand elasticities. The resulting elasticities are weakly related to national income.	Our data include both prices and quantities. We estimate elasticities without imposing Ramsey pricing, then examine whether observed prices are consistent with its inverse-elasticity ordering.

Continued on next page

Table A.1 — Continued

Literature and studies	Setting and question	Approach and findings	Relation to this paper
<i>B. Evidence on demand responsiveness</i>			
Cost sharing and medical care (Manning et al., 1987; Chandra et al., 2014; Brot-Goldberg et al., 2017)	Medical spending and service use under different copayments, coinsurance rates, or deductibles in U.S. insurance systems.	Use experimental or policy-induced variation in patient payments to estimate changes in utilization and spending. These studies cover medical care more broadly, with some reporting separate pharmaceutical outcomes.	They establish the importance of the price paid by the patient. Our estimates concern product sales and ex-manufacturer prices, so their elasticity magnitudes are contextual comparisons rather than estimates of the same response.
Drug demand and insurance design (Goldman et al., 2004; Landsman et al., 2005; Einav et al., 2018)	Prescription use under employer insurance or Medicare Part D, with differences in cost sharing across drugs and therapeutic classes.	Estimate how drug use responds to patient payments and benefit design. Einav et al. also relate drug-specific elasticities to the cost sharing chosen by private plans.	We extend the comparison of demand responsiveness across medicines to a multinational setting, using aggregate quantities and manufacturer prices. Insurance remains relevant to interpreting differences across countries, but we do not identify its separate causal effect.
Simultaneous changes in own and related-drug prices (Yeung et al., 2018)	A value-based formulary that raises copayments for some medicines and lowers them for others.	Shows how simultaneous changes in a drug's price and the prices of substitutes or complements produce a composite utilization response. The study estimates responses to the formulary change and its welfare implications.	We include own price and a same-molecule rival-price index jointly, so the own-price estimate does not absorb all variation in competitors' prices. The rival-price response summarizes an aggregate margin; it does not identify each substitution or complementarity relationship.
Cross-country molecule-level demand (Danzon and Chao, 2000a)	Manufacturer prices and molecule-level quantities in retail-pharmacy data for seven countries.	Combines international price comparisons with two-stage least-squares quantity equations. The demand estimates relate market-level quantities to manufacturer prices, rather than patient copayments.	This is a close empirical predecessor. We use product-level quantities in 63 countries, allow own-price responsiveness to vary by country, molecule, and year, and jointly estimate the response to a within-molecule rival-price index.

Continued on next page

Table A.1 — Continued

Literature and studies	Setting and question	Approach and findings	Relation to this paper
<i>C. Pharmaceutical regulation, procurement, and reference pricing</i>			
Structural demand under price regulation (Dubois and Lasio, 2018)	French anti-ulcer medicines, with U.S. and German data supporting restrictions on costs across markets.	Combines random-coefficients logit (RCL) demand with a pricing model to recover margins and assess binding price constraints. Simulates free pricing and external reference pricing.	We use a common semiparametric demand specification across many therapeutic markets and countries. This broader coverage comes with less structural detail: the global analysis estimates own- and rival-price responses without recovering country-specific costs or regulatory counterfactuals.
Procurement arrangements and buyer power (Waning et al., 2009; Dubois et al., 2021)	International antiretroviral procurement transactions; and sales of 40 molecules across seven low- and middle-income countries.	Waning et al. compare transaction prices across purchasing and pricing arrangements. Dubois et al. model buyer power and estimate lower prices under centralized procurement, with smaller savings when suppliers are more concentrated.	These studies show why prices can differ even when demand responsiveness is similar. Procurement and supplier concentration help interpret our price-elasticity patterns, especially for generics; our estimates do not separately identify these mechanisms.
Bargaining and international reference pricing (Dubois et al., 2022)	U.S. and Canadian hospital pharmaceuticals within selected therapeutic classes.	Estimates country-specific RCL demand, models U.S. pricing and Canadian bargaining, and simulates equilibrium responses to international reference pricing.	Provides our main structural benchmark for elasticity magnitudes. We compare U.S. estimates using aligned therapeutic classes and branded/generic groups, while recognizing differences in channels and periods. Our rival-price index does not reproduce the study's full substitution matrix.
Reference pricing, launch decisions, and transparency (Maini and Pammolli, 2023; Riccaboni et al., 2022)	How links between national prices affect firms' launch decisions and countries' access to medicines.	Maini and Pammolli estimate a structural entry model using moment inequalities and quantify reference-pricing delays. Riccaboni et al. use an agent-based model to examine net-price transparency, confidential discounts, and access.	Our historical MFN-style exercise compares observed foreign prices with a U.S. price benchmark informed by estimated elasticities.

Continued on next page

Table A.1 — Continued

Literature and studies	Setting and question	Approach and findings	Relation to this paper
<i>D. Structural demand and high-dimensional estimation</i>			
BLP and RCL demand systems (Berry et al., 1995; Nevo, 2000)	Consumer choice among differentiated products, with heterogeneous preferences and potentially endogenous prices.	Estimate preference parameters from market shares and instruments. The fitted system delivers own- and cross-price derivatives and, with supply assumptions, supports markup and policy analysis.	Proposition 2.1 establishes conditions under which a rich dictionary can approximate the price response of a fixed, identified logit or RCL model. Approximating its elasticities requires additional derivative conditions discussed in the main text.
Orthogonal estimation and dynamic panels (Chernozhukov et al., 2018; Semenova et al., 2023)	Estimation of effects with many controls; heterogeneous effects in panels with weak dependence.	Develop orthogonal estimation with learned nuisance functions. Semenova et al. extend the framework to dynamic panels and sketch a many-product own- and rival-price specification in Section 7, outside their main theoretical results.	We adapt this many-product specification to pharmaceutical demand across countries. Our empirical contribution is the common measurement of heterogeneous price responses; the RCL approximation and rank results clarify how that object relates to structural demand estimation.
Sparse moment models and structural demand (Fan and Liao, 2014; Sawada and Kawaguchi, 2020)	Models with endogenous regressors or many product characteristics relative to the available data.	Fan and Liao develop penalized focused GMM. Sawada and Kawaguchi study regularized and debiased RCL estimation, restricting how high-dimensional attributes enter preference heterogeneity.	Our rank result concerns <i>unrestricted</i> RCL moments. These studies show how sparsity and suitable identification conditions can support estimation in high dimensions. Notice that our regularized RCL exercise in the Appendix is intended as an illustrative simulation benchmark, not a new general estimator for which further analyses would be needed.

Continued on next page

Table A.1 — Continued

Literature and studies	Setting and question	Approach and findings	Relation to this paper
Orthogonal estimation within BLP (Jin, 2026)	BLP models with many potential demand controls and a small number of target parameters.	Develops Neyman-orthogonal estimation and inference for structural parameters, including price sensitivity, with high-dimensional nuisance functions and an approximate-sparsity implementation.	Our contribution concerns approximation of a fixed structural price response and the rank failure of an unrestricted RCL specification. Proposition 2.2 characterizes non-uniqueness of the linear parameters; elasticities are non-unique when an unidentified direction changes a price-response parameter. Our empirical DML model estimates demand responses directly.
Present paper	IQVIA MIDAS data on 47,589 products, 171 molecules, and 63 countries, 2008Q4–2020Q3. Quantities are standard units; prices are ex-manufacturer prices at fixed exchange rates.	Estimates own- and rival-price elasticities using DML, with additive country, molecule, and year heterogeneity. Compares U.S. estimates with structural benchmarks and relates international prices to estimated demand responsiveness.	Combines broad international evidence on pharmaceutical demand with a formal analysis of its relation to RCL. Continuous-choice models explain why income alone may poorly summarize price responsiveness and why elasticity alone does not determine the welfare implications of differential pricing.

Notes: The table groups studies by their role in this paper. DML denotes double machine learning; BLP denotes Berry–Levinsohn–Pakes; RCL denotes random-coefficients logit; GMM denotes generalized method of moments; MFN denotes most-favored-nation pricing. Our rival-price index is the log of the arithmetic mean of other products' prices within the same molecule, country, and quarter. Its elasticity is not a complete set of pairwise cross-price elasticities. Estimated responses vary additively across countries, molecules, and years and are shared by products within each such cell. The DML estimation sample starts in 2009Q4 because of lagged controls. A causal interpretation requires sequential conditional exogeneity given the chosen information set; machine learning does not remove this requirement.

B Inference in DML: the Ridge Correction and Its Relation to CLIME

The DML implementation uses regularization in both estimation stages to reduce overfitting. Regularization encourages simpler models, for example by shrinking coefficients toward zero or excluding variables. However, it can also weaken genuine relationships: even across repeated samples, estimates may then be systematically displaced from the true effects. This displacement is called regularization bias.

In the first stage, ML predicts quantities and prices from observed controls to estimate $\ell_0(\cdot)$ and $m_0(\cdot)$. For example, high past sales may signal persistent demand that supports both higher current prices and higher quantities. Accounting for past sales helps separate this shared influence from the effect of price. If regularization weakens this adjustment, residual price and quantity movements may still partly reflect past demand, potentially distorting the estimated price response. Under the assumptions stated in the main text, DML reduces sensitivity to these nuisance-estimation errors.

In the second stage, LASSO directly shrinks the coefficients governing price responsiveness, potentially suppressing genuine differences across countries, molecules, and years. This creates an additional bias that neither first-stage residualization nor the mentioned DML assumptions do remove. Conventional confidence intervals describe random variation around an estimate and do not automatically account for a systematic shift caused by shrinkage. Consequently, an interval can be narrow yet miss the true elasticity because it is centered on a biased estimate. This motivates the additional correction discussed below.

To address this second-stage regularization bias, Semenova et al. (2023) use CLIME (Constrained ℓ_1 Minimization for Inverse Matrix Estimation). The idea is to exploit the residual correlation of the residual quantity to the regressors. Indeed, after the LASSO fit, quantity residuals may remain associated with the second-stage regressors because the penalty has shrunk (but not eliminated) their coefficients. The score $\hat{s}_L$ summarizes this remaining association. Because the regressors are correlated, the association with one regressor can reflect estimation errors in several coefficients. CLIME constructs correction weights that account for these correlations, translating the score into an adjustment for each coefficient: $\hat{b}^C = \hat{b}^L + \hat{\Theta}^C \hat{s}_L$. The weighting matrix $\hat{\Theta}^C$ is a sparse approximate inverse of $\hat{\Gamma}$, the Gram matrix recording the joint variation of the residualized regressors. Under the required conditions, including approximate sparsity of the population inverse Gram matrix, the correction makes the remaining shrinkage bias negligible relative to sampling uncertainty, allowing valid confidence intervals for the elasticities.

Our specification allows price responses to differ across countries, molecules, and years because these comparisons are central to our research question.^{B.1} When some of these regressors move closely together, the data provide limited information about their separate contributions and attempting to fully correct LASSO shrinkage in such cases can amplify sampling noise.

We therefore use the ridge inverse $\hat{\Theta}_R = (\hat{\Gamma} + \lambda_R I_d)^{-1}$ to stabilize the one-step correction. Ridge moderates adjustments for combinations of coefficients that the data distinguish poorly, limiting the sensitivity of the correction to small sampling fluctuations. This approach does not require sparsity of the inverse Gram matrix with the trade-off that some LASSO shrinkage may remain uncorrected. Valid inference therefore still requires the remaining bias to be negligible relative to sampling uncertainty. Throughout the paper, ADL (adjusted debiased LASSO) denotes the ridge-adjusted one-step LASSO estimator: $\hat{b}^{\text{ADL}} \equiv \hat{b}^R = \hat{b}^L + (\hat{\Gamma} + \lambda_R I_d)^{-1} \hat{s}_L$ whose terms are explained in what follows.

^{B.1}As explained in Section 2.1, the second-stage regressors multiply residualized own and rival prices by the indicators in K_r . These price-indicator interactions form $\hat{V}_r$.

B.1 CLIME and the One-Step Bias Correction

To address the shrinkage induced by second-stage LASSO, Semenova et al. (2023) construct a one-step coefficient correction using CLIME (Constrained ℓ_1 Minimization for Inverse Matrix Estimation), developed by Cai et al. (2011). CLIME estimates the weights that translate the association remaining between regressors and unexplained quantities into adjustments to the LASSO coefficients.

The procedure is as follows. Let n denote the number of observations and d the number of residualized technical treatments. Define their sample Gram matrix and the score evaluated at the LASSO estimate as

$$\hat{\Gamma} = \frac{1}{n} \sum_{r=1}^n \hat{V}_r \hat{V}_r', \quad \hat{s}_L = \frac{1}{n} \sum_{r=1}^n \hat{V}_r \left(\hat{y}_r^\circ - \hat{V}_r' \hat{b}^L \right).$$

The Gram matrix records the joint variation of the residualized regressors through their sample second moments. The score measures how each regressor remains associated with the quantity variation left unexplained by the LASSO fit $\left(\hat{y}_r^\circ - \hat{V}_r' \hat{b}^L \right)$. This association does not necessarily indicate an error in that regressor's coefficient alone. For example, if the own- and rival-price regressors move together, an incorrectly estimated rival-price effect can leave unexplained quantity variation that is also associated with own price. The correction therefore needs to account for these joint movements to determine how much to adjust each coefficient.

To intuitively understand the CLIME correction, let us start from a generic adjustment to the estimated coefficient when regularization techniques are adopted, i.e. $\hat{b}^L$: $\hat{b}(\hat{\Theta}) = \hat{b}^L + \hat{\Theta} \hat{s}_L$, where $\hat{\Theta}$ acts as an approximate inverse of the Gram matrix. It is easy to see that if $\hat{\Gamma}$ is invertible, its exact inverse converts the score into the adjustment that recovers OLS using the same outcome and regressors, without an additional intercept. Indeed, the least-squares normal equations give $0 = \hat{s}_L - \hat{\Gamma} (\hat{b}^{OLS} - \hat{b}^L)$, $\hat{b}^{OLS} - \hat{b}^L = \hat{\Gamma}^{-1} \hat{s}_L$, such that the correction reduces the coefficients to the coefficients of an OLS. However, the sample Gram matrix can be singular or nearly singular, making direct inversion unavailable or sensitive to sampling noise. As a result, CLIME estimates a sparse approximate inverse, i.e., more formally – defining the operations $\|A\|_{\max} = \max_{j,k} |A_{jk}|$ and $\|A\|_1 = \sum_{j,k} |A_{jk}|$ – it first solves

$$\tilde{\Theta}^C = \arg \min_{\Theta \in \mathbb{R}^{d \times d}} \|\Theta\|_1 \quad \text{subject to} \quad \left\| \hat{\Gamma} \Theta - I_d \right\|_{\max} \leq \lambda_{\Gamma}, \quad (\text{B.1})$$

and then symmetrizes the solution to obtain $\hat{\Theta}^C$, retaining – among each transposed pair of entries of the matrix Θ – the entry with the smaller absolute value. The resulting coefficient estimator is $\hat{b}^C = \hat{b}^L + \hat{\Theta}^C \hat{s}_L$.

Intuitively, the correction aims to offset the difference between the LASSO coefficients $\hat{b}^L$ and the true coefficients b_0 . Ideally, we would add $b_0 - \hat{b}^L$ to the LASSO estimate, but this difference is unknown. The residual score provides information about it. Indeed – temporarily setting aside nuisance-estimation error – substituting the second-stage demand equation into the definition of the score, $\hat{s}_L = \hat{\Gamma}(b_0 - \hat{b}^L) + \frac{1}{n} \sum_{r=1}^n \hat{V}_r u_r$. Expanding, we get:

$$\hat{s}_{L,j} = \underbrace{\hat{\Gamma}_{jj} (b_{0,j} - \hat{b}_j^L)}_{\text{contribution from coefficient } j} + \underbrace{\sum_{k \neq j} \hat{\Gamma}_{jk} (b_{0,k} - \hat{b}_k^L)}_{\text{contributions from other coefficients}} + \underbrace{\frac{1}{n} \sum_{r=1}^n \hat{V}_{r,j} u_r}_{\text{sampling noise}}.$$

The score, however, as seen contains the coefficient errors of several coefficients mixed together through $\hat{\Gamma}$, plus sampling noise arising from the demand shocks u_r . Multiplying the score $\hat{s}_L$ by an inverse of $\hat{\Gamma}$ would

undo this mixing ^{B.2}. The inverse combines the scores so that the error in coefficient j receives weight one, while errors in all other coefficients receive weight zero. This cancellation follows from $\hat{\Gamma}^{-1}\hat{\Gamma} = I_d$. It separates the coefficient errors but does not eliminate the sampling noise in the score. CLIME seeks to achieve this approximately, in case an inverse of Γ is not available, by requiring $\hat{\Theta}^C\hat{\Gamma} \approx I_d$, where I_d is the $d \times d$ identity matrix and, as a reminder, $\hat{\Theta}$ denotes the solution of the constrained optimization problem, *before* symmetrization. To understand this requirement, consider the adjustment to coefficient j . The (j, j) entry of this product determines the weight placed on $b_{0,j} - \hat{b}_j^L$, the adjustment that coefficient j needs. An entry close to one therefore retains approximately the full required adjustment. Each off-diagonal entry (j, k) determines how much the error in another coefficient k enters that same adjustment; keeping these entries close to zero limits this mixing. Importantly, these statements concern the entries of the product $\hat{\Theta}^C\hat{\Gamma}$, not those of either matrix separately.

The tolerance λ_Γ controls the permitted departure from the identity matrix. Among matrices satisfying this constraint, the ℓ_1 objective favors relatively few substantial correction weights. The solution is then symmetrized to obtain $\hat{\Theta}^C$. For any correction matrix $\hat{\Theta}$, substituting the score into the one-step estimator gives

$$\hat{b}(\hat{\Theta}) - b_0 = \underbrace{(I_d - \hat{\Theta}\hat{\Gamma})(\hat{b}^L - b_0)}_{\text{remaining LASSO estimation error}} + \underbrace{\hat{\Theta} \frac{1}{n} \sum_{r=1}^n \hat{V}_r u_r}_{\text{transformed demand shocks}}. \quad (\text{B.2})$$

The first term shows how much of the original coefficient error, including LASSO shrinkage, remains after correction. Exact inversion would eliminate this term; approximate inversion leaves a remainder whose size also depends on the initial LASSO error. The second term shows why even exact inversion does not recover the true coefficients perfectly: sampling noise remains. With estimated nuisance functions, an additional remainder enters this decomposition. Orthogonality and cross-fitting help make it negligible under the required learning-rate and dependence conditions.

Assumptions 4.1–4.6 of Semenova et al. (2023) ensure that CLIME estimates the population inverse Gram matrix sufficiently accurately to control the first term in Equation (B.2). They require adequate residual treatment variation, accurate first-stage predictions, sparsity of both the coefficient vector and the population inverse Gram matrix, and restrictions on tails and dependence, including independence across panel units and sufficiently weak dependence over time. With the additional central-limit and variance-consistency conditions, these results support pointwise inference for linear combinations of b_0 whose weights have bounded ℓ_1 norm; consistency of their variance estimator also invokes Assumption 4.7 (Van de Geer et al., 2014; Chernozhukov et al., 2018; Semenova et al., 2023).

^{B.2}To see this, consider component j :

$$\begin{aligned} \left[\hat{\Gamma}^{-1}\hat{\Gamma}(b_0 - \hat{b}^L) \right]_j &= \sum_{k=1}^d \underbrace{\left[\sum_{h=1}^d (\hat{\Gamma}^{-1})_{jh} \hat{\Gamma}_{hk} \right]}_{(I_d)_{jk}} (b_{0,k} - \hat{b}_k^L) \\ &= \underbrace{1 \cdot (b_{0,j} - \hat{b}_j^L)}_{\text{own coefficient error retained}} + \underbrace{\sum_{k \neq j} 0 \cdot (b_{0,k} - \hat{b}_k^L)}_{\text{other coefficient errors cancelled}} \\ &= b_{0,j} - \hat{b}_j^L. \end{aligned}$$

where the 0 on the other coefficients arises because that multiplier is an off-diagonal entry of the identity matrix $\sum_{h=1}^d (\hat{\Gamma}^{-1})_{jh} \hat{\Gamma}_{hk} = (I_d)_{jk} = \begin{cases} 1, & k = j, \\ 0, & k \neq j. \end{cases}$

B.2 The Ridge Adjustment and Its Implications

Our specification contains 488 interactions of residualized own and rival prices with country, molecule, and year indicators, allowing elasticities to vary across these dimensions (Section 2.1). When regressors move closely together, separating their coefficient errors can require large correction weights that also amplify sampling noise. In particular, the inverse Gram matrix can require large weights to distinguish their separate effects. Each weight $\hat{\Theta}_{jk}$ determines how much score component k contributes to the adjustment of coefficient j . These weights also multiply the score’s sampling-noise component, $n^{-1} \sum_r \hat{V}_r u_r$. Consequently, small chance associations between residualized prices and demand shocks can produce large coefficient adjustments. As shown in Equation (B.1), CLIME minimizes the total absolute size of the correction weights while requiring $\hat{\Theta}$ to remain close to I_d , so that coefficient errors are approximately separated. This favors a sparse correction matrix, and its theoretical guarantees rely on approximate sparsity of the population inverse Gram matrix.

Ridge provides an alternative form of regularization: the idea is that adding $\lambda_R I_d$ to the Gram matrix before inversion limits noise amplification, where residual price variation provides little information about separate effects. It does so without requiring a sparse inverse, but can leave some LASSO shrinkage uncorrected. For technical details on how this is done, the reader is referred to the Online Appendix. Namely, we use the ridge inverse in the same one-step adjustment as CLIME, but we now define $\hat{\Theta}_R$ as

$$\hat{\Theta}_R = (\hat{\Gamma} + \lambda_R I_d)^{-1}, \quad \lambda_R = \max \left\{ 0.01, \sqrt{\frac{\log d}{n}} \right\}, \quad (\text{B.3})$$

$$\hat{b}^R = \hat{b}^L + \hat{\Theta}_R \hat{s}_L.$$

Here, R denotes the ridge adjustment, and $\hat{b}^R$ is the ADL estimator reported in the paper. Semenova et al. (2023) also use a similar ridge inverse in their empirical application to simplify computation.^{B.3} Holding the cross-fitted residuals, LASSO estimate, and score fixed, the two corrections differ only in the matrix translating the score into coefficient adjustments: $\hat{b}^C - \hat{b}^R = (\hat{\Theta}^C - \hat{\Theta}_R) \hat{s}_L$. While CLIME controls the approximate inverse error entry by entry, while favoring a sparse matrix, ridge uses a generally dense matrix and satisfies the exact identity $I_d - \hat{\Theta}_R \hat{\Gamma} = \lambda_R \hat{\Theta}_R$. Hence, our ridge inverse retains most of the assumptions listed for CLIME in the previous section, but does not require this sparsity restriction for its construction. Instead, valid inference requires the contribution of the remaining term $\lambda_R \hat{\Theta}_R (\hat{b}^L - b_0)$ to each reported elasticity to be negligible relative to its standard error.

C Pooled DML and IV Estimates of Price Elasticities

This appendix compares pooled DML and IV estimates using pharmaceutical prices abroad. Pooling retains only the constant in K_{igct} , reducing D_{igct} to T_{igct} and replacing the 488 price-response coefficients with two common elasticities, β_0 . Reestimating DML under this restriction provides a natural benchmark for pooled IV since both methods use the same sample, controls – i.e. both use the same conditioning set, $\mathcal{W}_{igct}^{IV}$, which excludes the two foreign-price instruments – and price-response specification, while relying on different identifying variation. Pooled DML assumes that domestic price movements remaining after conditioning on $\mathcal{W}_{igct}^{IV}$ are unrelated to domestic demand shocks, while IV allows these residual prices to remain endogenous and uses variation predicted by foreign prices. Conditional on $\mathcal{W}_{igct}^{IV}$, the instruments must jointly predict

^{B.3}CLIME results are available upon request.

domestic own and rival prices, be unrelated to domestic demand shocks, and affect domestic sales only through domestic prices.

The own-price instrument is the log mean price in levels of the same molecule–corporation–local-product key in other countries during the same quarter (Dubois et al., 2021). For the rival-price instrument, we replace each domestic rival’s price with its foreign mean and apply the same averaging rule:

$$z_{igct}^{\text{own}} = \log \left[\frac{1}{|\mathcal{F}_{igt}^{-c}|} \sum_{c' \in \mathcal{F}_{igt}^{-c}} \exp(p_{igc't}) \right],$$

$$z_{igct}^{\text{rival}} = \log \left[\frac{1}{|\mathcal{C}_{gct}| - 1} \sum_{\substack{j \in \mathcal{C}_{gct} \\ j \neq i}} \exp(z_{jgct}^{\text{own}}) \right].$$

Here, $\mathcal{F}_{igt}^{-c}$ contains foreign countries with an exact product match, and $\mathcal{C}_{gct}$ is the domestic same-molecule product set. Prices are ex-manufacturer unit values at constant exchange rates. We require foreign matches for the focal product and every domestic rival so that the two rival indexes cover the same basket. Requiring four quarterly lags then leaves 66,264 observations across 62 countries, 166 molecules, and 44 quarters: 15.4% of the baseline DML sample. The comparison therefore concerns a selected sample of internationally matched products.

For this comparison, both DML and IV use a *pooled* model. We retain individual observations and assume that the own-price and rival-price elasticities are each constant across countries, molecules, and years. The two coefficients are $\beta_0 = (\beta_0^{\text{own}}, \beta_0^{\text{rival}})'$.^{C.1} Estimating the main model’s 488 coefficients by conventional 2SLS would require instruments strong enough to distinguish the country, molecule, and year components of both responses. Interacting the shifters with K_{igct} provides candidate instruments whose joint strength would need to be assessed (Sanderson and Windmeijer, 2016). Given, however, the smaller matched sample and the limited own-price instrument strength documented below, we restrict the comparison to one common own-price elasticity and one common rival-price elasticity. Pooling reduces the number of effects the instruments must identify and allows us to compare these two responses under different identifying assumptions. Weak-instrument concerns remain in this simpler model.

We compare pooled DML, estimated as a partially linear regression (PLR), with 2SLS that instruments both domestic prices jointly. To make the comparison informative, both methods use the same observations and conditioning set, $\mathcal{W}_{igct}^{\text{IV}}$. We first remove the components of quantities, domestic prices, and instruments explained by these controls.^{C.2} Writing the domestic prices and foreign-price instruments as

$$T_{igct} = \begin{pmatrix} p_{igct} \\ p_{igct}^- \end{pmatrix}, \quad Z_{igct}^{\text{FP}} = \begin{pmatrix} z_{igct}^{\text{own}} \\ z_{igct}^{\text{rival}} \end{pmatrix},$$

the corresponding population residuals are

$$y_{igct}^{\circ} = y_{igct} - E[y_{igct} | \mathcal{W}_{igct}^{\text{IV}}],$$

$$T_{igct}^{\circ} = T_{igct} - E[T_{igct} | \mathcal{W}_{igct}^{\text{IV}}],$$

$$Z_{igct}^{\text{FP}, \circ} = Z_{igct}^{\text{FP}} - E[Z_{igct}^{\text{FP}} | \mathcal{W}_{igct}^{\text{IV}}],$$

^{C.1}For related pharmaceutical 2SLS applications, see Danzon and Chao (2000a).

^{C.2}Controls include panel-unit, therapeutic-class, country and molecule indicators; seasonal indicators; product type; aggregate and country-specific trends; four lags of own price and quantity; and country and molecule interactions with those lags.

where the conditional means are estimated separately using elastic net. The predictions have been constructed with ten contiguous neighbors-left-out time folds shared across the equations. Subtracting these predictions gives the cross-fitted residuals. Under the IV assumptions, the remaining instrument variation is uncorrelated with the component of residual quantity left unexplained by domestic prices:

$$E \left[Z_{igct}^{FP,\circ} (y_{igct}^{\circ} - (T_{igct}^{\circ})' \beta_0) \right] = 0. \quad (C.1)$$

For estimation, we proceeded with 2SLS. Different from Pooled DML that regresses residual quantity on the two price residuals, 2SLS instruments that price residuals with the two instrument residuals. Country clustering allows observations within a national market to be related. Table C.1 reports the first-stage regressions of 2SLS, showing how the foreign-price instruments predict the domestic price variation used to estimate the demand responses. Each domestic price is regressed on both instruments, allowing us to distinguish the predictive contribution of the focal product’s foreign price from that of its rivals’ foreign prices.

Table C.1. IV First-stage

Dependent price residual	Primary		Additional fixed effects	
	$\hat{z}^{own,\circ}$	$\hat{z}^{rival,\circ}$	$\hat{z}^{own,\circ}$	$\hat{z}^{rival,\circ}$
Own price	0.0354** (0.0149)	0.0031** (0.0015)	0.0669*** (0.0228)	0.0050*** (0.0016)
Rival-price index	-0.0268 (0.0243)	0.5567*** (0.0832)	0.0376* (0.0222)	0.5058*** (0.0955)

Notes: Both cross-fitted instrument residuals enter each regression. Country-clustered standard errors are in parentheses. Stars use the same two-sided normal-approximation tests as in Table C.2: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.2 brings together the estimated demand responses and the diagnostics needed to interpret them. Panel A reports the pooled DML elasticities and the second-stage 2SLS estimates, while Panel B assesses how strongly the foreign-price instruments predict domestic prices in the first stage. This is key since statistically significant of the first-stage coefficients reported in Table C.1 alone do not establish that both demand responses can be estimated precisely. Panel C reports Anderson–Rubin confidence sets for the IV elasticities, allowing for the uncertainty associated with weak instruments (Anderson and Rubin, 1949; Sanderson and Windmeijer, 2016; Andrews et al., 2019).

Table C.2. Pooled price responses, IV second-stage and diagnostics

	Primary		Additional fixed effects	
	PLR	IV	PLR	IV
<i>Panel A: Pooled estimates</i>				
Own-price elasticity	−0.5373*** (0.0782)	−0.5873** (0.2827)	−0.5368*** (0.0941)	−0.4342*** (0.1596)
Rival-price-index elasticity	0.0115** (0.0058)	0.0177*** (0.0053)	0.0129** (0.0059)	0.0236*** (0.0062)
<i>Panel B: System diagnostics</i>				
Observations	66,264	66,264	65,781	65,781
First-stage rank	–	2 of 2	–	2 of 2
Minimum scale-free singular value	–	0.1015	–	0.1435
Condition number	–	6.44	–	4.23
Partial R^2 , own price	–	0.0107	–	0.0218
Partial R^2 , rival-price index	–	0.4269	–	0.3678
Conditional F , own (country)	–	8.22	–	12.62
Conditional F , rival (country)	–	51.96	–	28.84
Conditional F , own (country–molecule)	–	26.38	–	60.31
Conditional F , rival (country–molecule)	–	66.10	–	54.31
Conditional F , own (two-way)	–	8.02	–	11.12
Conditional F , rival (two-way)	–	24.47	–	16.75
<i>Panel C: 95% Anderson–Rubin projections</i>				
Own price, country	–	Finite bounds not found	–	[−1.503, 0.145]
Rival-price index, country	–	Finite bounds not found	–	[0.0060, 0.0418]
Own price, country–molecule	–	[−1.426, 0.145]	–	[−1.021, 0.0617]
Rival-price index, country–molecule	–	[0.0028, 0.0333]	–	[0.0062, 0.0456]

Notes: PLR and IV use the same observations, controls and folds within each specification. Country-clustered standard errors are in parentheses. Stars use two-sided normal-approximation Wald tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. IV stars describe conventional significance; Panel C provides inference allowing for weak instruments. Conditional F statistics are the Sanderson–Windmeijer statistics for each price in the joint system. Two-way clustering is by country and molecule. Panel C projects the joint cluster-robust Anderson–Rubin region onto each coefficient. “Finite bounds not found” means the accepted region still reached the search-grid boundary after eight expansions.

Panel A of Table C.2 shows similar pooled own-price estimates under DML and IV, i.e. −0.537 and −0.587, respectively. These imply that – under the identifying assumptions – a 1% increase in own price reduces quantity by roughly 0.5–0.6%, holding rival prices fixed. Both estimates remain robust with additional fixed effects. Panel B shows that foreign prices predict domestic rival prices more strongly than own prices, after accounting for the controls. Own-price instrument strength improves with additional fixed effects moving the country-clustered conditional F statistic from 8.22 to 12.62. Nevertheless, the confidence sets in Panel C indicate substantial uncertainty about the own-price response. For example, with additional fixed effects, the country-clustered confidence set is [−1.503, 0.145], including both the DML estimate and zero.

The pooled DML and IV point estimates give a similar picture of own-price responsiveness on the matched sample, although the IV confidence sets remain wide. Notice, however, that imposing common coefficients does not account for the heterogeneity of the effects, which can affect the magnitude of the estimated coefficients (Goldsmith-Pinkham et al., 2024).

D Limitations of DML

Moving from a structural demand system to a reduced-form DML specification is useful because it allows for flexible heterogeneity in the estimation of price elasticities. However, this flexibility comes at a cost: the DML approach does not directly deliver the full product-by-product demand Jacobian needed for marginal-cost recovery. This appendix clarifies the trade-off involved in using a more flexible CATE-based elasticity specification. Fix a molecule–country–quarter market (g, c, t) , let $a = a(t)$ denote the calendar year containing quarter t , and suppress the corresponding market subscripts. Write $\mathcal{J}^g := \mathcal{C}_{gct}$ for the set of products containing molecule g in this country and quarter. Consistently with the main text, let $P_i := P_{igct} = \exp(p_{igct})$, $Q_i := Q_{igct} = \exp(y_{igct})$, and $P_i^- := \exp(p_{igct}^-)$. Let $P^g = (P_i)_{i \in \mathcal{J}^g}$ be the vector of prices. The within-molecule negative demand Jacobian is

$$\Delta_{\ell i}^g := -\frac{\partial Q_\ell}{\partial P_i}, \quad \ell, i \in \mathcal{J}^g,$$

where rows index the demand being differentiated and columns index the price that changes. We impose zero cross-effects across molecules, so the full product-level negative demand Jacobian is block diagonal, with one block Δ^g for each molecule.

The DML specification used in the main analysis does not estimate a separate cross-price effect for each rival product. Instead, for each product i , it summarizes rival prices through the leave-one-out index $P_i^- := \frac{1}{|\mathcal{J}^g| - 1} \sum_{j \in \mathcal{J}^g, j \neq i} P_j$. Thus, importantly, conditional on controls and fixed effects, the estimated demand equation is interpreted as depending on P_i and P_i^- , rather than on the full vector $\{P_j : j \in \mathcal{J}^g, j \neq i\}$. Because the heterogeneity dictionary contains country, molecule, and year terms but no product term, the DML elasticity coefficients are common across products within this fixed cell. They satisfy

$$\epsilon_{gca}^{\text{own}} = \frac{\partial \log Q_i}{\partial \log P_i}, \quad \epsilon_{gca}^{\text{rival}} = \frac{\partial \log Q_i}{\partial \log P_i^-}.$$

The own-price derivative is immediately $\frac{\partial Q_i}{\partial P_i} = \epsilon_{gca}^{\text{own}} \frac{Q_i}{P_i}$. For a rival product $j \in \mathcal{J}^g \setminus \{i\}$, applying the chain rule gives $\frac{\partial Q_i}{\partial P_j} = \frac{\partial Q_i}{\partial P_i^-} \frac{\partial P_i^-}{\partial P_j} = \epsilon_{gca}^{\text{rival}} \frac{Q_i}{P_i^-} \frac{1}{|\mathcal{J}^g| - 1}$. This allocation rule is the key maintained assumption needed to construct a full Jacobian from the DML estimates. It assigns the same quantity response to an equal change in any individual rival’s price in levels. Therefore, the DML-based within-molecule negative demand Jacobian is

$$\Delta_{\ell i}^g = - \begin{cases} \epsilon_{gca}^{\text{own}} \frac{Q_\ell}{P_\ell}, & i = \ell, \\ \epsilon_{gca}^{\text{rival}} \frac{Q_\ell}{(|\mathcal{J}^g| - 1) P_\ell}, & i \in \mathcal{J}^g \setminus \{\ell\}. \end{cases}$$

The off-diagonal derivatives vary across rows through quantities and prices but use the same cell-level rival-price elasticity and are constant across rival products within a row. Constructing the product-by-product demand Jacobian requires knowing how demand responds to each rival’s price separately. As seen, however, our specification estimates one response to the rival-price index. Indeed, to obtain the individual derivatives above, we assume that rival prices affect demand only through their arithmetic mean P_i^- .

This assumption determines how the estimated index response is allocated across rivals. Applying the chain rule gives $\frac{\partial \log Q_i}{\partial \log P_j} = \epsilon_{gca}^{\text{rival}} \frac{P_j}{\sum_{k \in \mathcal{J}^g, k \neq i} P_k}$, $j \in \mathcal{J}^g \setminus \{i\}$. Each rival’s cross-price elasticity is therefore proportional to its share of total rival prices. This follows from the construction of the index: a given percentage increase in a higher-priced rival changes the arithmetic mean by more.

These price shares sum to one, so applying the summation on both sides we get $\sum_{\substack{j \in \mathcal{J}^g \\ j \neq i}} \frac{\partial \log Q_i}{\partial \log P_j} = \epsilon_{gca}^{\text{rival}}$.

Intuitively, increasing all rival prices by the same percentage raises their mean by that percentage. Under the maintained index assumption, the index elasticity thus equals the sum of the individual cross-price elasticities. Our restriction concerns how this total response is distributed across rivals. For example, two equally priced rivals are assigned the same cross-price elasticity, even if consumers regard one as a closer substitute. Our specification does not estimate such differences separately. The resulting Jacobian therefore combines estimated price responsiveness with an assumed pattern of substitution across rivals. Marginal-cost recovery inherits this restriction.

The structure of the Jacobian Δ^g has also consequences on the markups once a conduct model is assumed. The idea is that, if changing a product's price also affects sales of the firm's products in other molecules, those effects would enter the pricing conditions and could change the recovered markups and costs. Recovery therefore would depend on the assumed allocation of substitution within molecules, the exclusion of cross-price effects between molecules, and the conduct. In order to show how the assumed substitution patterns enter the calculation of firms' markups (and hence marginal costs), suppose for instance, that we operate under Bertrand pricing. Under such a conduct each firm chooses its prices to maximize profits across its products, taking other firms' prices as given. It therefore considers how changing one product's price affects sales of its other products.

Let $o(i)$ denote the firm owning product i . Define $H_{\ell i}^g := \mathbf{1}\{o(\ell) = o(i)\}$, $\Omega^g := H^g \odot \Delta^g$, where $\odot$ denotes element-by-element multiplication. The ownership matrix H^g selects the demand derivatives relevant to these within-firm profit effects.

The assumption of zero cross-price effects across molecules makes the full demand Jacobian and its ownership-weighted counterpart block diagonal. This allows the pricing equations to be solved separately for each molecule, including when a firm owns products in several molecules.

Let $Q^g = (Q_i)_{i \in \mathcal{J}^g}$ and $MC^g = (MC_i)_{i \in \mathcal{J}^g}$ denote the quantity and marginal-cost vectors. The firm's pricing conditions imply $(\Omega^g)'(P^g - MC^g) = Q^g$. Provided $(\Omega^g)'$ is invertible, the implied markups are

$$P^g - MC^g = [(\Omega^g)']^{-1} Q^g.$$

Marginal costs are then obtained by subtracting these markups from observed prices.

E Existence of a Dictionary Approximation for Logit-Based Demand

Let $(X_m, p_m) \in \mathcal{S}_{\text{state}} \subset \mathbb{R}^{d_{\text{state}}}$ collect the relevant market-level state variables, where d_{state} is the dimension of that state, and assume that $\mathcal{S}_{\text{state}}$ is compact. Throughout, the model parameters are held fixed ^{E.1}. We use $(\cdot)$ for previously specified market-variable arguments when these are understood, retaining any consumer-heterogeneity argument explicitly. Establishing continuity and polynomial approximation for simple-logit shares is relatively easy and is done as follows. In the simple logit model, where $p_j = \log P_j$ denotes log price (suppressing the market subscript on product-level variables),

$$s_j^L(p_m, X_m) = \frac{\exp(\delta_j(x_j, p_m, \xi_j))}{1 + \sum_{\ell} \exp(\delta_{\ell}(x_{\ell}, p_m, \xi_{\ell}))},$$

$$\delta_j(x_j, p_m, \xi_j) = x_j' \beta - \alpha p_j + \xi_j.$$

^{E.1}Indeed, we study how shares and price responses vary with (X_m, p_m) , holding the structural parameters fixed. In RCL, preferences still differ across consumers, but the parameters governing this heterogeneity remain unchanged (Nevo, 2000).

First, let's consider the following observations. The mean utilities $\delta_j(\cdot)$ are continuous, then $s_j^L(\cdot)$ is continuous. Since $\mathcal{S}_{\text{state}}$ is compact and the logit probabilities, $s_j^L(\cdot)$, are strictly positive – because their exponential numerator is positive and its denominator is finite – for finite utilities, $s_j^L(\cdot)$ is also bounded away from zero on $\mathcal{S}_{\text{state}}$. Hence $\log s_j^L(\cdot)$ is continuous as well.

To establish that the (simple) logit share function can be approximated uniformly, we construct polynomial terms from the individual entries of (X_m, p_m) , such as product characteristics and log prices. These terms include constants, powers, and interactions. Following Semenova et al. (2023), a dictionary is a finite vector of selected transformations of the model variables, whose entries are combined linearly to represent a function of interest. In what follows, we consider polynomial dictionaries in (X_m, p_m) , including constants, powers, and interactions, to establish the approximation result. Namely, let $\mathcal{A}$ collect all finite linear combinations of polynomial terms of arbitrary degree, with real coefficients. Since polynomials are continuous and remain polynomials under addition and multiplication, $\mathcal{A}$ is a real subalgebra of $C(\mathcal{S}_{\text{state}}, \mathbb{R})$, the space of continuous real-valued functions on the market-state domain. We measure accuracy throughout this domain using the uniform norm, $\|f\|_\infty := \sup_{(X_m, p_m) \in \mathcal{S}_{\text{state}}} |f(X_m, p_m)|$, so that $\|f - h\|_\infty$ is the largest absolute difference between the target function f and its approximation h .

The polynomial collection contains constants and distinguishes any two different market states: if $(X_m, p_m) \neq (\tilde{X}_m, \tilde{p}_m)$, at least one coordinate differs, and the corresponding coordinate function belongs to $\mathcal{A}$. These properties, together with compactness of $\mathcal{S}_{\text{state}}$, allow us to apply the Stone–Weierstrass theorem: $\overline{\mathcal{A}}^{\|\cdot\|_\infty} = C(\mathcal{S}_{\text{state}}, \mathbb{R})$. The closure on the left consists of functions that can be approximated arbitrarily closely in the uniform norm by polynomials in $\mathcal{A}$. Consequently, for every $\varepsilon > 0$, there exist a finite polynomial dictionary $K_s(X_m, p_m)$ and a coefficient vector γ_s such that

$$\sup_{(X_m, p_m) \in \mathcal{S}_{\text{state}}} |s_j^L(p_m, X_m) - K_s(X_m, p_m)' \gamma_s| < \varepsilon.$$

Thus, a weighted sum of sufficiently many polynomial terms can approximate the logit share with an error smaller than any chosen tolerance at every market state in the specified domain. The approximation remains linear in its coefficients while allowing nonlinear dependence on prices and characteristics. This guarantee concerns increasingly rich dictionaries whose linear spans are dense in $C(\mathcal{S}_{\text{state}}, \mathbb{R})$. The same argument applies directly to any continuous logit-implied price-response function. For example, because price enters utility in logs, the simple-logit own-price elasticity is $f_j(X_m, p_m) = \frac{\partial \log s_j^L(p_m, X_m)}{\partial p_j} = -\alpha(1 - s_j^L(p_m, X_m))$. Since f_j is continuous, for every $\varepsilon > 0$, there exist a sufficiently rich finite polynomial dictionary $H_j(X_m, p_m)$ and a coefficient vector ρ_j such that $\sup_{(X_m, p_m) \in \mathcal{S}_{\text{state}}} |f_j(X_m, p_m) - H_j(X_m, p_m)' \rho_j| < \varepsilon$. This is a direct approximation of the elasticity function. In this example, $H_j(\cdot)' \rho_j$ directly approximates the own-price elasticity $\partial \log s_j^L(\cdot) / \partial p_j$. Obtaining elasticities by differentiating an approximation to log demand requires accuracy in the relevant price derivatives as well as in function values. This additional condition is discussed after the proof of Proposition 2.1.

As mentioned, the preceding argument concerns simple logit. The next proposition establishes continuity and uniform approximation for RCL shares and log shares as well. This supplies the RCL share properties used in the subsequent proof of Proposition 2.1, which covers both models. As a note for readability, in simple logit, systematic utility is common across consumers, so the choice probability $s_j^L(\cdot)$ also gives the aggregate market share. RCL retains the logit formula for each consumer but allows systematic utility to vary with consumer preferences. We denote the resulting choice probability, conditional on v , by $\sigma_{jm}(\cdot, v)$; averaging these probabilities over F gives the aggregate share $s_{jm}^{\text{RCL}}(\cdot)$.

Proposition E.1 (Continuity and approximation of RCL shares). *Let $(X_m, p_m) \in \mathcal{S}_{\text{state}} \subset \mathbb{R}^{d_{\text{state}}}$, with $\mathcal{S}_{\text{state}}$ compact and the choice set fixed. Let $\mathbf{v}$ denote consumer heterogeneity with distribution F . The individual RCL choice probability is*

$$\sigma_{jm}(p_m, X_m, \mathbf{v}) = \frac{\exp\{\delta_{jm}(x_{jm}, p_m, \xi_{jm}) + \mu_{jm}(X_m, p_m, \mathbf{v})\}}{1 + \sum_{\ell \in \mathcal{J}_m} \exp\{\delta_{\ell m}(x_{\ell m}, p_m, \xi_{\ell m}) + \mu_{\ell m}(X_m, p_m, \mathbf{v})\}},$$

and the aggregate RCL share is

$$s_{jm}^{\text{RCL}}(p_m, X_m) = \int \sigma_{jm}(p_m, X_m, \mathbf{v}) dF(\mathbf{v}).$$

Assume that, for F -almost every $\mathbf{v}$, the functions $\delta_{\ell m}(\cdot)$ and $\mu_{\ell m}(\cdot, \mathbf{v})$ are continuous in (X_m, p_m) for every product ℓ in the choice set $\mathcal{J}_m$. Then $s_{jm}^{\text{RCL}}(\cdot)$ is continuous on $\mathcal{S}_{\text{state}}$. If $s_{jm}^{\text{RCL}}(p_m, X_m) > 0$ for all $(X_m, p_m) \in \mathcal{S}_{\text{state}}$, then compactness implies that s_{jm}^{RCL} is bounded away from zero, and hence $\log s_{jm}^{\text{RCL}}(\cdot)$ is continuous and can be uniformly approximated on $\mathcal{S}_{\text{state}}$ by a sufficiently rich finite polynomial dictionary.

Proof. Let $(X_m^{(n)}, p_m^{(n)})$ be a sequence of market states converging to (X_m, p_m) , where n indexes the states in the sequence. For a fixed consumer type $\mathbf{v}$, $\sigma_{jm}(p_m^{(n)}, X_m^{(n)}, \mathbf{v})$ denotes the probability of choosing product j at the n -th market state. Continuity of the utility functions and the logit formula implies that this probability converges to its value at the limiting state:

$$\sigma_{jm}(p_m^{(n)}, X_m^{(n)}, \mathbf{v}) \rightarrow \sigma_{jm}(p_m, X_m, \mathbf{v}).$$

This holds for F -almost every $\mathbf{v}$, meaning that any exceptions form a set of probability zero under F . Moreover, $0 \leq \sigma_{jm}(p_m^{(n)}, X_m^{(n)}, \mathbf{v}) \leq 1$. Since F is a probability measure, the dominated convergence theorem implies

$$\begin{aligned} s_{jm}^{\text{RCL}}(p_m^{(n)}, X_m^{(n)}) &= \int \sigma_{jm}(p_m^{(n)}, X_m^{(n)}, \mathbf{v}) dF(\mathbf{v}) \\ &\rightarrow \int \sigma_{jm}(p_m, X_m, \mathbf{v}) dF(\mathbf{v}) = s_{jm}^{\text{RCL}}(p_m, X_m). \end{aligned}$$

Thus s_{jm}^{RCL} is continuous. If it is strictly positive on compact $\mathcal{S}_{\text{state}}$, it attains a positive minimum, so $\log s_{jm}^{\text{RCL}}$ is continuous. The approximation claim then follows from the same real Stone–Weierstrass argument that we previously applied to the simple logit case. $\square$

The preceding proposition extends the continuity and approximation argument from simple logit to RCL. We now use this result to prove Proposition 2.1 – which refers to general logit-based models. Proposition 2.1 concerns the price-dependent component of log quantity for either simple logit or RCL. As a reminder from the main text, we write $m_j^{\text{LB}}(X_m, p_m) = a_j(X_m) + \tau_j(X_m, p_m)$. Here, $a_j(X_m)$ collects baseline demand determinants that do not depend on current prices, including the market-size term. The remaining component $\tau_j(X_m, p_m)$ contains the dependence on prices and allows that dependence to vary with market characteristics. Intuitively, under the assumed continuity of τ_j , the same Stone–Weierstrass argument shows that this component can be uniformly approximated by a sufficiently rich polynomial dictionary ^{E.2}.

^{E.2}This is an existence result for a sufficiently rich general dictionary; whether a particular finite empirical dictionary provides an adequate approximation is an additional specification condition.

Proof of Proposition 2.1. By assumption, the price-response component $\tau_j(X_m, p_m) = m_j^{\text{LB}}(X_m, p_m) - a_j(X_m)$ is continuous on the compact set $\mathcal{S}_{\text{state}}$. Consider the algebra $\mathcal{A}$ of real polynomials in the coordinates of (X_m, p_m) , restricted to this set. As established in Proposition E.1, its functions are continuous, it contains constants, and it separates points because distinct market states differ in at least one coordinate. The real Stone–Weierstrass theorem therefore implies that $\mathcal{A}$ is dense in $C(\mathcal{S}_{\text{state}}, \mathbb{R})$. In particular, for any $\varepsilon > 0$, there exists a polynomial whose uniform distance from τ_j is smaller than ε . Every such polynomial contains finitely many terms. Collecting these terms into the dictionary $D_j(X_m, p_m)$ and their coefficients into γ_D gives

$$\sup_{(X_m, p_m) \in \mathcal{S}_{\text{state}}} |\tau_j(X_m, p_m) - D_j(X_m, p_m)' \gamma_D| < \varepsilon.$$

The dictionary and its coefficients may depend on the desired accuracy ε . Define the approximation error by $\tau_{jm} = \tau_j(X_m, p_m) - D_j(X_m, p_m)' \gamma_D$. Substituting in τ_{jm} the definition of τ_j yields

$$m_j^{\text{LB}}(X_m, p_m) = a_j(X_m) + D_j(X_m, p_m)' \gamma_D + \tau_{jm}, \quad \sup_{(X_m, p_m) \in \mathcal{S}_{\text{state}}} |\tau_{jm}| < \varepsilon,$$

which proves the claimed representation E.3. $\square$

Proposition 2.1 establishes that the price-response component of log demand can be approximated uniformly for both simple logit and RCL. Elasticities, however, depend on the derivatives of this component with respect to log prices. Accuracy in function values alone does not guarantee accuracy in these derivatives. For RCL, we first establish conditions under which the elasticity functions are continuous, allowing them to be approximated directly by polynomial dictionaries. We then state the additional condition needed when elasticities are obtained by differentiating a dictionary approximation to log demand.

To approximate elasticities, we first establish that the price derivatives of aggregate shares are continuous. Hold the choice set, the structural parameters, and the distribution F fixed. For F -almost every $\mathbf{v}$, assume that the log-price derivatives of $\sigma_{jm}(p_m, X_m, \mathbf{v})$ exist and are jointly continuous in (X_m, p_m) . Impose these conditions on $\mathcal{S}_{\text{state}}$ and a neighborhood in the price coordinates. These smoothness conditions hold in standard logit and random-coefficients logit specifications with smooth systematic utilities and a fixed choice set (Berry et al., 1995; Nevo, 2001).

We want to calculate how the market share changes when a price changes. Since market share averages consumers' choice probabilities, we ask whether its price response can be calculated by averaging their individual price responses. To do so, suppose a nonnegative function $B(\mathbf{v})$ bounds the absolute price derivatives throughout this neighborhood. In particular, for every $k \in \mathcal{J}_m$,

$$\sup_{(X_m, p_m) \in \mathcal{S}_{\text{state}}} \left| \frac{\partial \sigma_{jm}(p_m, X_m, \mathbf{v})}{\partial p_{km}} \right| \leq B(\mathbf{v}), \quad \int B(\mathbf{v}) dF(\mathbf{v}) < \infty.$$

The envelope $B(\mathbf{v})$ bounds the size of each consumer type's response across market states. Requiring this bound to have a finite population average keeps the average absolute response finite. Together with the assumed continuity of individual price derivatives in (X_m, p_m) , this integrable bound allows us to differentiate the aggregate share by averaging the individual derivatives:

$$\frac{\partial s_{jm}^{\text{RCL}}(p_m, X_m)}{\partial p_{km}} = \int \frac{\partial \sigma_{jm}(p_m, X_m, \mathbf{v})}{\partial p_{km}} dF(\mathbf{v}).$$

^{E.3}Notice that the error τ_{jm} varies with the market state (X_m, p_m) ; the uniform bound ensures that its absolute value remains below ε at every state in $\mathcal{S}_{\text{state}}$.

The left-hand side is the market's price response, while the right-hand side averages individual responses using their population weights. The same conditions ensure that small changes in market conditions produce small changes in this aggregate response. In the standard RCL model with a linear random coefficient on log price, the required envelope exists if the average absolute value of that coefficient is finite.

Since the shares are continuous and bounded away from zero, dividing a share derivative by the corresponding share gives a continuous elasticity:

$$\varepsilon_{jk,m}^{\text{RCL}}(X_m, p_m) = \frac{1}{s_{jm}^{\text{RCL}}(p_m, X_m)} \frac{\partial s_{jm}^{\text{RCL}}(p_m, X_m)}{\partial p_{km}} = \frac{\partial \log s_{jm}^{\text{RCL}}(p_m, X_m)}{\partial p_{km}}.$$

The Stone–Weierstrass argument therefore applies to each elasticity function: a sufficiently rich polynomial dictionary can approximate it uniformly on $\mathcal{S}_{\text{state}}$.

A further condition is needed when elasticities are obtained by differentiating an approximation to log demand. An approximation can reproduce function values closely while reproducing their slopes poorly. Since $a_j(X_m)$ and market size are held fixed when prices change, the price derivatives of $\tau_j(X_m, p_m)$ are the corresponding elasticities. We therefore require the same dictionary approximation to reproduce both τ_j and its price derivatives. Specifically, for every $\varepsilon > 0$, assume that there exist a finite differentiable dictionary D_j and a coefficient vector γ_D such that

$$\begin{aligned} & \sup_{(X_m, p_m) \in \mathcal{S}_{\text{state}}} |\tau_j(X_m, p_m) - D_j(X_m, p_m)' \gamma_D| \\ & + \max_{k \in \mathcal{J}_m} \sup_{(X_m, p_m) \in \mathcal{S}_{\text{state}}} \left| \frac{\partial \tau_j(X_m, p_m)}{\partial p_{km}} - \frac{\partial [D_j(X_m, p_m)' \gamma_D]}{\partial p_{km}} \right| < \varepsilon. \end{aligned}$$

The first term controls approximation error in the price-dependent component of log quantity, while the second controls the largest error among its own- and cross-price derivatives. This condition ensures that differentiating the demand approximation also approximates the elasticities. Estimating these derivatives from data, however, can be more demanding than estimating the function itself. Under their smoothness and regularity conditions, Chen et al. (2025) establish slower uniform convergence rates for derivative estimates than for function estimates. The existence of an accurate approximation, therefore, does not guarantee equally precise estimates of log demand and its elasticities.

F High-Dimensional RCL: Rank Failure and Regularization

In RCL, estimation proceeds by fixing the nonlinear parameters θ_2 , inverting the observed market shares to recover the mean utilities $\delta_{jm}(\theta_2)$, and then forming the structural demand residual

$$\xi_{jm}(\theta) = \delta_{jm}(\theta_2) - X'_{1,jm} \theta_1,$$

with $X'_{1,jm} = (-p_{jm}, x'_{jm})$ and the linear mean-utility coefficients $\theta_1 = (\alpha, \beta')'$ as in the main text. $\delta_{jm}(\theta_2)$ is usually recovered from the observed shares by BLP inversion at the candidate value θ_2 . The GMM estimator jointly estimates $\theta = (\theta'_1, \theta'_2)'$ by making the sample demand moments as close to zero as possible. Let $\theta_0 = (\theta'_{1,0}, \theta'_{2,0})'$ denote the true structural parameter vector and let z_{jm}^{IV} denote the observation-level RCL instrument vector. The conditional restriction is $E[\xi_{jm}(\theta_0) | z_{jm}^{\text{IV}}] = 0$. As discussed by Backus et al. (2021), empirical GMM estimation commonly uses weaker unconditional moment restrictions implied by this condition. Here, these are

$$E[z_{jm}^{\text{IV}} \xi_{jm}(\theta_0)] = 0.$$

For these moments to identify the parameters locally, they must provide enough distinct information to distinguish changes in each parameter. Under standard regularity conditions, a sufficient condition is

$$\text{rank} \left(E \left[z_{jm}^{\text{IV}} \frac{\partial \xi_{jm}(\theta)}{\partial \theta'} \Big|_{\theta=\theta_0} \right] \right) = \dim(\theta). \quad (\text{F.1})$$

This is the classical full column rank condition on the Jacobian, which requires at least as many moments as parameters and rules out parameter changes whose effects cancel in every moment to first order.

Hold θ_2 fixed and consider whether the moments uniquely determine $\theta_1 = (\alpha, \beta')'$ ^{F.1}. Since $\xi_{jm}(\theta) = \delta_{jm}(\theta_2) - X'_{1,jm} \theta_1$, we have $\frac{\partial \xi_{jm}(\theta)}{\partial \theta'_1} = -X'_{1,jm}$. The Jacobian for the linear block is therefore $-E[z_{jm}^{\text{IV}} X'_{1,jm}]$ and – since multiplication by minus one does not change its rank – a necessary identification condition implied by Equation (F.1) is the following

$$\text{rank} (E[z_{jm}^{\text{IV}} X'_{1,jm}]) = k_x + 1, \quad (\text{F.2})$$

where $k_x + 1$ counts the price coefficient and the k_x characteristic coefficients. Thus, exogenous instruments must also provide enough independent information to distinguish these coefficients. If this condition fails, different θ_1 vectors generate the same population moments at any fixed θ_2 . Searching over θ_2 cannot remove that ambiguity, because the matrix governing it does not depend on θ_2 . The linear rank condition is therefore necessary for joint identification (although satisfying it alone does not establish identification of the nonlinear parameters). We will formalize this intuition later in this subsection.

The sample counterpart is additionally constrained by the number of observations. This requires the additional condition that the rank is also lower than the number of observations. This is formalized in what follows. Stacking observations, let $J_m := |\mathcal{J}_m|$ and $N = \sum_m J_m$ denote the total number of product–market observations. Let Z be the $N \times L_Z$ RCL instrument matrix, where L_Z is the number of instruments, and let X_1 be the $N \times (k_x + 1)$ matrix of price and characteristics. The rows of Z and X_1 corresponding to observation (j, m) are $(z_{jm}^{\text{IV}})'$ and $X'_{1,jm}$, respectively. The sample analogue satisfies

$$\text{rank} \left(\frac{1}{N} Z' X_1 \right) \leq \min\{L_Z, N, k_x + 1\}.$$

Notice that, when taking into account the sample analog, the sample size N (and the number of instruments) matters. Indeed, additionally to condition (F.2) we need to take into account that if $k_x + 1 > N$, the sample counterpart of the full-rank condition in Equation (F.2) cannot hold, and the unrestricted vector θ_1 is not uniquely determined by the sample moments of the standard non-regularized BLP/GMM procedure.^{F.2}

As intuitively mentioned, a failure of the linear rank condition in Equation (F.2) also implies that the complete vector $\theta = (\theta'_1, \theta'_2)'$ cannot satisfy the regular first-order full-column-rank condition. To see this, define the population Jacobian blocks

$$G_1^{\text{RCL}} = -E[z_{jm}^{\text{IV}} X'_{1,jm}], \quad G_2^{\text{RCL}} = E \left[z_{jm}^{\text{IV}} \frac{\partial \delta_{jm}(\theta_2)}{\partial \theta'_2} \Big|_{\theta_2=\theta_{2,0}} \right],$$

^{F.1}This does not assume that θ_2 is known in estimation. For each trial value of θ_2 , the BLP algorithm can solve for the linear coefficients and then search over θ_2 . This computational profiling step remains part of the joint GMM estimation problem described in Nevo (2000).

^{F.2}Equivalently, after share inversion BLP produces only N recovered mean utilities δ_{jm} . More than N unrestricted linear coefficients cannot be uniquely fitted to these values. In the extreme case in which X_1 has full row rank N , for every candidate θ_2 there exists some θ_1 such that $X_1 \theta_1 = \delta_N(\theta_2)$, implying $\xi_N(\theta_1, \theta_2) = 0$ and hence a zero demand-side GMM objective for every θ_2 .

and write $G^{\text{RCL}} = [G_1^{\text{RCL}} \ G_2^{\text{RCL}}]$. If G_1^{RCL} is rank deficient – i.e. Equation (F.2) fails to hold – there exists a nonzero vector v_1 such that $G_1^{\text{RCL}} v_1 = 0$. Hence there is a direction in the linear parameter space along which moving the parameters does not change the moments at first order. The corresponding null direction of the full Jacobian G^{RCL} is $(v_1', 0')'$. The conventional GMM estimator, therefore, does not provide a point-identified estimate of the complete unrestricted parameter vector since the moment conditions are locally insensitive to perturbations of the form $(\theta_{1,0} + hv_1, \theta_{2,0})$, where h is a scalar, and cannot distinguish θ_0 from nearby alternatives lying in the direction v_1 .

As a conclusive note, nonuniqueness of θ_1 does not by itself imply nonuniqueness of θ_2 . To identify θ_2 , we must rule out changes in θ_2 that can be offset by adjustments to θ_1 , leaving the population moments equal to zero. Otherwise, the same moment conditions would be satisfied at different values of θ_2 , and the ambiguity would extend to θ_2 . A particularly important sample failure occurs when $\text{rank}(X_1) = N$. In this case, the linear component is flexible enough to reproduce any vector of N recovered mean utilities. Thus, for every candidate θ_2 , we can choose θ_1 such that $X_1 \theta_1 = \delta_N(\theta_2)$, $\xi_N(\theta_1, \theta_2) = 0$. Every candidate θ_2 therefore achieves a zero demand-side GMM objective after adjusting θ_1 since $\xi_N(\theta_1, \theta_2) = 0 \ \forall \theta_2$. The sample demand moments cannot distinguish among these parameter combinations, so they do not uniquely determine the joint structural parameter vector.

To formalize the requirement on θ_2 (at first order), let R_1^{RCL} project out the moment changes that the linear parameters can generate, using the GMM weighting metric. Then $R_1^{\text{RCL}} G_2^{\text{RCL}}$ describes the remaining effect of changes in θ_2 . Regular first-order identification of θ_2 requires

$$\text{rank}(R_1^{\text{RCL}} G_2^{\text{RCL}}) = \dim(\theta_2),$$

which requires that every nonzero direction of change in θ_2 leave some effect on the moments that adjustments to θ_1 cannot cancel.

Proposition 2.2 in the main text formalizes the observations that we discussed so far. The proof of Proposition 2.2 follows. For convenience, we remind here the definition $g_N(\theta_1, \theta_2) = \frac{1}{N} Z' \xi_N(\theta_1, \theta_2)$.

Proof of Proposition 2.2. For fixed θ_2 , $\frac{\partial g_N(\theta_1, \theta_2)}{\partial \theta_1'} = -\frac{1}{N} Z' X_1$. If this matrix has rank smaller than $k_x + 1$, there exists $v_1 \neq 0$ such that $Z' X_1 v_1 = 0$. Therefore, for every scalar h ,

$$g_N(\theta_1 + hv_1, \theta_2) = g_N(\theta_1, \theta_2),$$

and hence the GMM objective is constant along this direction ^{F.3}. This proves nonuniqueness, i.e. if the matrix is rank deficient, different coefficients give the same moments. We should now prove the converse: having more coefficients than observations or instruments necessarily makes the matrix rank-deficient. Because Z' is $L_Z \times N$ and X_1 is $N \times (k_x + 1)$, simple algebra gives $\text{rank}(Z' X_1) \leq \min\{L_Z, N, k_x + 1\}$. Thus, if we have more coefficients than observations, i.e. $k_x + 1 > N$ (or $k_x + 1 > L_Z$), the rank is necessarily smaller than $k_x + 1$, and the preceding argument establishes nonuniqueness of θ_1 . This completes the proof of the

^{F.3}To see this, using the definition of the structural residual, $g_N(\theta_1 + hv_1, \theta_2) = \frac{1}{N} Z' [\delta_N(\theta_2) - X_1(\theta_1 + hv_1)]$. Subtracting the original moment vector gives

$$\begin{aligned} & g_N(\theta_1 + hv_1, \theta_2) - g_N(\theta_1, \theta_2) \\ &= \frac{1}{N} Z' [\delta_N(\theta_2) - X_1 \theta_1 - h X_1 v_1] - \frac{1}{N} Z' [\delta_N(\theta_2) - X_1 \theta_1] \\ &= -\frac{h}{N} Z' X_1 v_1 = 0. \end{aligned}$$

Thus, moving the coefficients along v_1 leaves the sample moments unchanged. The moments themselves need not be zero; their difference is zero.

first claim of the proposition.

Let us now move to the second claim. The full sample Jacobian includes derivatives with respect to both parameter blocks:

$$G_N^{\text{RCL}}(\theta_1, \theta_2) := \frac{\partial g_N(\theta_1, \theta_2)}{\partial \theta'} = \begin{bmatrix} -\frac{1}{N} Z' X_1 & \frac{1}{N} Z' \frac{\partial \delta_N(\theta_2)}{\partial \theta_2'} \end{bmatrix}.$$

Moving θ_1 along v_1 while leaving θ_2 unchanged corresponds to the nonzero joint direction $(v_1', 0')'$. Since $Z' X_1 v_1 = 0$,

$$G_N^{\text{RCL}}(\theta_1, \theta_2) \begin{pmatrix} v_1 \\ 0 \end{pmatrix} = -\frac{1}{N} Z' X_1 v_1 = 0.$$

Thus, adding the columns associated with θ_2 cannot remove the linear dependence already present among the θ_1 columns. The full sample Jacobian therefore cannot have full column rank, proving item 2.

If $\text{rank}(X_1) = N$, its columns span $\mathbb{R}^N$. Because $\delta_N(\theta_2) \in \mathbb{R}^N$, for every admissible θ_2 there exists a solution to $X_1 \theta_1 = \delta_N(\theta_2)$. This gives

$$\xi_N(\theta_1, \theta_2) = 0, \quad g_N(\theta_1, \theta_2) = 0, \quad Q_N(\theta_1, \theta_2) = 0.$$

Thus the sample demand moments cannot distinguish among these parameter vectors. $\square$

Proposition 2.2 concerns the uniqueness of the structural parameter vector. Elasticities may nevertheless remain unique if all parameter combinations producing the same sample moments imply the same price responses. For example, including both product fixed effects and product characteristics that are constant across markets prevents their separate coefficients from being uniquely determined within the demand equation. Provided that α and θ_2 are identified, reallocating the contributions of these characteristics and fixed effects leaves mean utilities and price elasticities unchanged (Nevo, 2000, Section 3.5).

In Simulation 2, we present a case where the elasticity might also not be uniquely determined as a consequence of the nonuniqueness of the structural parameter vector. In that case, holding θ_2 fixed, changes in the mean price coefficient α are offset by changes in the characteristic coefficients, preserving the recovered mean utilities and sample demand moments, but changing price responses. For a simple illustration, suppose log price p_{jm} equals one characteristic $x_{jm,1}$ throughout the observed sample: the mean-utility specifications $-p_{jm} + 2x_{jm,1}$ and $-2p_{jm} + 3x_{jm,1}$ then give identical utilities and shares, holding all other terms fixed. Their derivatives with respect to log price, holding the characteristic fixed, are nevertheless -1 and -2 , so they imply different own-price elasticities. A more concrete example is presented in the main text.

F.1 An illustrative regularized structural route

Suppose, as in the above, that the available data and instruments leave several RCL parameter vectors consistent with the same sample demand moments, but these vectors imply different elasticities. For example, the observed purchases may not distinguish a medicine's benefit from consumers' sensitivity to its price. Our concern is whether regularization can resolve this ambiguity.

One possible naive approach would be to use Proposition 2.1 to represent demand through a rich dictionary and regularize its coefficients. Choosing $a_j(X_m) = \log M_m$ gives $\tau_j = \log s_{jm}^{\text{RCL}}$. Under the proposition's assumptions, for every $\varepsilon > 0$, there exist a finite dictionary D_j and coefficients γ_D such that

$$\sup_{(X_m, p_m) \in \mathcal{S}_{\text{state}}} |\log s_{jm}^{\text{RCL}}(p_m, X_m) - D_j(X_m, p_m)' \gamma_D| < \varepsilon.$$

This representation suggests estimating γ_D with a penalty that favors fewer terms or smaller coefficients. Such a penalty may resolve numerical nonuniqueness by selecting among representations that fit the observed shares equally well, but does not guarantee the selection of the correct price response. The latter, indeed, also requires assumptions under which the available variation distinguishes the relevant sparse alternatives. The difficulty is that different dictionary representations can also fit the observed shares equally well while implying different responses to a price change. Even two sparse coefficient vectors may satisfy

$$\begin{aligned} D_j(\cdot)' \gamma_D^{(A)} &= D_j(\cdot)' \gamma_D^{(B)} && \text{at all observed states,} \\ \frac{\partial [D_j(\cdot)' \gamma_D^{(A)}]}{\partial p_{jm}} &\neq \frac{\partial [D_j(\cdot)' \gamma_D^{(B)}]}{\partial p_{jm}} && \text{at some of those states.} \end{aligned}$$

A penalty may select one representation, but that choice alone does not establish which price response is correct. If the two representations also have the same penalty, it may not distinguish them at all. Sparsity therefore helps only when the available identifying variation can distinguish the relevant sparse alternatives, thus not solving the original problem. The approximation theorem, indeed, establishes neither this condition nor the accuracy of the estimated price derivatives.

Hence, to address the structural estimation problem, we need a more direct approach. In what follows, we propose placing the sparsity restriction on the RCL characteristic coefficients β and penalizing them within the GMM criterion, leaving α unpenalized. This approach retains the structural demand model and its instrument restrictions (Fan and Liao, 2014; Sawada and Kawaguchi, 2020). Its potential comes from restricting the parameter combinations that need to be distinguished. Successful recovery still requires the true model to satisfy the sparsity restriction and valid instruments to distinguish the relevant sparse alternatives. We now present an illustrative route for putting such a direct regularization into practice. Establishing its consistency and valid inference is beyond the scope of this paper.^{F.4} Our proposed approach targets nonuniqueness arising from an unrestricted mean-utility specification with many candidate characteristics. The idea is that if the true characteristic coefficients are sparse, regularization can exploit this restriction without requiring the relevant characteristics to be known in advance. Recovery additionally requires the instruments to distinguish the relevant sparse parameter combinations, including those with different price coefficients and nonlinear parameters. Its potential gain is therefore consistent estimation under sparsity and restricted identification, even when the unrestricted sample problem has multiple solutions.

Following Nevo (2000), BLP estimation searches over the nonlinear parameters θ_2 . For each trial value, share inversion recovers the mean utilities that reproduce observed market shares. The linear coefficients $\theta_1 = (\alpha, \beta')'$ are then chosen to minimize the GMM objective Q_N defined in Proposition 2.2. Comparing the resulting objective values across trial values of θ_2 gives

$$\hat{\theta}_1(\theta_2) \in \arg \min_{\theta_1} Q_N(\theta_1, \theta_2), \quad \hat{\theta}_2 \in \arg \min_{\theta_2} Q_N(\hat{\theta}_1(\theta_2), \theta_2),$$

where $\hat{\theta}_1(\theta_2)$ denotes the linear estimates associated with each trial value of θ_2 . This profiling procedure implements joint estimation of both parameter blocks.

Suppose that β – the vector of characteristic coefficients in mean utility whose high-dimensionality may create the discussed rank failure – is sparse in the sense that only a small subset of the candidate characteristics has nonzero coefficients. The idea is to introduce a penalty term to the moment conditions in order to force a shrinkage of the coefficients toward zero, following the general approach of penalized GMM

^{F.4}Our simulation uses a regularized moment-constraint estimator followed by a post-selection IV/GMM refit.

studied in Caner (2009) and Caner and Zhang (2014). An illustrative extension replaces the linear step with the "inner-step"

$$\hat{\theta}_1^\lambda(\theta_2) \in \arg \min_{\theta_1=(\alpha, \beta)'} \{Q_N(\theta_1, \theta_2) + \lambda_N \|\beta\|_1\}, \quad (\text{F.3})$$

where $\|\beta\|_1$ is the sum of absolute characteristic coefficients and $\lambda_N \geq 0$ controls the penalty. The mean-price coefficient α remains unpenalized. For fixed θ_2 and a fixed positive-definite weighting matrix W_N ; this is a convex, LASSO-type problem that admits a closed-form solution.

After estimating $\hat{\theta}_1^\lambda(\theta_2)$ for each trial value of θ_2 , we compare those candidates using their total GMM fit and coefficient penalty. This extends the procedure described by Nevo (2000) to the following "outer-step"

$$\hat{\theta}_2^\lambda \in \arg \min_{\theta_2} \left\{ Q_N(\hat{\theta}_1^\lambda(\theta_2), \theta_2) + \lambda_N \|\hat{\beta}^\lambda(\theta_2)\|_1 \right\},$$

where $\hat{\beta}^\lambda(\theta_2)$ contains the characteristic coefficients estimated at that trial value. Thus, for each trial value of θ_2 , the inner step holds θ_2 fixed and estimates the linear coefficients $\theta_1 = (\alpha, \beta)'$ by minimizing $Q_N(\theta_1, \theta_2) + \lambda_N \|\beta\|_1$. Each trial value of θ_2 produces its own linear estimates $\hat{\theta}_1^\lambda(\theta_2)$ and associated penalty. The outer step compares these candidates using the total $Q_N + \lambda_N \|\hat{\beta}^\lambda(\theta_2)\|_1$ and selects the smallest value. Although the coefficients were estimated using regularization, Q_N itself contains only the moment-fit term. We therefore include $\lambda_N \|\hat{\beta}^\lambda(\theta_2)\|_1$ when comparing candidates, so that both steps minimize the same total objective. Together, the two steps minimize a single objective:

$$(\hat{\theta}_1^\lambda, \hat{\theta}_2^\lambda) \in \arg \min_{\theta_1, \theta_2} \{Q_N(\theta_1, \theta_2) + \lambda_N \|\beta\|_1\},$$

with $\hat{\theta}_1^\lambda = \hat{\theta}_1^\lambda(\hat{\theta}_2^\lambda)$. The proposed procedure assumes that sparsity restricts the structural utility coefficients directly. As for Nevo (2000), recovering the true parameters still requires valid instruments and enough information in the moments to distinguish the relevant sparse parameter combinations, including those with different α and θ_2 . Together with suitable penalty and sampling conditions, the latter identification condition provides a basis for consistency. Establishing consistency and valid inference for this profiled RCL extension remains beyond our scope.

G Simulations

This Appendix describes the two simulation exercises discussed in Section 2.2. The first generates quantities from a known random-coefficients logit (RCL) model and applies DML to the resulting product-market panel. Comparing the estimated elasticities with those computed from the data-generating process assesses how closely the DML specifications recover the model's own-price responses.

The second simulation examines the rank failure in Proposition 2.2. Data are generated from sparse RCL models, but the unrestricted mean-utility specification contains more coefficients than product-market observations. We compare the elasticities from observationally indistinguishable unrestricted RCL solutions with oracle sparse and regularized RCL estimates. The panel with exogenous prices also reports a regularized DML price-response estimate.

G.1 Simulation results: Proposition 2.1

Motivated by Proposition 2.1, the simulation examines how closely DML, using a rich dictionary of observable variables, can recover the own-price elasticities generated by the RCL model. The details on how the

latter is built are in the Online Appendix. We summarize the simulation setup here for reference. In particular, we simulate product–market data using a known RCL demand model inspired by Nevo (2000). The model generates quantities and supplies benchmark own-price elasticities, holding rival prices and market conditions fixed. Using DML, we estimate price responses from log quantities and observable product–market variables. We consider the following specification. We first constructed the baseline specification, which is linear in log own price, with a slope that varies with product characteristics, income, and rival prices. Its implied elasticity, therefore differs across products and markets, but remains constant as the own price changes with these other variables held fixed. Because RCL elasticities generally also vary with their own price, we augmented the baseline specification by constructing a richer specification that includes non-linear price terms and their interactions with product and market characteristics. We then perform the steps described in the main text for computing elasticities in DML: namely, we perform the residualization step of DML by removing the variation predicted by controls using models trained on other markets when cross-fitting, then regress the residualized outcome on the residualized dictionary entries. Differentiating the fitted response with respect to log own price, holding rival prices and non-price characteristics fixed, gives the estimated own-price elasticities.

Figure G.1 shows the results when orthogonal OLS, orthogonal lasso, and the ridge-adjusted one-step estimator (labeled “Debiased LASSO”) are adopted in the second step of the DML procedure described in the main text. The figure also includes an oracle benchmark, which uses the true RCL elasticities known in the simulation. We differentiate the dictionary with respect to log own price to obtain the terms entering the elasticity formula, then choose their coefficients to fit the true elasticities directly^{G.1}. This checks how well the dictionary can represent the elasticity profile when that profile is known, helping distinguish limitations of the dictionary from difficulties in learning elasticities from quantities. Observations are sorted from most negative to least negative true elasticity and grouped into bins. Each curve reports averages within the same bins; the horizontal axis records position in this ordering. Lasso and the one-step estimator closely follow the true RCL curve, while OLS shows larger discrepancies. These results support recovery of the average elasticity profile in this design and illustrate the practical relevance of Proposition 2.1.

^{G.1}As an example, following the notation of Proposition 2.1, consider a dictionary containing only p_{jm} and p_{jm}^2 . The fitted log-quantity response is $\gamma_{D,1}p_{jm} + \gamma_{D,2}p_{jm}^2$, whose derivative with respect to log own price is $\gamma_{D,1} + 2\gamma_{D,2}p_{jm}$. The oracle fit therefore regresses the known RCL elasticities on 1 and $2p_{jm}$, choosing γ_D^{oracle} to approximate those elasticities directly. Feasible DML instead estimates $\hat{\gamma}_D$ from log quantities and then computes the implied elasticities.

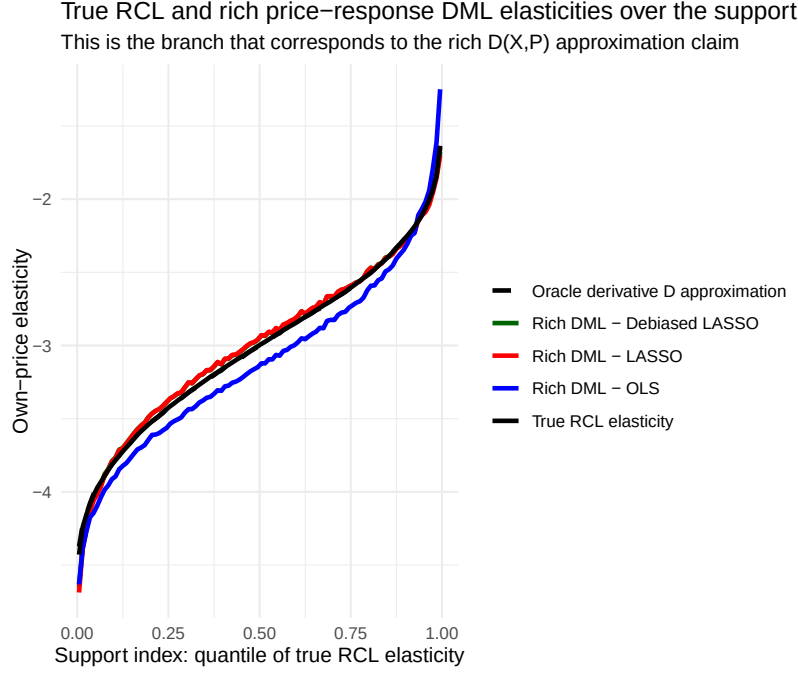


Figure G.1. Replication exercise: true RCL and DML-implied own-price elasticities. The figure compares the true RCL own-price elasticities with the elasticities recovered by the rich DML dictionary in its orthogonal OLS, orthogonal lasso, and ridge-adjusted one-step specifications. Observations are sorted by the true elasticity and averaged within support bins. The oracle derivative-dictionary curve is included only as a diagnostic benchmark for the approximation capacity of the chosen dictionary.

G.2 Simulation results: Proposition 2.2

The second simulation illustrates the rank failure in Proposition 2.2 and examines whether regularization helps recover price responses when the true demand model is sparse. We generate 480 product-market observations from a known RCL model, so its own-price elasticities provide a benchmark. We assumed sparsity in the sense that only a subset of the 700 candidate characteristics enters true mean utility with a coefficient different from 0. Nevertheless, the researcher does not know the model ex-ante and as a consequence decides to be agnostic by including in the unrestricted, i.e., non-regularized, specification all 700 characteristics and price, giving 701 linear coefficients for 480 observations. Full implementation details are provided in the Online Appendix.

In the simulated designs, this unrestricted specification is flexible enough to reproduce the observed shares exactly with many different coefficient vectors. We construct two such vectors by changing the mean-price coefficient and adjusting the characteristic coefficients to preserve the same fit at every observed price. Both solutions attain the same zero sample GMM objective, but imply different responses when own price changes while characteristics and rival prices remain fixed. Their elasticity curves therefore illustrate the central consequence of the proposition: the unrestricted sample criterion cannot select a unique price response.

We compare these solutions with two structural estimators. The oracle estimator is told which character-

istics enter the true model and estimates their coefficients together with the price coefficient. The regularized estimator, discussed in Appendix F.1, considers all candidate characteristics, uses the demand moments to select regressors, and then re-estimates the retained coefficients without regularization. The idea is to assess what regularization can achieve without knowing the relevant characteristics in advance. For all structural curves, the parameters governing consumer heterogeneity (θ_2) are fixed at their true values, isolating the recovery of the linear coefficients.

The two panels of Figure G.2 consider different price-generating processes. In Panel A, prices are exogenous, i.e., independent of the unobserved demand disturbance conditional on observed characteristics. In Panel B, that disturbance also enters the price equation, making price endogenous. Cost shifters that affect price without entering utility provide excluded instruments in this panel.

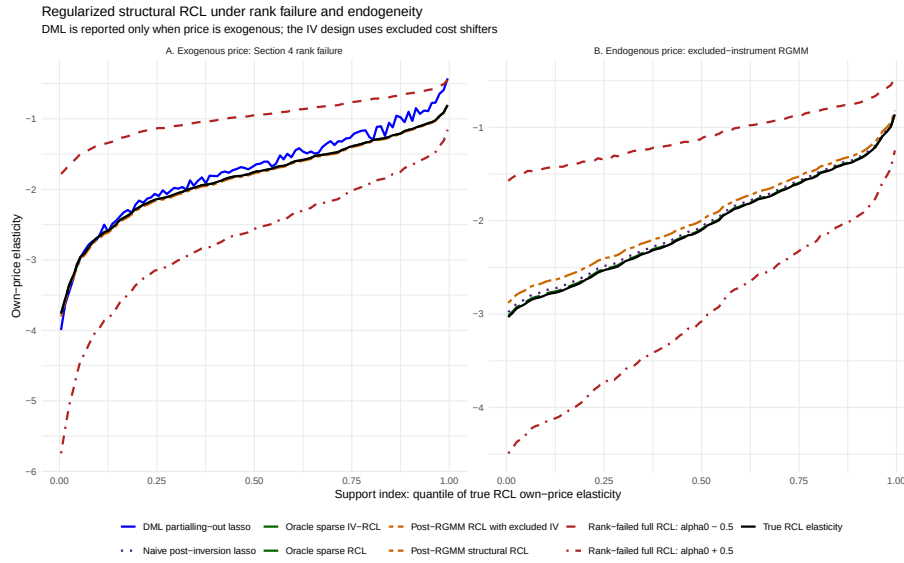


Figure G.2. Rank failure and regularization in an RCL model with many candidate characteristics. The black curve is the true elasticity benchmark. The two red curves correspond to unrestricted exact-fit solutions whose mean-price coefficients are 0.5 below and above the true value. Both panels include oracle and regularized structural estimates. Panel A additionally reports DML; Panel B reports a naive lasso comparison without excluded instruments. Within each panel, observations are ordered from the most negative to the least negative true elasticity, and all curves are averaged within the same 100 bins. The horizontal axis indicates position in this ordering.

In Panel A, the oracle and regularized structural curves closely follow the true elasticity benchmark. DML also captures the overall pattern, with larger discrepancies where the true elasticities are closest to zero. This estimator obtains the price response directly from log shares and observable variables, without using the known active characteristics or the RCL heterogeneity parameters.

In Panel B, the oracle and regularized structural estimators use instrumental-variable moments and remain close to the benchmark. The naive lasso curve provides a comparison that applies regularization without using the excluded instruments. Its inclusion distinguishes the task of selecting relevant regressors from that of addressing price endogeneity. The DML specification used here requires conditional price exogeneity

and is therefore reported only in Panel A.

Overall, Figure G.2 shows that exploiting sparsity can help recover elasticities even when the unrestricted sample criterion admits different price responses. Its interpretation has two qualifications. First, agreement between curves concerns bin averages and can conceal differences at individual observations. Second, all structural elasticities are evaluated using the choice probabilities recovered from observed shares and the known consumer heterogeneity.

To summarize and enforce the main message again, with consumer heterogeneity held fixed, the structural elasticity curves differ because their mean-price coefficients differ. For the two unrestricted solutions, this illustrates the implication of Proposition 2.2: the same observed shares and sample demand moments can support different price responses^{G.2}. Forcing the regularization of the moment conditions arising from RCL seems to resolve the latter ambiguity.

H Further validity checks on estimated elasticities

This appendix assesses the robustness of the DML estimates through placebo stress tests and examines whether the estimated elasticities reflect systematic price responses and are economically plausible. In particular, in Subsection H.1, we first check for spurious responses by repeating the estimation after randomly reassigning quantities across product observations from different molecule–country cells. In such a case, we expect a non-significant result for all the reshuffled molecules.

We then show – in Section H.2 – the corresponding of Figure 2 when elasticities are aggregated at the molecule level providing some evidence of their robustness.

Finally we perform further descriptive analyses in Subsection H.3, which aim to assess whether the estimated elasticities replicate known patterns documented in the pharmaceutical literature.

H.1 Permutation placebo exercise

We conduct a permutation placebo exercise to assess whether the estimation procedure generates systematic non-zero price-response estimates after the observed matching between quantities and prices has been deliberately broken, i.e., after we randomly reassign molecule–countries’ quantities to other molecule–countries’ prices. For this exercise, recall that y_{igt} denotes log standard units for product i , molecule g , country c , and period t . p_{igt} denotes log own price and p_{igt}^- is the leave- i -out rival-price index constructed from the other products in the same country–molecule–period cluster $\mathcal{C}_{gct}$. The baseline price-response equation is given by Equation (1).

We construct the placebo outcome by randomly reassigning quantity observations across different molecule–country cells while keeping the quarter fixed. Specifically, an observation for product i , molecule g , and country c in quarter t receives the log quantity of a randomly assigned source observation for product i' , molecule g' , and country c' in the same quarter: $y_{igt}^{\text{pl}} := y_{i'g'c't}$. Here the primed indices identify the source observation supplying the replacement quantity. The source and destination molecules, g' and g , belong to therapeutically unrelated ATC classes, and the source and destination countries, c' and c , have substantially different economic profiles^{H.1}. Germany and Ecuador provide one such country pair.

^{G.2}Matching the true elasticity curve, however, does not mean that all characteristics coefficients have been correctly estimated. The exercise also leaves open how well the estimators would perform if the parameters governing consumer heterogeneity had to be estimated.

^{H.1}The reassignment was conducted mainly at random across countries located on different continents. Once a pair of countries was selected, we randomly reshuffled the molecules and reassigned the observations accordingly.

The own price p_{igct} , rival-price index p_{igct}^- , conditioning set $\mathcal{W}_{igct}$, and heterogeneity dictionary K_{igct} remain attached to the destination observation. Thus, basically the placebo pairs quantities from one molecule–country cell with prices and covariates from another cell in the same quarter. We use the superscript pl to denote objects associated with the reshuffled outcome: b_0^{pl} is the coefficient vector, g_0^{pl} the nuisance function, and u_{igct}^{pl} the disturbance. As in the baseline model, b_0^{pl} is partitioned into own-price and rival-price blocks, $b_0^{\text{own,pl}}$ and $b_0^{\text{rival,pl}}$.

In order to estimate elasticities after the discussed reshuffling, we applied the same residualization procedure, cross-fitting scheme, heterogeneity dictionary, and final-stage estimators used in the baseline analysis. Specifically, we estimate the placebo version of the baseline model:

$$y_{igct}^{\text{pl}} = D_{igct}' b_0^{\text{pl}} + g_0^{\text{pl}}(\mathcal{W}_{igct}) + u_{igct}^{\text{pl}},$$

where the technical-treatment vector remains $D_{igct} = \begin{pmatrix} p_{igct} K_{igct} \\ p_{igct}^- K_{igct} \end{pmatrix}$. In the second stage, for estimating elasticities, we used only the ADL method and the resulting placebo elasticities are denoted as $\hat{\epsilon}_{gca}^{\text{own,pl}} = K_{igct}' \hat{b}^{\text{own,pl}}$, $\hat{\epsilon}_{gca}^{\text{rival,pl}} = K_{igct}' \hat{b}^{\text{rival,pl}}$, where $\hat{b}^{\text{own,pl}}$ and $\hat{b}^{\text{rival,pl}}$ are the own-price and rival-price blocks of the estimated placebo coefficient vector. The estimator label is suppressed for readability [H.2](#).

We finally checked whether the estimation procedure produces systematically non-zero price-response estimates after the original matching between quantities and prices has been disrupted. Figure [H.1](#) reports the results for Germany and Ecuador as an example. Across the displayed placebo estimates, 96% are not statistically different from zero at the 5 percent level and are shown as missing. This high non-rejection rate is consistent with the absence of a systematic placebo signal.

We repeat the quantity-reassignment exercise for several other deliberately "dissimilar" country pairs and obtain qualitatively similar results [H.3](#). For brevity, we report only the Germany–Ecuador exercise.

^{H.2}Furthermore, as in Section 2.1, $a = a(t)$ denotes the calendar year containing quarter t , and K_{igct} is common within each country–molecule–year cell. The model, therefore, assigns common placebo elasticities to observations within that cell.

^{H.3}These are available upon request.

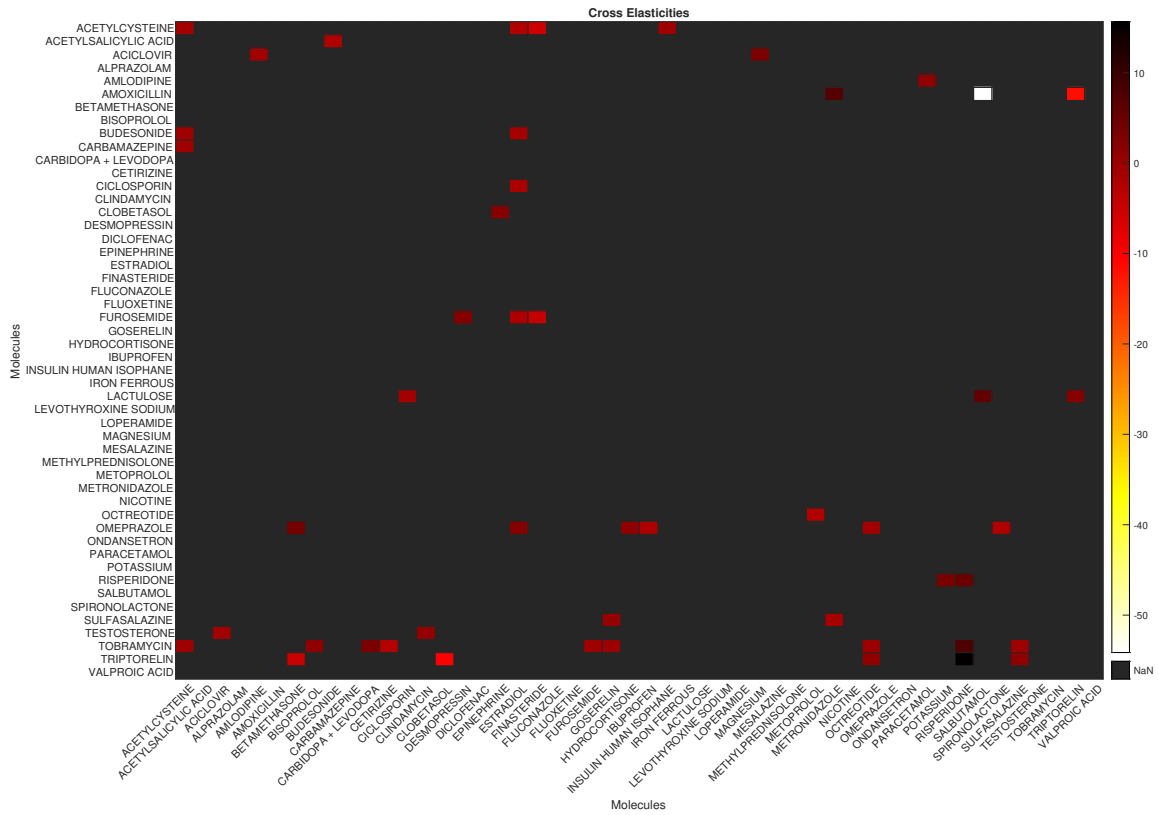


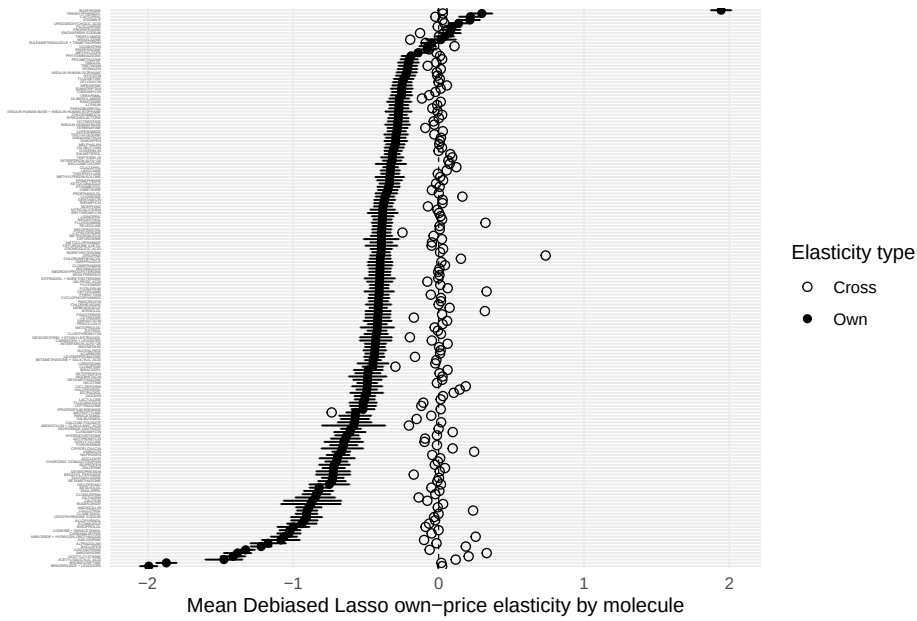
Figure H.1. Permutation-placebo estimates for Germany and Ecuador. Rows and columns identify the molecules paired in the placebo exercise. The heatmap reports price-response estimates obtained by applying the baseline DML specification to the permuted outcome. Estimates that are not statistically different from zero at the 5 percent level are shown as missing. The molecule pairs identify the placebo reassignments.

H.2 Price elasticities by molecule

Figure H.2 reports molecule-level averages of the own- and cross-price elasticities estimated using ADL. To reduce differences arising from the composition of molecules, Figure H.2b repeats the analysis using a common set of molecules observed in every country throughout the sample period.

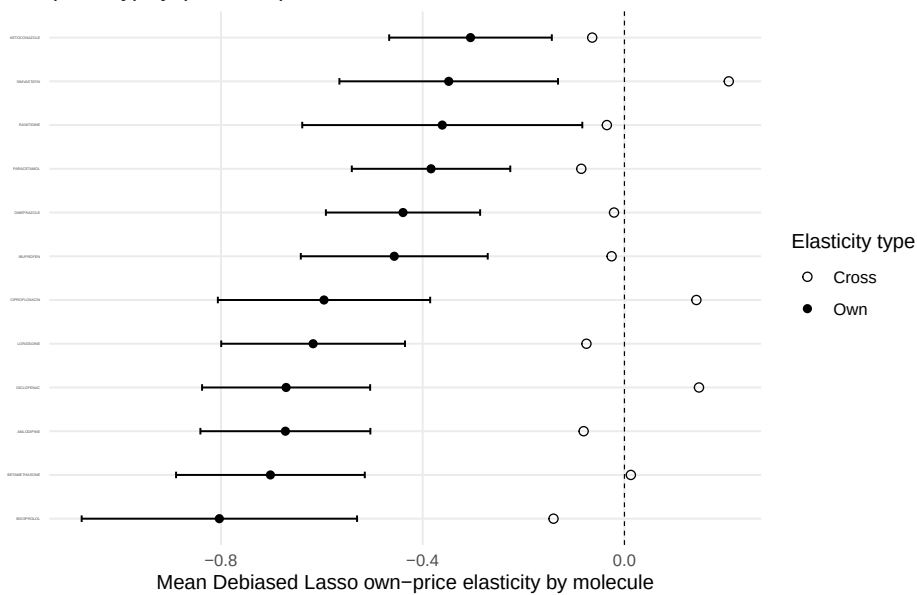
Figure H.3 uses the latter to examine whether the same molecules repeatedly display particularly high or low rival-price-index elasticities across countries. For each country, we first average each molecule's elasticity over time and identify the molecules in the top 5% and bottom 5% of the resulting distribution. We then rank molecules by the number of countries in which they appear in each tail (the number inside the circles of Figure H.3). Large positive elasticities are consistent with stronger substitution between a molecule's products and the competing products represented in the rival-price index, whereas large negative elasticities are consistent with stronger complementarity.

Own-price elast. (with CI) and cross elast. (overlay) by molecule



(a) Baseline DML sample.

Own-price elast. (with CI) and cross elast. (overlay) by (common) molecule



(b) Common-molecule restricted sample.

Figure H.2. Price elasticities by molecule. The figure reports molecule-level mean own-price elasticities from the baseline ADL specification. Each row corresponds to a molecule. White diamonds denote the baseline DML point estimates, and horizontal bars show 95 percent confidence intervals based on clustered standard errors. The vertical dashed line marks zero. Molecules are ordered by the baseline point estimate, with more negative values indicating greater price responsiveness.

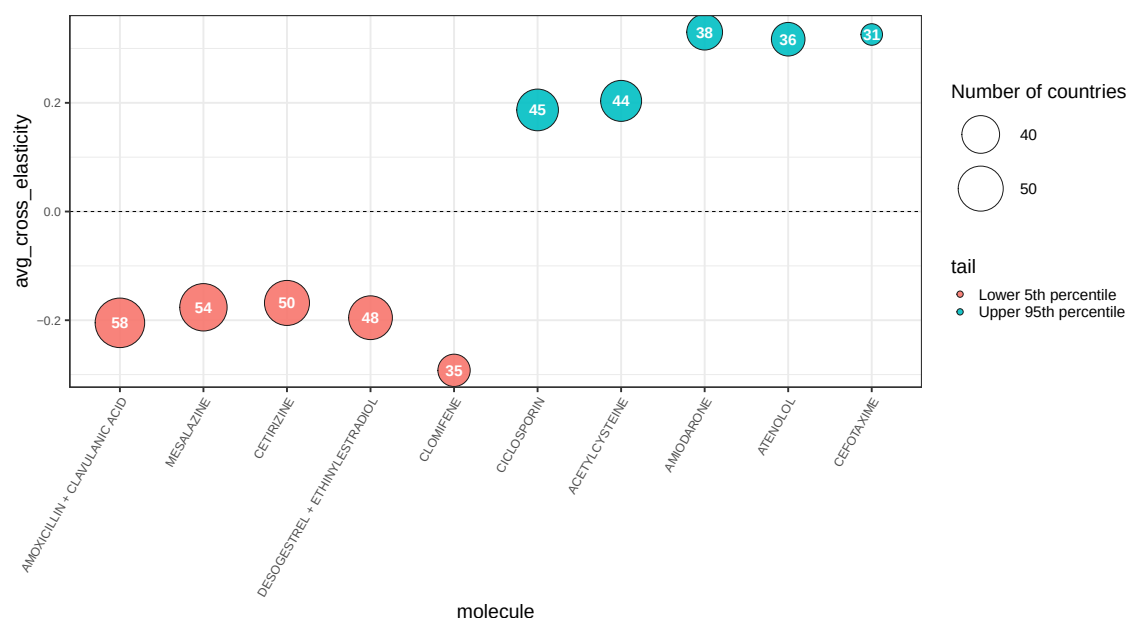


Figure H.3. Top molecules in the negative and positive tails of the rival-price-index elasticity distribution. The vertical axis reports the average rival-price-index elasticity across countries in which the molecule belongs to the relevant tail. Molecules are ordered from the most negative to the most positive average rival-price-index elasticity. Marker size is proportional to the number of countries in which the molecule appears in the lower 5th percentile or upper 95th percentile of the country-specific rival-price-index elasticity distribution.

For molecules in the positive tail (shown in light blue), the estimates suggest strong substitution between their products and the competing products represented in the rival-price index. One possible explanation is competition between branded and generic versions of the same molecule, which can be close substitutes depending on the regulatory setting and the choices of patients and prescribers (Griliches Z., 1993; Brekke et al., 2007; Dubois and Lasio, 2018; Castanheira et al., 2019). These molecules have fewer branded products than the average molecule in the market, but a similar number of generics.^{H.4} This composition may indicate that brand differentiation plays a smaller role in these markets, making relative prices more important in product choice. Product counts alone, however, cannot establish this mechanism. Related studies document generic uptake in markets for cephalosporins and antihypertensive drugs (Ellison et al., 1997; Shrank et al., 2011; Branstetter et al., 2016).

The negative-tail molecules (shown in red) require a more cautious interpretation. Their negative rival-price-index elasticities are consistent with complementarity, but do not by themselves establish that the products are economic complements. These molecules have substantially more branded products and fewer generics than the average molecule^{H.5} though this evidence is not conclusive on complementarity since generics may complement branded as well. The literature also points to brand loyalty as an alternative channel. Patients may be reluctant to switch to a cheaper generic because they are uncertain about its effects

^{H.4}These descriptive results are available upon request.

^{H.5}These descriptive results are available upon request.

or prefer to continue using a familiar product. For some of the products contained in the molecules of the negative-tail, moreover, prescribers' concerns about generics can further discourage switching (Janssen and Granlund, 2023; Thaver et al., 2020; Lundin, 2000; Carrera and Villas-Boas, 2023).^{H.6}

Similarly, generic adoption may also differ across patients: those who are more sensitive to price have a stronger incentive to try cheaper generics, while less price-sensitive patients may continue using branded products (Ching, 2010). Product characteristics provide a further consideration, since medicines containing the same molecule can differ in their delivery systems^{H.7}. Mesalazine formulations, for example, differ in where and how the active ingredient is released in the gastrointestinal tract, which may influence treatment choice and substitution across formulations (Ye and van Langenberg, 2015).

While we are aware that these findings help explain why products containing the same molecule may be imperfect substitutes, but do not directly establish complementarity, to our knowledge, evidence addressing these negative responses for the specific molecules identified here, in settings comparable to ours, remains limited.

H.3 Additional elasticity-related descriptives

In what follows, we use the elasticities estimated via DML – with ADL in the second stage – to provide additional descriptive evidence on whether the estimated price responses display patterns consistent with established features of pharmaceutical markets. We first compare own-price elasticities between branded and generic products. This comparison is motivated by evidence that generic markets tend to feature closer within-molecule substitutes and stronger price competition relative to branded products (Danzon and Chao, 2000b; Granlund and Bergman, 2018). Since median elasticities of branded drugs may be driven by the entry of generics when their patents expire, we consider a related comparison at the molecule level, distinguishing molecules whose products remain predominantly branded, with little or no generic competition for most of the sample period, from molecules whose branded products present, instead, a conspicuous number of generic competitors.

Figure H.4 presents the former exercise displaying the yearly average price-elasticity of generics and branded in the whole sample.

^{H.6}Thaver et al. (2020), for instance, documents reluctance to prescribe generic oral contraceptives among clinicians who are skeptical about generics, while Carrera and Villas-Boas (2023) emphasizes how generic aversion takes time to fade. This aversion may lead to persistence in branded choice, inducing the mechanisms described in Lundin (2000)

^{H.7}In our specification, we are indeed not able to distinguish between release mechanisms and delivery systems.

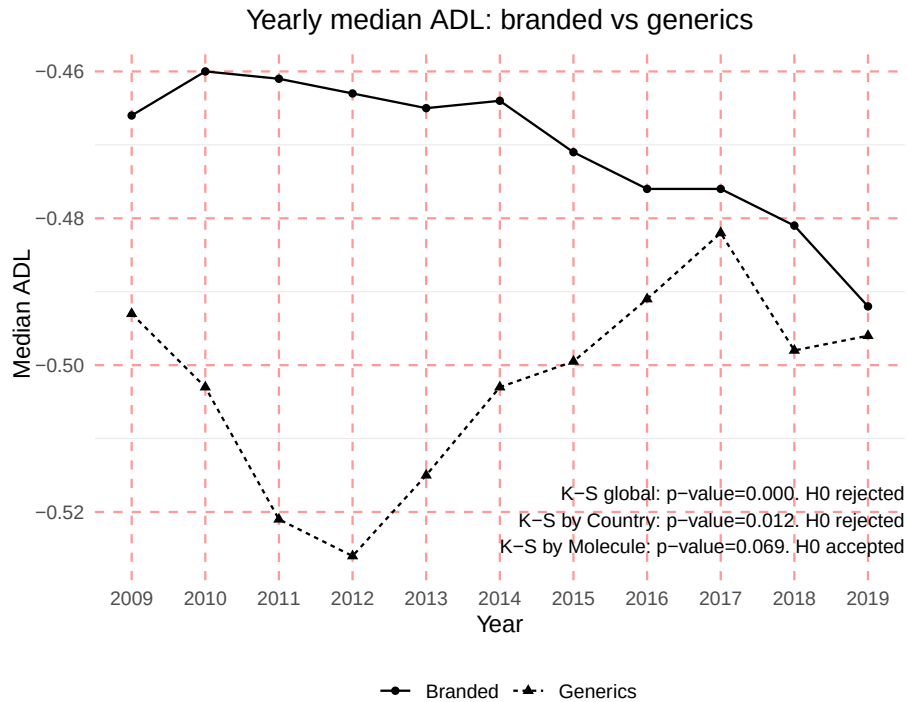


Figure H.4. Yearly median own-price elasticities for branded and generic products. The figure plots, for each year, the median own-price elasticity estimated by ADL separately across branded and generic product observations. More negative values indicate greater estimated price responsiveness. The annotation reports Kolmogorov–Smirnov tests comparing the branded and generic elasticity distributions in the pooled sample, across country-level medians, and across molecule-level medians. Equality of the distributions is rejected in the pooled sample ($p < 0.001$) and across country-level medians ($p = 0.012$), but not across molecule-level medians at the 5 percent level ($p = 0.069$) (only at the 10%).

As expected, every year, the absolute value of the median price elasticity is bigger for generic than for branded products, indicating greater estimated price responsiveness for the typical generic observation. The reported Kolmogorov–Smirnov tests reject equality of the branded and generic distributions in the pooled sample and across country-level medians, but not across molecule-level medians at the 5 percent level. The convergence observed in later years may partly reflect increasing exposure of branded products to generic competition as patents expire and generic entry occurs. In particular, products that lost exclusivity earlier in the sample may have faced generic competition for a longer period, while additional branded products may have become exposed to such competition toward the end of the sample. However, this pattern may also reflect changes in product composition or other market conditions over time.

To explore these possibilities further, we examine (if and) how the estimated elasticities of branded incumbents change in response to generic entry. The intuition is that the arrival of a generic version gives patients an additional, typically lower-priced alternative containing the same active ingredient. This creates another opportunity to substitute between products in response to relative prices. We therefore examine the responsiveness of branded demand to competitors’ prices, as captured by the rival-price-index elasticity.

Specifically, we focus on branded products already present before the first observed generic entry in each country–ATC-4–molecule market. We compare their estimated rival-price-index elasticities before and after entry, accounting for country–year variation, and plot their evolution relative to the first observed entry year. Figure H.5 shows greater estimated responsiveness to the rival-price index after generic entry. This pattern is consistent with the idea that the availability of a close generic substitute strengthens the connection between branded demand and competing products’ prices, thus providing additional descriptive support to the above interpretation of the convergence of the distributions in the latest years of Figure H.4.

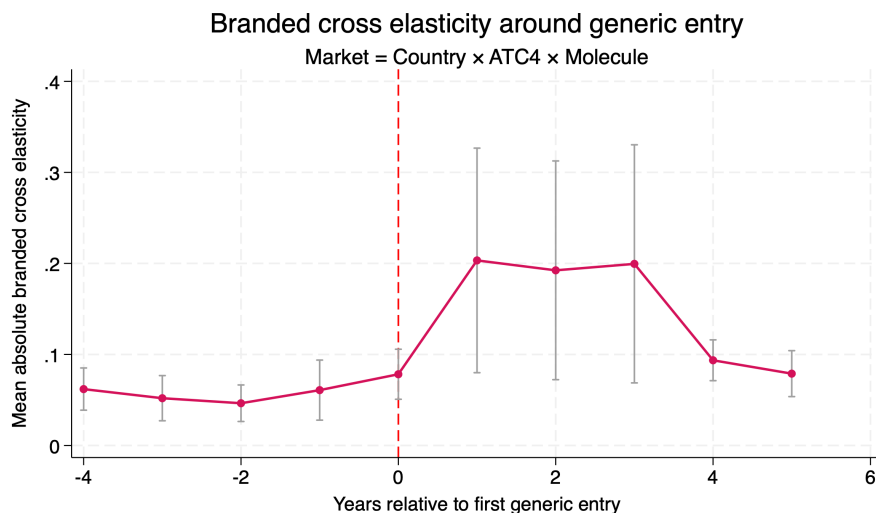


Figure H.5. Branded rival-price-index elasticities around generic entry. The figure plots the average absolute value of the rival-price-index elasticity of branded products by event time, where event time zero is the first year in which a generic product is observed in the same country–ATC-4–molecule market. The vertical dashed line marks the generic-entry year. Error bars report approximate 95 percent confidence intervals around event-time means.

We complemented the graphical evidence with a simple regression that relates the estimated rival-price-index elasticity of branded products to an indicator for whether a generic has already entered the same country–ATC-4–molecule–year market. As observed descriptively in the Figure, once a generic enters the market, the branded product faces a closer substitute, so its demand should become more responsive to rivals’ prices ^{H.8}. Consistent with this idea, the coefficient on the post-generic-entry indicator is positive: after generic entry, branded rival-price-index elasticities are higher by about 0.066 elasticity points, controlling for country–year shocks. Relative to the average pre-entry rival-price-index elasticity of about 0.010, this corresponds to an increase to roughly 0.076, or about 7.6 times the pre-entry level. Equivalently, the post-entry elasticity is significantly about 0.066 points higher, suggesting that branded demand becomes more sensitive to competing products’ prices after generics enter the market.

^{H.8}Related evidence comes from Gugler et al. (2026), who find that own-price demand becomes more elastic following clopidogrel generic entry in Austria.

I Details of the Structural RCL Benchmark

This appendix describes the adjustments underlying the comparison in Section 4.1. We retain the structural notation of Section 2.2, but define markets for this benchmark as $m = (c, h, t)$, where c denotes country, h the ATC-4 therapeutic class, and t the year-quarter. This definition follows Dubois et al. (2022) and replaces, for this exercise, the country-molecule-quarter grouping used in the baseline analysis.

As is standard when computing RCL price elasticities, we hold each market’s potential size fixed while allowing product shares and the outside-option share to respond to prices. Notice that, under this convention, Equation (5) implies that share and quantity elasticities coincide.

To make the DML estimates more comparable with the RCL benchmark, we broaden the set of competitors used to construct the DML rival-price index. Competitors now include all other products sold in the same country, ATC-4 therapeutic class, and quarter, rather than only products containing the same molecule as in the main specification. Differences in product aggregation and demand modeling remain. Dubois et al. (2022) aggregate sales by molecule and corporation within each market, whereas our DML analysis uses more detailed product-level observations. Their RCL model describes buyers’ choices among the available medicines and the option of purchasing none of them. These choices depend on prices and product characteristics, including patent and generic status. To address price endogeneity, they use instruments describing the number and composition of competing products. Their analysis focuses on hospital purchases, treating hospitals as informed purchasers who internalize drug prices. Our DML estimates instead come from a semiparametric log-log specification, and both our DML and RCL estimates use data that do not separately identify hospital and retail sales. They therefore describe demand across these channels. Section 4.1 discusses the implications of these differences for interpreting the comparison.

DML implementation for comparison with Dubois et al. (2022). For implementing DML in the robustness checks of Section 4.1, we retain the log-log specification introduced in Section 2.1, applying the rival-price-index formula to the broader set of competitors defined above (country – ATC-4 – time). The country index is removed because we only compare the elasticities of Dubois et al. (2022) in the U.S., as explained in the main text.

Estimation follows the residualization and time-block cross-fitting procedure described in Section 3. First, the nuisance regressions remove the components of log quantities and prices predicted by the controls in $\mathcal{W}$. These controls capture product and market characteristics, time effects, lagged information, and the competitive environment. Foreign-price information also enters this prediction step.

The second stage uses the remaining variation to estimate price responses. Here, the dictionaries $K^{US,LD}$ and $K^{US,HD}$ determine how those responses can differ across observations, through interactions with residualized prices. As detailed in Section 4.1, the superscript "LD" is for "low-dimensional" specification. The latter allows elasticities to differ across the ATC-4 $\times$ branded/generic groups used for reporting the comparison. The high-dimensional specification – denoted with the superscript "HD" on the corresponding dictionary – allows additional variation within those groups.

As explained in the main text, foreign-price and competitive-environment variables serve as predictors in this DML exercise, rather than as excluded instruments. Their identification, therefore, continues to rely on the conditional-exogeneity restriction in Equation (3).

Table 2 compares the RCL estimates from Dubois et al. (2022) with DML estimates.

Our RCL vs Dubois et al. (2022) Using the benchmark markets defined above, we estimate an RCL model that closely follows the demand specification of Dubois et al. (2022). Mean utility includes log prices, generic status, molecule fixed effects, ATC-4-by-year fixed effects, and quarter fixed effects. Using

log prices is particularly appropriate here because pharmaceutical prices differ by orders of magnitude across therapeutic classes.

To identify price effects, we follow their approach of using changes in the set of available products as a source of variation in competitive conditions. We construct instruments from the variables available in our data, considering product and generic counts at different therapeutic and molecule levels. After accounting for the fixed effects and assessing the empirical relevance of the candidate instruments, we retain two excluded instruments: the number of products and the number of generics containing the same molecule. These counts reflect changes in competition associated with entry, generic availability, and product turnover. We first estimate the model using these instruments, then use the resulting estimates to construct Chamberlain-style optimal instruments Reynaert and Verboven (2014). As in Dubois et al. (2022), we re-estimate the model with these instruments to improve precision and numerical stability.

The data report quantities sold, but neither the potential market size nor the outside share. We therefore construct market shares using a market-size calibration procedure following Huang and Rojas (2013) and Huang and Rojas (2014), as implemented in Dubois and Lasio (2018). We first estimate an auxiliary logit specification using differences between products within the same market. These differences remove the common outside-good term, allowing the price and generic-status coefficients to be estimated without first specifying potential market size. We then choose market size for each ATC-4 class so that the corresponding coefficients estimated from the logit equation with an outside option match these initial estimates as closely as possible. This calibration provides the product shares and implied outside share needed for RCL estimation, without fixing an arbitrary outside share in advance.

We do not use micro-moments for two reasons. First, we seek to maintain comparability with Dubois et al. (2022), who do not use them. Second, the publicly available statistics that could inform these moments generally do not match our market definitions or quantity units. They are typically reported at the ATC-3-by-year level in patients or prescription counts, whereas our benchmark uses country-ATC-4-quarter markets and quantities measured in standard units.

The sales-channel difference discussed above also matters for interpreting the estimates. Hospital purchasers bear the cost of the drugs they buy and can be viewed as relatively informed, price-conscious decision makers. Retail demand reflects a more complex combination of patient preferences, prescribing decisions, insurance coverage, formularies, and cost-sharing arrangements. Patients may face copayments rather than the full drug price, and brand loyalty or persistence in treatment choices may further limit their response to price changes. A sample combining hospital and retail sales like ours may therefore produce lower own-price elasticities in absolute value than a hospital-only sample. Evidence that prescription drug utilization responds to insurance design and cost sharing support the relevance of these channels (Goldman et al., 2007; Duggan and Morton, 2010; Einav et al., 2015).

J The inverse elasticity rule with continuous choice

The canonical pharmaceutical differential-pricing applications on which our interpretation relies, formulate the inverse-elasticity problem in terms of differentiable market-demand schedules and demand elasticities (Ramsey, 1927; Baumol and Bradford, 1970; Danzon, 1997; Danzon and Towse, 2003; Barros and Martinez-Giralt, 2008). In these models, prices across markets are chosen subject to a revenue or cost-recovery requirement, and the resulting first-order condition links price-cost margins to demand responsiveness. Notably these models adopt a continuous-quantity or differentiable aggregate-demand formulation. Showing the consequences of using such continuous quantity (differentiable aggregate-demand) formulation is important in our context since it allows to add – at the same time – three elements that are relevant to the

cross-country empirical exercise. First, non-homothetic demand that separates income or expenditure effects from price responsiveness. Second, insurance and reimbursement wedges that can alter the price faced by the decision maker and the resulting elasticity. Third, country-specific constraint tightness which implies that the welfare consequences of Ramsey pricing depend jointly on demand elasticity and on the shadow value of the financing or coverage constraint.

Namely, in the present Appendix we will adopt a non-homothetic CES (NH-CES) framework for demand. Relative to the RCL model, the NH-CES imposes non-homotheticity directly instead of allowing heterogeneous tastes across consumers and absorbing rich fixed effects. While in RCL substitution arises from the structure of discrete choice with heterogeneity, in NH-CES substitution patterns are described via a constant-elasticity aggregator.

The present Appendix focuses on the following. First, a common shortcut in reduced-form cross-country analyses is to use income measures (e.g. GDP per capita) as a proxy for price sensitivity or demand elasticities (Danzon and Towse, 2003).^{J.1} However, GDP per capita is not, in general, a reliable proxy for the relevant elasticity once insurance wedges and non-homothetic preferences shape demand.

Moreover, our theoretical results highlight an identification challenge for welfare statements based on such coarse proxies. In particular, when the proxy co-moves with the country-specific tightness of the budget/coverage constraint (the Ramsey shadow cost), reduced-form comparisons cannot separately identify the role of demand elasticity from constraint tightness in shaping the efficiency of Ramsey pricing (as captured by deadweight-loss shares). Accordingly, our empirical strategy conditions directly on estimated elasticities and absorbs part of the unobserved constraint tightness through rich fixed effects.

Modeling entry and exit is out of the scope of the present paper.

J.1 Intensive margin

In this section, we model consumer behavior within a period since, conditional on period t prices and coverage, choices reflect current tastes; common time shocks are absorbed by time fixed effects. We focus on how individuals simultaneously allocate their income across branded drugs, generic alternatives, and a numeraire good. Moreover, consumers are assumed to face relative prices and a budget constraint, and make consumption choices that maximize utility. To reflect realistic expenditure patterns and preference heterogeneity across income groups, we employ a non-homothetic CES framework inspired by Comin et al. (2021) and Hanoch (1975). We believe that the choice of a NH-CES specification—allowing branded-generic budget shares to vary with utility through $f_k(U_l)$ and with insurance through effective prices—to be consistent with some stylized facts in the pharmaceutical industry. For example macro evidence point to the non-homotheticity of health spending—its share rises with income (Hall and Jones, 2007). Similarly, micro evidence suggest that coverage design shifts the composition of drug use between branded and generics (Dafny et al., 2017).

Consumers are divided into groups denoted by l . This heterogeneity allows to construct relative results for each group. It is for instance easy to see that the definition of a good as luxury or necessity in the NH-CES specification below is l dependent.

To simplify the analysis, we assume a group-specific but type-invariant coverage rate δ_l , so that

$$\tilde{P}_{k,l} = (1 - \delta_l)P_k, \quad P_k > 0, \quad 0 \leq \delta_l < 1,$$

^{J.1}The empirical and institutional link between income and price responsiveness is not mechanical and may vary with coverage, reimbursement rules, and other country-specific institutions.

and relative prices between B and G are unaffected by insurance within group l . The coverage rates are therefore free to vary across income groups l through $\tilde{P}_{k,l}$. To derive the demand functions and expenditure shares, Comin et al. (2021) consider the consumer's utility maximization problem, where preferences are represented by a non-homothetic CES (NH-CES) utility aggregator:

$$1 = \sum_{k \in \{B, G, y\}} (f_k(U_l))^{\frac{1}{\sigma'}} C_{k,l}^{\frac{\sigma'-1}{\sigma'}},$$

where $k \in \{B, G, y\}$ indexes a representative drug of type (B)—i.e. branded—, a representative drug of type (G)—i.e. generic—and a numeraire good (y); $f_k(U_l)$ is an income-dependent taste parameter for group $l \in \{L_R, L_P, H_R, H_P\}$ —that is, $f_k(U_l)$ and subsequent quantities are specified at the group level, representing the behavior of a representative consumer in each consumer group. We therefore have a system of equations, one for each consumer group l . Notice that the following results hold also when k varieties are defined at a more granular level (e.g. ATC-4 groups).

In the following, each group l solves its own problem independently. Group-specific allocations may differ through coverage, total drug expenditure, and the utility-dependent taste terms $f_k(U_l)$. $\sigma' > 1$ represents the constant elasticity of substitution ^{J.2}. Since the CES block is defined separately for each subgroup l , σ' is the elasticity of substitution only between goods $k \in \{B, G, y\}$ within a given l . Quantities with different l belong to different representative consumers and are not traded off in a single optimization. In particular, for instance, “ B for L_R ” and “ B for L_P ” are not substitutes.

Imposing a single elasticity $\sigma' > 1$ makes the compensated substitution patterns among all goods in the block depend on the same parameter (here: B with G , B with y , and G with y). This may be restrictive when drug spending should not substitute closely with other consumption. We therefore assume weak separability between pharmaceuticals and the numeraire. Thus, exploiting weak separability, we restrict attention to the drugs block only and model the allocation across $k \in \{B, G\}$, since the final composition of the consumer's consumption bundle, including the numeraire good y , is not central to our analysis. For each group l , let $E_{D,l} > 0$ denote realized total drug expenditure in the conditional Marshallian within-drugs allocation; it is taken as given when the representative consumer allocates this expenditure between B and G . Preferences over B and G are represented by the following NH-CES normalized aggregator where the normalization is a choice of units (“one unit of drug utility”). It only fixes the scale of the utility index so that prices and the taste terms $f_k(U_l)$ have consistent units. Marginal rates of substitution and optimal choices are unaffected by that normalization:

$$1 = \sum_{k \in \{B, G\}} f_k(U_l)^{1/\sigma_{BG}} C_{k,l}^{(\sigma_{BG}-1)/\sigma_{BG}}, \quad \forall l \in \{L_R, L_P, H_R, H_P\}, \quad \sigma_{BG} > 1, \quad f_k(U_l) > 0. \quad (\text{J.1})$$

In order to keep the notation as clean as possible and given the above clarifications, we will avoid writing every time $\forall l \in \{L_R, L_P, H_R, H_P\}$ and rename σ_{BG} as σ only. Since $f_k(U_l) > 0$, the NH-CES implies an interior allocation across B and G : the representative consumer in l allocates a positive budget share to each. This should be interpreted as an aggregate allocation for group l —e.g., different individuals (or conditions) within group l are treated with G versus B .

A convenient way to derive compensated demand is to fix a utility target U_l and solve the expenditure-minimization problem

$$e_l(\tilde{P}_l, U_l) = \min_{\{C_{k,l}\}_{k \in \{B, G\}}} \sum_{k \in \{B, G\}} \tilde{P}_{k,l} C_{k,l}$$

^{J.2} Assuming a constant elasticity of substitution may seem restrictive, however when thinking about it as the average elasticity of substitution between the goods in the economy, the latter becomes more reasonable.

subject to Equation (J.1), where $\tilde{P}_l = (\tilde{P}_{B,l}, \tilde{P}_{G,l})'$. Following Comin et al. (2021), $f_k(\cdot)$ is written as a function of the utility index U_l . In this dual problem, U_l is fixed rather than chosen. The corresponding Lagrangian is

$$\mathcal{L}_l^e = \sum_{k \in \{B,G\}} \tilde{P}_{k,l} C_{k,l} + \rho_l \left[1 - \sum_{k \in \{B,G\}} f_k(U_l)^{1/\sigma} C_{k,l}^{(\sigma-1)/\sigma} \right].$$

The first-order condition for good k is

$$\tilde{P}_{k,l} = \rho_l \frac{\sigma-1}{\sigma} f_k(U_l)^{1/\sigma} C_{k,l}^{-1/\sigma}.$$

Writing $h_{k,l}(\tilde{P}_l, U_l)$ for Hicksian demand and multiplying this condition by $h_{k,l}$ gives

$$\tilde{P}_{k,l} h_{k,l} = \rho_l \frac{\sigma-1}{\sigma} f_k(U_l)^{1/\sigma} h_{k,l}^{(\sigma-1)/\sigma}. \quad (\text{J.2})$$

Summing Equation (J.2) across k and using the binding normalized aggregator yields

$$e_l(\tilde{P}_l, U_l) = \rho_l \frac{\sigma-1}{\sigma}.$$

Define the compensated expenditure share as

$$s_{k,l} := \frac{\tilde{P}_{k,l} h_{k,l}(\tilde{P}_l, U_l)}{e_l(\tilde{P}_l, U_l)}.$$

Dividing Equation (J.2) by compensated expenditure gives

$$s_{k,l} = f_k(U_l)^{1/\sigma} h_{k,l}^{(\sigma-1)/\sigma}. \quad (\text{J.3})$$

Combining this expression with the definition of $s_{k,l}$ gives Hicksian demand:

$$h_{k,l}(\tilde{P}_l, U_l) = \left[\frac{e_l(\tilde{P}_l, U_l)}{\tilde{P}_{k,l}} \right]^\sigma f_k(U_l). \quad (\text{J.4})$$

Taking logarithms,

$$\ln h_{k,l} = \sigma \ln e_l(\tilde{P}_l, U_l) - \sigma \ln \tilde{P}_{k,l} + \ln f_k(U_l). \quad (\text{J.5})$$

The compensated expenditure share is therefore

$$s_{k,l} = e_l(\tilde{P}_l, U_l)^{\sigma-1} \tilde{P}_{k,l}^{1-\sigma} f_k(U_l).$$

Finally, substituting Equation (J.4) into compensated expenditure yields

$$e_l(\tilde{P}_l, U_l)^{1-\sigma} = \sum_{k \in \{B,G\}} \tilde{P}_{k,l}^{1-\sigma} f_k(U_l). \quad (\text{J.6})$$

The compensated and conditional Marshallian problems hold different objects fixed. In the compensated problem, utility is fixed and minimum expenditure adjusts with prices. In the conditional Marshallian problem, total drug expenditure $E_{D,l}$ is fixed and utility adjusts with prices. Let $V_l(\tilde{P}_l, E_{D,l})$ denote the utility attained in the conditional Marshallian problem. Then

$$e_l(\tilde{P}_l, V_l(\tilde{P}_l, E_{D,l})) = E_{D,l},$$

and conditional Marshallian demand satisfies

$$C_{k,l}^M(\tilde{P}_l, E_{D,l}) = h_{k,l}(\tilde{P}_l, V_l(\tilde{P}_l, E_{D,l})).$$

Thus, $e_l(\tilde{P}_l, U_l)$ and $E_{D,l}$ are distinct objects: they coincide only when compensated demand is evaluated at the utility attained by the conditional Marshallian allocation.

In this Section the main theoretical findings are presented. We will denote with a superscript H the Hicksian elasticities and with a superscript M the conditional Marshallian ones.

J.2 GDP (per capita) as a measure of elasticity

GDP does not take into account insurance-induced distortion effects and non-homothetic consumer tastes. Moreover, GDP per capita is not the same economic object as total drug expenditure $E_{D,l}$. The NH-CES system therefore does not imply a direct mapping from GDP to price responsiveness. At a given allocation, the compensated own-price elasticity depends on the good's expenditure share, whereas the conditional Marshallian elasticity also depends on how demand changes with total drug expenditure. Along a fixed-price expenditure–utility path, the compensated elasticity changes little with $E_{D,l}$ when demand for the good is approximately unit-elastic with respect to total drug expenditure. It is therefore important to estimate price elasticities directly rather than infer them from GDP. The corresponding exploratory Slutsky-style analyses are reported in Table 3, in the rows currently labeled “Log GDP (PPP) Hicksian.” A more formal statement follows.

Theorem 1 (NH–CES income dependence). *Suppose that the implicit NH–CES representation defines regular, differentiable preferences and that the within-drug allocation is interior. Define the compensated and conditional Marshallian own-price elasticities as*

$$\epsilon_{kk,l}^H = \left. \frac{\partial \ln h_{k,l}}{\partial \ln \tilde{P}_{k,l}} \right|_{U_l}, \quad \epsilon_{kk,l}^M = \left. \frac{\partial \ln C_{k,l}^M}{\partial \ln \tilde{P}_{k,l}} \right|_{E_{D,l}},$$

and define the conditional Marshallian expenditure elasticity as $\eta_{E_{D,l}}^{C_{k,l}} = \left. \frac{\partial \ln C_{k,l}^M}{\partial \ln E_{D,l}} \right|_{\tilde{P}_l}$. At the common Hicksian–

Marshallian allocation, $\epsilon_{kk,l}^H = -\sigma(1 - s_{k,l})$, and the Slutsky equation implies $\epsilon_{kk,l}^M = \epsilon_{kk,l}^H - s_{k,l}\eta_{E_{D,l}}^{C_{k,l}} = -\sigma(1 - s_{k,l}) - s_{k,l}\eta_{E_{D,l}}^{C_{k,l}}$. In the homothetic CES special case, $\eta_{E_{D,l}}^{C_{k,l}} = 1$, so $\epsilon_{kk,l}^M = -\sigma + (\sigma - 1)s_{k,l}$.

To describe how the compensated elasticity changes with total drug expenditure, define $\bar{\epsilon}_{kk,l}^H(\tilde{P}_l, E_{D,l}) = \epsilon_{kk,l}^H(\tilde{P}_l, V_l(\tilde{P}_l, E_{D,l}))$. Then

$$\left. \frac{\partial \bar{\epsilon}_{kk,l}^H}{\partial \ln E_{D,l}} \right|_{\tilde{P}_l} = \sigma s_{k,l} (\eta_{E_{D,l}}^{C_{k,l}} - 1).$$

Proof. Holding U_l fixed, Equation (J.6) implies

$$\frac{\partial \ln e_l(\tilde{P}_l, U_l)}{\partial \ln \tilde{P}_{k,l}} = \frac{f_k(U_l) \tilde{P}_{k,l}^{1-\sigma}}{\sum_j f_j(U_l) \tilde{P}_{j,l}^{1-\sigma}} = s_{k,l}.$$

Differentiating Equation (J.5) with respect to $\ln \tilde{P}_{k,l}$, holding utility fixed, therefore gives

$$\epsilon_{kk,l}^H = \sigma s_{k,l} - \sigma = -\sigma(1 - s_{k,l}).$$

At the same prices and attained utility, Hicksian and conditional Marshallian quantities coincide. The own-price Slutsky equation is therefore $\varepsilon_{kk,l}^H = \varepsilon_{kk,l}^M + s_{k,l}\eta_{E_{D,l}}^{C_{k,l}}$, which yields $\varepsilon_{kk,l}^M = -\sigma(1 - s_{k,l}) - s_{k,l}\eta_{E_{D,l}}^{C_{k,l}}$. Under homothetic CES preferences, conditional demand is homogeneous of degree one in total drug expenditure, so $\eta_{E_{D,l}}^{C_{k,l}} = 1$, giving

$$\varepsilon_{kk,l}^M = -\sigma + (\sigma - 1)s_{k,l}.$$

Finally, evaluate compensated demand at $V_l(\tilde{P}_l, E_{D,l})$. Because compensated and conditional Marshallian quantities coincide at this utility,

$$s_{k,l} = \frac{\tilde{P}_{k,l} C_{k,l}^M(\tilde{P}_l, E_{D,l})}{E_{D,l}}.$$

Holding prices fixed,

$$\frac{\partial \ln s_{k,l}}{\partial \ln E_{D,l}} = \eta_{E_{D,l}}^{C_{k,l}} - 1,$$

and hence

$$\frac{\partial s_{k,l}}{\partial \ln E_{D,l}} = s_{k,l} (\eta_{E_{D,l}}^{C_{k,l}} - 1).$$

Differentiating $\bar{\varepsilon}_{kk,l}^H = -\sigma(1 - s_{k,l})$ along this fixed-price expenditure–utility path gives

$$\left. \frac{\partial \bar{\varepsilon}_{kk,l}^H}{\partial \ln E_{D,l}} \right|_{\tilde{P}_l} = \sigma s_{k,l} (\eta_{E_{D,l}}^{C_{k,l}} - 1),$$

which completes the proof. $\square$

We next introduce a separate aggregate isoelastic benchmark. It is not obtained by aggregating the preceding NH–CES system: its country-level elasticity η_c and tightness parameter μ_c are distinct theoretical objects.

J.3 Efficiency of Ramsey pricing across countries

We consider a Ramsey framework where each country $c \in \{H, L\}$ —standing for high and low income—has its own regulator and cost-recovery or budget constraint. Equivalently, a single planner may face country-specific budget pressure or welfare weights. Let μ_c denote the corresponding country-specific tightness parameter.

For each country, aggregate molecule demand at posted price P is

$$C_c(P) = A_c P^{-\eta_c}, \quad A_c > 0, \quad \eta_c > 1,$$

where η_c is the absolute Marshallian elasticity of this aggregate isoelastic demand schedule. Insurance wedges are absorbed into the scale parameter.^{J.3} Here $C_c(P)$ is aggregate molecule demand in country c , summed over consumer groups. We assume constant marginal cost MC (Filson, 2012; Bokhari and Mariuzzo, 2018).

^{J.3}If effective demand is $\widehat{C}_c(\tilde{P}) = \widehat{A}_c \tilde{P}^{-\eta_c}$, with $\tilde{P} = (1 - \delta_c)P$, where δ_c denotes the country-specific coverage rate, as in Barros and Martinez-Giralt (2008), then $C_c(P) = A_c P^{-\eta_c}$, where $A_c = \widehat{A}_c (1 - \delta_c)^{-\eta_c}$.

The Ramsey first-order condition is

$$\frac{P_{R,c} - MC}{P_{R,c}} = \frac{\mu_c}{\eta_c}, \quad 0 < \mu_c < 1,$$

where $P_{R,c}$ is the Ramsey price. The monopoly formula is obtained at the limiting value $\mu_c = 1$ (Joskow, 2007). A higher μ_c corresponds to a tighter revenue requirement and therefore to a larger optimal markup. The pricing rule implies

$$\frac{P_{R,c}}{MC} = \frac{\eta_c}{\eta_c - \mu_c}.$$

To interpret the area under the Marshallian demand curve as exact consumer surplus, this partial-equilibrium benchmark assumes quasi-linear preferences, or equivalently constant marginal utility of income. At any price P , define

$$CS_c(P) = \int_P^\infty C_c(x) dx, \quad PS_c(P) = (P - MC)C_c(P), \quad TS_c(P) = CS_c(P) + PS_c(P).$$

At the Ramsey price,

$$DWL_c = \int_{MC}^{P_{R,c}} C_c(P) dP - (P_{R,c} - MC)C_c(P_{R,c}) = TS_c(MC) - TS_c(P_{R,c}).$$

In particular,

$$CS_c(P) = \frac{A_c}{\eta_c - 1} P^{1-\eta_c}, \quad TS_c(MC) = \frac{A_c}{\eta_c - 1} MC^{1-\eta_c},$$

and

$$TS_c(P_{R,c}) = \frac{A_c MC^{1-\eta_c}}{\eta_c - 1} \frac{\eta_c + \mu_c(\eta_c - 1)}{\eta_c - \mu_c} \left(\frac{\eta_c - \mu_c}{\eta_c} \right)^{\eta_c}.$$

Consequently, defining $R_c = DWL_c / [CS_c(P_{R,c}) + PS_c(P_{R,c})]$, we obtain

$$R_c + 1 = F(\eta_c, \mu_c) := \frac{\eta_c - \mu_c}{\eta_c(1 + \mu_c) - \mu_c} \left(\frac{\eta_c}{\eta_c - \mu_c} \right)^{\eta_c}. \quad (\text{J.7})$$

Theorem 2. Suppose $\eta_L > \eta_H > 1$ and fix $\mu_H \in (0, 1)$. There exists a unique $\mu_L^* \in (0, 1)$ such that $R_L = R_H$. Moreover,

$$R_L > R_H \iff \mu_L > \mu_L^*, \quad R_L < R_H \iff \mu_L < \mu_L^*.$$

Proof. For fixed $\eta > 1$, Equation (J.7) implies

$$\frac{\partial \log F(\eta, \mu)}{\partial \mu} = \frac{\mu \eta (\eta - 1)}{(\eta - \mu) \{ \eta + \mu(\eta - 1) \}} > 0 \quad \text{for } \mu \in (0, 1).$$

Thus, $F(\eta, \mu)$ is continuous and strictly increasing in μ , with $F(\eta, 0) = 1$. In addition,

$$\frac{\partial \log F(\eta, 1)}{\partial \eta} = \log \left(\frac{\eta}{\eta - 1} \right) - \frac{2}{2\eta - 1} > 0.$$

Hence, because $\eta_L > \eta_H$ and $\mu_H \in (0, 1)$,

$$1 < F(\eta_H, \mu_H) < F(\eta_H, 1) < F(\eta_L, 1).$$

The intermediate value theorem therefore gives a unique $\mu_L^* \in (0, 1)$ satisfying $F(\eta_L, \mu_L^*) = F(\eta_H, \mu_H)$. Strict monotonicity in μ_L establishes the two rankings. $\square$

The theorem shows that the ranking of deadweight-loss shares depends on the pair (η_c, μ_c) , not on demand elasticity alone. A symmetric threshold for μ_H exists only when the target value $F(\eta_L, \mu_L)$ lies below the endpoint $F(\eta_H, 1)$ and it is therefore not guaranteed for every (η_L, μ_L) . Furthermore, notice that the aggregate theoretical elasticity η_c is not the same object as the paper’s country–molecule–year DML elasticity. Moreover the empirical analysis estimates the signed cell-level response $\hat{\epsilon}_{gca}^{\text{own,ADL}}$ and may report its absolute magnitude $\hat{\chi}_{gca}^{\text{ADL}} = -\hat{\epsilon}_{gca}^{\text{own,ADL}}$. Without an explicit aggregation model, $\hat{\chi}_{gca}^{\text{ADL}}$ is not set equal to the aggregate isoelastic parameter η_c . The theorem therefore motivates conditioning directly on estimated demand responsiveness rather than using income alone; it does not identify η_c , μ_c , or country-level welfare from the DML estimates ^{J.4}.

A practical difficulty in testing Theorem 2 directly is that neither the deadweight-loss share $R_c \equiv \frac{DWL_c}{CS_c(P_{R,c}) + PS_c(P_{R,c})}$ nor the threshold μ_L^* is directly observed in the data. Indeed, DWL_c is a counterfactual welfare object (it requires the full demand schedule and the efficient benchmark at $P = MC$), while μ_L^* is defined implicitly by the equality $F(\eta_L, \mu_L^*) = F(\eta_H, \mu_H)$ and therefore depends on the unobserved tightness parameters. We can however notice that, if two countries c and c' , $c \neq c'$, have similar aggregate elasticities and market size, the Ramsey first-order condition implies $L_c \equiv \frac{P_{R,c} - MC}{P_{R,c}} = \frac{\mu_c}{\eta_c} = \mu_c \left(\frac{1}{\eta_c} \right)$, where L_c is the Lerner index. If marginal costs were observed and μ_c were constant within the relevant comparison, a regression of L_c on $1/\eta_c$ would recover that common tightness parameter. Because neither condition is established in the global panel, we do not implement this as a direct welfare test. The theorem instead implies that, holding η_c fixed, a larger μ_c produces a larger markup and a larger deadweight-loss share.

Economic interpretation of Theorem 1 and Table 3. Theorem 1 explains why income alone need not rank countries by price responsiveness. Even for a common substitution parameter σ , the conditional Marshallian elasticity, $\epsilon_{kk,l}^M = -\sigma(1 - s_{k,l}) - s_{k,l}\eta_{E_{D,l}}^{C_{k,l}}$, also depends on the drug’s expenditure share $s_{k,l}$ and on how its demand responds to total drug expenditure $\eta_{E_{D,l}}^{C_{k,l}}$. GDP does not determine either object: countries with similar income can differ in insurance coverage, reimbursement, regulation, and product composition.

Table 3 asks whether GDP nevertheless provides a useful empirical summary of the estimated elasticities. The small and statistically weak GDP coefficients indicate that it does not. The exploratory Slutsky-style adjustment produces similarly small GDP associations, suggesting that this conclusion is not driven solely by the expenditure effect contained in the conditional Marshallian response. By contrast, the positive HIC-versus-other-market coefficient indicates that demand is less price responsive in high-income markets after conditioning on molecule and year. This association may reflect broader institutional differences that GDP alone does not capture.

The theorem and the table therefore deliver a common, economically relevant message: income is not a sufficient statistic for pharmaceutical demand responsiveness. Evaluating international pricing patterns consequently requires directly estimated elasticities rather than elasticities inferred from countries’ income levels.

Commentary on the empirical evidence for Theorem 2. Theorem 2 clarifies why cross-country welfare statements based only on elasticities, or on income proxies for elasticities, are incomplete. In the Ramsey environment of Appendix J, each country c has an aggregate isoelastic-demand magnitude $\eta_c > 1$ and a

^{J.4}The theorem therefore motivates conditioning directly on estimated demand responsiveness rather than using income alone; it does not identify η_c , μ_c , or country-level welfare from the DML estimates. In simple terms, the DML elasticities describe differences in demand responsiveness, but they cannot by themselves rank countries’ welfare losses because those losses also depend on the unobserved tightness parameter μ_c and on how cell-level demand is aggregated to the country level.

country-specific tightness parameter μ_c , which summarizes the shadow cost of the budget, coverage, or revenue constraint. The Ramsey condition is $L_c \equiv \frac{P_{R,c} - MC}{P_{R,c}} = \frac{\mu_c}{\eta_c}$. Theorem 2 shows that the deadweight-loss share, $R_c = \frac{DWL_c}{CS_c(P_{R,c}) + PS_c(P_{R,c})}$, is ranked by the pair (η_c, μ_c) , not by η_c alone. In particular, for given (η_H, μ_H) , there exists a threshold μ_L^* such that the low-income country has a higher deadweight-loss share if and only if $\mu_L > \mu_L^*$. Thus, even if low-income countries have more elastic demand, their Ramsey deadweight-loss share can be higher or lower depending on the tightness of their budget or coverage constraint.

This result cannot be tested directly in Table 3. The deadweight-loss share R_c is a counterfactual welfare object, and the threshold μ_L^* is not observed. Moreover, constructing the Lerner index L_c for the full global panel would require marginal costs, which are not observed. The empirical relevance of the theorem is therefore mainly diagnostic: it explains why GDP per capita is not a sufficient statistic for Ramsey efficiency. A GDP proxy may combine two distinct objects, the elasticity η_c and the tightness parameter μ_c , without separating them.

The GDP rows in Table 3 are consistent with this warning. Log GDP (PPP) is not systematically related to estimated elasticities, while the listed-HIC–other-market indicator is sometimes positive, indicating less price-sensitive demand in high-income countries. These patterns do not identify μ_c or the welfare ranking R_c . They instead support the theorem’s main empirical lesson: income categories alone cannot determine Ramsey efficiency, because the elasticity channel and the tightness channel are distinct.

K Interpreting Price–Elasticity Associations with Heterogeneous Marginal Costs

The inverse-elasticity benchmark in Appendix J relates demand responsiveness to price–cost margins assuming a fixed marginal cost. This appendix examines when the price–elasticity associations in Section 5 provide evidence about those margins despite differences in marginal costs. In this scenario, the concern is that while the main regressions in Table 3 ask whether higher average log prices are associated with less price-responsive demand, higher prices where demand is less elastic could reflect higher costs of supplying those medicines rather than larger markups. We therefore investigate whether observable characteristics explain cost differences associated with elasticities, and whether removing predictable price variation may help reduce their contribution to the observed association.

We first identify a sufficient condition for interpreting the price coefficient’s sign as evidence about markups in subsection K.1. We then explain – in the subsequent subsections – what price residualization can remove and present auxiliary structural diagnostics of the cost–elasticity relationship. Namely, to do the latter, we need to obtain markups, which are unknown because marginal costs are unobserved. We therefore combine our RCL demand estimates with static Bertrand pricing, assuming that each firm sets prices jointly for the products it owns.^{K.1} Subtracting the resulting markups, measured in price units, from observed prices gives model-implied marginal costs.^{K.2}

Using these model-implied costs, we first assess how well a separate set of diagnostic controls predicts log marginal costs. We then compare the regression of log marginal costs on the estimated ADL elasticity measure before and after residualizing both variables using these controls. A weaker association after this

^{K.1}For pharmaceutical applications of differentiated-products Bertrand pricing and marginal-cost recovery, see Dubois and Lasio (2018), Björnerstedt and Verboven (2016), and Gugler et al. (2026). Related uses include the unregulated patent counterfactuals of Chaudhuri et al. (2006), who bound observed costs because of price regulation, and the illustrative Bertrand oligopoly model in Costinot et al. (2019).

^{K.2}We retain prices in the main regressions because replacing them with recovered markups would make the international comparison depend on the conduct assumptions required to recover those markups.

adjustment would suggest that observable characteristics help explain why costs and elasticities vary together within the auxiliary model. Such evidence would support the rationale for using residualized prices as regressors in Table 3, without establishing that these price residuals isolate markups.

K.1 Prices, costs, and the inverse-elasticity benchmark

The Ramsey benchmark motivates the comparison in Table 3. It considers a regulator who chooses medicine prices to meet a financing requirement while limiting the welfare loss from reduced consumption. Prices above marginal cost contribute toward fixed costs, such as R&D expenditure, but also discourage medicine use. A given percentage price increase produces a smaller percentage reduction in consumption when demand is less elastic. For a given financing pressure, the benchmark therefore assigns larger proportional margins to markets with less elastic demand.

Appendix J has formalized this trade-off using a simple aggregate-demand model. Briefly, let $\eta_c > 1$ denote the magnitude of the aggregate molecule-demand elasticity in country c , and let $0 < \mu_c < 1$ summarize the tightness of its financing constraint. The resulting Ramsey price $P_{R,c}$ satisfies

$$\frac{P_{R,c} - MC}{P_{R,c}} = \frac{\mu_c}{\eta_c}, \quad \log\left(\frac{P_{R,c}}{MC}\right) = -\log\left(1 - \frac{\mu_c}{\eta_c}\right). \quad (\text{K.1})$$

Holding financing pressure $\mu_c = \mu$ fixed, less elastic demand implies a larger proportional markup. If marginal costs are also common across markets, as in Equation K.1, markets with less elastic demand also have higher prices. Under these conditions, the price–elasticity relationship is informative about markups, because differences in prices reflect differences in proportional margins rather than marginal costs. Differences in costs or financing pressure, however, can change this ordering. The analysis below focuses on the cost concern. We examine how cost differences contribute to the observed price–elasticity association and identify a weaker sufficient condition under which the price coefficient’s sign remains informative about the association between elasticities and price–cost ratios. This condition allows marginal costs to differ across markets and provides a sufficient restriction for interpreting the observed-price coefficients in Table 3 as evidence about markups. In later subsections, we will examine whether it is possible to relax this assumption through partialling out.

More specifically, we ask whether the association between higher prices and less elastic demand reflects larger markups. Medicines with less elastic demand might instead cost more to produce, making their prices higher even with the same proportional markup. The concern is therefore whether costs vary systematically with elasticities within the comparisons made by our regressions.

To make the cost concern precise, we begin with Equations (7) and (8), which relate estimated elasticities to average log prices. The following identity shows how costs and markups enter the price variable used in these regressions. For positive prices and marginal costs,

$$p_{igct} = \log MC_{igct} + \log\left(\frac{P_{igct}}{MC_{igct}}\right), \quad p_{igct} = \log P_{igct}. \quad (\text{K.2})$$

Thus, log prices combine log marginal costs and the log price–cost ratio. The next step is to show how each component contributes to the price–elasticity association.

We begin with the main regressions that use observed prices. Because these regressions include fixed effects, their coefficients describe the price–elasticity association that remains after accounting for those effects. To decompose this same association, we average log prices, log marginal costs, and the log price–cost ratio at the country–year or molecule–year level used in the relevant regression, and remove the same fixed

effects from each component. An overbar denotes these averages, calculated using the same observations and averaging weights for all three variables. The operator R_{FE} denotes the subsequent removal of fixed effects, using the regression weights. This allows us to separate the cost and markup contributions to the association actually measured by the regression. Suppressing indices and applying the covariance operator on both sides of Equation K.2, we obtain

$$\begin{aligned} \text{Cov}(R_{FE}\hat{\epsilon}^{\text{own}}, R_{FE}\bar{p}) &= \text{Cov}(R_{FE}\hat{\epsilon}^{\text{own}}, R_{FE}\overline{\log MC}) \\ &\quad + \text{Cov}(R_{FE}\hat{\epsilon}^{\text{own}}, R_{FE}\overline{\log(P/MC)}), \end{aligned}$$

where all covariances use the same weights. The first term on the right captures the association with costs, while the second captures the association with the log price–cost ratio.

By the Frisch–Waugh–Lovell theorem, the price coefficient equals the covariance on the left-hand side divided by $\text{Var}(R_{FE}\bar{p})$. Namely, dividing by the variance of average log prices after removing fixed effects we can write

$$\begin{aligned} \hat{\beta} &= \frac{\text{Cov}(R_{FE}\hat{\epsilon}^{\text{own}}, R_{FE}\bar{p})}{\text{Var}(R_{FE}\bar{p})} \\ &= \frac{\text{Cov}(R_{FE}\hat{\epsilon}^{\text{own}}, R_{FE}\overline{\log MC})}{\text{Var}(R_{FE}\bar{p})} \\ &\quad + \frac{\text{Cov}(R_{FE}\hat{\epsilon}^{\text{own}}, R_{FE}\overline{\log(P/MC)})}{\text{Var}(R_{FE}\bar{p})}. \end{aligned}$$

The first term captures the contribution associated with marginal costs; the second captures the contribution associated with the log price–cost ratio.

Since $\text{Var}(R_{FE}\bar{p})$ is positive, the coefficient’s sign depends on the sum of the cost and price–cost-ratio associations. Our aim is to assess whether higher prices where demand is less elastic reflect larger price–cost ratios rather than higher costs. To check the latter, let’s first assess when this happens. A sufficient condition for the price coefficient to have the same sign as the association between estimated elasticities and the log price–cost ratio is $\text{Cov}(R_{FE}\hat{\epsilon}^{\text{own}}, R_{FE}\overline{\log MC}) = 0$, i.e., the contribution of the (averaged without fixed effects) marginal-cost term to the numerator of the price coefficient is zero. So in this case, we have that higher prices where demand is less elastic reflect larger price–cost and the price patterns can be interpreted as evidence about markups^{K.3}. Marginal costs can still differ across observations since the condition only requires that the remaining variation in average log marginal costs be uncorrelated with estimated elasticities. The numerator, $\text{Cov}(R_{FE}\hat{\epsilon}^{\text{own}}, R_{FE}\overline{\log(P/MC)})$, then contains only the covariance with the log price–cost ratio, which determines the price coefficient’s sign (division by $\text{Var}(R_{FE}\bar{p}) > 0$ does not change the sign).

To summarize, we have shown that a positive price coefficient reflects higher price–cost ratios where demand is less elastic if average log marginal costs and estimated elasticities are uncorrelated after removing fixed effects. The following subsections examine whether observable characteristics explain why marginal costs might be higher where demand is less elastic, helping to mitigate the reliance on the latter independence assumption via residualization. If so, removing predictable price variation could reduce the extent to which these cost differences, rather than larger markups, explain the observed price–elasticity association.

^{K.3}E.g. a positive price coefficient indicates that less elastic demand is associated with larger markups, measured by average log price–cost ratios, after accounting for the regression’s fixed effects.

We first explain what price residualization can remove. Because marginal costs are unobserved, Subsection K.2 then uses model-implied costs from the auxiliary RCL–Bertrand model to investigate whether observables predict log marginal costs and whether the cost–elasticity association weakens after removing the predicted components of both variables. A weaker association would suggest that observables account for part of the potential cost explanation, supporting the rationale for price residualization. This diagnostic provides indirect evidence, rather than establishing that residualized prices isolate markups.

K.2 Price residualization and structural cost diagnostics

We now examine whether price residualization can reduce the contribution of costs to the price–elasticity association. The idea is that removing price variation predicted by observable characteristics may remove systematic cost differences associated with elasticities. However, the same adjustment may also partly remove predictable markup variation.

Recall that the estimated residualized log price is $\hat{p}_{igct}^\circ = p_{igct} - \hat{m}^{p,(-k)}(\mathcal{W}_{igct})$, where $\hat{m}^{p,(-k)}$ predicts log price using the demand-stage controls. To understand what this adjustment removes, first consider its population counterpart. Applying the linearity of conditional expectations to Equation (K.2) and suppressing observation indices gives

$$\underbrace{p - \mathbb{E}[p \mid \mathcal{W}]}_{p^\circ} = \underbrace{\log MC - \mathbb{E}[\log MC \mid \mathcal{W}]}_{(\log MC)^\circ} + \underbrace{\log(P/MC) - \mathbb{E}[\log(P/MC) \mid \mathcal{W}]}_{[\log(P/MC)]^\circ}. \quad (\text{K.3})$$

Residual prices contain both remaining log-cost variation and remaining log price–cost-ratio variation. To see how each contributes to the regression association, we apply the same averaging and fixed-effect adjustment to both components and take their covariance with estimated elasticities. Linearity gives

$$\begin{aligned} \text{Cov}(R_{\text{FE}}\hat{\epsilon}^{\text{own}}, R_{\text{FE}}\bar{p}^\circ) &= \text{Cov}\left(R_{\text{FE}}\hat{\epsilon}^{\text{own}}, R_{\text{FE}}\overline{(\log MC)^\circ}\right) \\ &\quad + \text{Cov}\left(R_{\text{FE}}\hat{\epsilon}^{\text{own}}, R_{\text{FE}}\overline{[\log(P/MC)]^\circ}\right), \end{aligned} \quad (\text{K.4})$$

where all covariances use the regression weights. It is easy to see that a sufficient condition for costs to contribute nothing to this association is

$$\text{Cov}\left(R_{\text{FE}}\hat{\epsilon}^{\text{own}}, R_{\text{FE}}\overline{(\log MC)^\circ}\right) = 0,$$

since, under this condition, the first term disappears, leaving only the association with the remaining log price–cost ratio. Similarly to the previous subsection, dividing (both sides of) Equation (K.3) by the positive variance of $R_{\text{FE}}\bar{p}^\circ$ gives the residual-price coefficient, so its sign reflects that remaining log-ratio association. Provided residual prices retain positive variance after removing fixed effects, the price coefficient’s sign then reflects the association between estimated elasticities and – importantly – the *remaining* log price–cost-ratio component after residualization.

We investigate whether $\text{Cov}\left(R_{\text{FE}}\hat{\epsilon}^{\text{own}}, R_{\text{FE}}\overline{(\log MC)^\circ}\right) \neq 0$ using two complementary diagnostics. First, we measure how much log-cost (estimated assuming Bertrand) variation the observables predict. Second, we examine whether the cost–elasticity association weakens after removing the components of both variables predicted by the diagnostic controls. In other words, the former diagnostic describes how much cost variation

the adjustment captures and the second whether the variation left behind remains associated with demand responsiveness.

We estimate the marginal cost, as usual, by combining estimated demand with assumptions about how firms set prices and which products they price jointly. In this context, demand estimates describe how sales respond to prices, while the pricing and ownership assumptions determine how firms take these responses into account when choosing prices. Put together, these ingredients imply price–cost margins, which we subtract from observed prices to obtain model-implied marginal costs.

For the diagnostics reported here, we use the auxiliary U.S. RCL demand estimates from Section 4.1 and assume static differentiated-products Bertrand pricing where each firm chooses its prices to maximize profits, taking other firms’ prices as given. We compare two assumptions about which products are priced jointly.^{K.4}

Our preferred specification follows observed corporate ownership. Each corporation chooses prices jointly for its products, accounting for how a price change affects sales of its other products. Let $H_{m,jj'} = 1$ when products j and j' belong to the same corporation in market m , and zero otherwise. The pricing conditions are

$$s_{jm} + \sum_{j' \in \mathcal{J}_m} H_{m,jj'} \left(P_{j'm} - \widehat{MC}_{j'm}^B \right) \frac{\partial s_{j'm}}{\partial P_{jm}} = 0.$$

Given estimated demand responses and observed prices and shares, these conditions allow us to recover the Bertrand-implied marginal costs $\widehat{MC}^B$. As a robustness check, we assume that each product is instead priced separately to maximize its own profits. In this single-product robustness check, setting $H_m = I_{J_m}$ simplifies the pricing conditions to $P_{jm} - \widehat{MC}_{jm}^B = \frac{s_{jm}}{-\partial s_{jm}/\partial P_{jm}}$. In what follows, we conduct the diagnostics under both observed multiproduct ownership – our preferred specification – and single-product pricing – our robustness check. This comparison assesses how recovered costs and the diagnostic results depend on firms taking into account how one product’s price affects sales of their other products. We use the same candidate diagnostic controls in both specifications, although the usable samples differ because log-cost analysis requires positive recovered costs. For simplicity, the notation below suppresses the distinction between ownership specifications. As mentioned, within each specification, we separately predict log model-implied marginal costs and the already estimated ADL elasticity measure using these controls. We remove the predicted component from each variable to examine whether costs and elasticities remain associated beyond the patterns explained by the observables. Comparing this remaining association with the unadjusted relationship indicates whether the controls account for part of their co-movement, thus providing auxiliary evidence about predictable cost patterns.

Let r index matched product–quarter observations in the auxiliary U.S. sample. We assign each observation the own-price elasticity estimated for its ATC-4 class and branded/generic status in the low-dimensional DML benchmark ^{K.5}. We use its sign-reversed value, $\widehat{\chi}_r^{\text{ADL}} = -\widehat{\epsilon}_r^{\text{own,ADL}}$, in the diagnostics below.

Let W_r collect the explanatory variables used to predict model-implied log marginal costs and the assigned elasticity measure. These include product and market characteristics, time indicators, lagged prices and sales, and information on competitors and foreign prices.

^{K.4}The general approach can also be applied under other conduct assumptions, using their corresponding optimality conditions. The two specifications reported here vary the products whose profits are considered jointly within the Bertrand framework.

^{K.5}The low-dimensional specification in Section 4.1 was chosen to match the ATC-4-by-branded/generic reporting level of the external benchmark. We retain these estimates in the cost diagnostic to keep the two exercises consistent. Although estimation uses product–quarter observations, this specification allows elasticities to differ only across these groups. Matching therefore assigns each observation its group’s estimated response, while model-implied costs may still vary across products and quarters within the group. The diagnostic does not assess elasticity heterogeneity within these groups.

We estimate separate cost and elasticity prediction functions using elastic net with held-out time blocks and neighboring-block exclusion in the sense that for an observation in time block $k(r)$, both models are estimated excluding that block and its neighboring blocks. This prevents the predictions from being fitted to the observations being evaluated and limits the use of closely related observations from adjacent periods. Namely, for each observation, define

$$\begin{aligned}\hat{e}_r^{MC} &= \log \widehat{MC}_r^B - \hat{f}_{MC}^{(-k(r))}(W_r), \\ \hat{e}_r^\chi &= \hat{\chi}_r^{\text{ADL}} - \hat{f}_\chi^{(-k(r))}(W_r).\end{aligned}$$

First, we assess how much of the variation in model-implied log marginal costs the auxiliary controls can predict. We compare the model's held-out prediction errors with the errors obtained by assigning the same sample-average log cost to every observation. Pooling predictions over the valid sample $\mathcal{S}$, we compute

$$R_{\text{oos}}^2 = 1 - \frac{\sum_{r \in \mathcal{S}} (\hat{e}_r^{MC})^2}{\sum_{r \in \mathcal{S}} \left(\log \widehat{MC}_r^B - \log \widehat{MC}_{\mathcal{S}}^B \right)^2}.$$

Intuitively, a higher R_{oos}^2 indicates that the controls provide more information about cost differences.

Table K.1 reports an out-of-sample R^2 for predicting model-implied log marginal costs of 0.815 under observed multiproduct ownership and 0.824 under single-product ownership. Thus, the predictions seem to reduce squared prediction error by about 82 percent relative to assigning the sample-average log cost to every observation. The auxiliary controls therefore predict substantial cost variation, but the remaining prediction errors may still be associated with estimated elasticities. The next diagnostic examines that association specifically asking whether cost differences left unexplained by the controls are associated with differences in the estimated elasticity measure that those controls also fail to predict. In particular, we average each prediction error within ATC-4-by-branded/generic groups where a positive mean cost error indicates that a group's model-implied log costs are higher than predicted, and a positive mean elasticity error indicates that its sign-reversed ADL elasticity measure is higher than predicted.

We then regress the mean cost errors on the mean elasticity errors, including an intercept and giving each group equal weight. The slope, denoted $\hat{\rho}^{\text{res}}$, measures whether groups with higher-than-predicted elasticity measures also tend to have higher- or lower-than-predicted costs. A zero slope corresponds to no linear association between these two remaining components at the group level, i.e. groups with higher-than-predicted model-implied log costs do not tend to have higher-than-predicted values of the (sign-reversed) elasticity.

Under multiproduct ownership, the estimated slope is 0.128 (standard error 0.112), and the correlation between the group-average prediction errors is 0.108. Under single-product ownership, the corresponding slope is 0.101 (standard error 0.110), with correlation 0.083. The observation-weighted slopes are 0.071 and 0.016, respectively. The slope estimates are not statistically distinguishable from zero. However, their imprecision and reliance on model-implied costs mean that these results do not establish that true marginal costs are unrelated to elasticities in the main regressions.

Third, we repeat the same group-level regressions before subtracting the predicted components of log costs and sign-reversed elasticities, obtaining $\hat{\rho}^{\text{raw}}$. The raw slopes are approximately -0.688 under multiproduct ownership and -0.687 under single-product ownership: higher sign-reversed elasticity estimates are associated with lower model-implied log costs. After adjustment, the slopes become positive, with magnitudes equal to 18.6 and 14.8 percent of the raw slopes, respectively. This change suggests that observable differences across groups help explain the original negative association, although the adjusted estimates remain imprecise.

Table K.1. Structural cost-side diagnostics

Demand	Conduct	Ownership	N	OOS R^2	$\hat{\rho}^{\text{res}}$	s.e.	Corr. res.	$ \hat{\rho}^{\text{res}} / \hat{\rho}^{\text{raw}} $
Selected RCL	Bertrand	Single-product	80021	0.824	0.101	0.110	0.083	0.148
Selected RCL	Bertrand	Observed multiproduct	76727	0.815	0.128	0.112	0.108	0.186

Notes: N counts matched observations with positive model-implied costs and finite ADL elasticities. OOS R^2 uses held-out predictions of log costs from the auxiliary dictionary W . Slopes and correlations use 243 equally weighted ATC-4-by-branded/generic cell averages; standard errors are heteroskedasticity-robust. The residual regressions use $\hat{e}^{MC}$ and $\hat{e}^{\mathcal{A}}$; the raw regressions use $\log \widehat{MC}^B$ and $\hat{\chi}^{\text{ADL}}$. Observation-weighted alternatives use product–quarter observations and standard errors clustered by reporting cell. Reported standard errors treat the estimated inputs as given and do not propagate uncertainty from preceding estimation steps. All results depend on the estimated demand model and assumed pricing conduct.

Table K.1, therefore, provides reassurances. Indeed since observable characteristics explain much of the co-movement between costs and elasticities^{K.6} – which would have potentially made a price–elasticity relationship look like a markup–elasticity relationship – residualizing for these characteristics may reduce this cost concern.

Finally, because the main price–elasticity analysis covers 63 countries, extending the cost diagnostics beyond the U.S. would be informative. However, doing so would require credible marginal-cost estimates for those countries. Recovering these costs from our data would entail additional structural assumptions about demand, pricing conduct, and ownership across different institutional settings. We therefore limit the structural cost diagnostics to the auxiliary U.S. sample.

^{K.6}Table K.1 shown that even though such co-movements are present, i.e. $\hat{\rho}^{\text{res}} \neq 0$, they largely attenuate when we control for observables $\hat{\rho}^{\text{res}}$ is almost 0

References

- Anderson, Theodore W and Herman Rubin (1949). “Estimation of the parameters of a single equation in a complete system of stochastic equations”. In: *The Annals of mathematical statistics* 20.1, pp. 46–63.
- Andrews, Isaiah, James H Stock, and Liyang Sun (2019). “Weak instruments in instrumental variables regression: Theory and practice”. In: *Annual Review of Economics* 11.1, pp. 727–753.
- Armstrong, Mark and John Vickers (2018). “Multiproduct pricing made simple”. In: *Journal of Political Economy* 126.4, pp. 1444–1471.
- Backus, Matthew, Christopher Conlon, and Michael Sinkinson (2021). *Common ownership and competition in the ready-to-eat cereal industry*. Tech. rep. National Bureau of Economic Research.
- Barros, P. P. and X. Martinez-Giralt (2008). “On international cost-sharing of pharmaceutical R&D”. In: *International Journal of Health Care Finance and Economics* 8.4, pp. 301–312.
- Baumol, W. J. and D. F. Bradford (1970). “Optimal Departures From Marginal Cost Pricing”. In: *The American Economic Review* 60.3, pp. 265–283. ISSN: 00028282. URL: <http://www.jstor.org/stable/1817977>.
- Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen (2014). “Inference on treatment effects after selection among high-dimensional controls”. In: *Review of Economic Studies* 81.2, pp. 608–650.
- Berndt, Ernst R., Patricia M. Danzon, and Gregory B. Kruse (2007). “Dynamic competition in pharmaceuticals: cross-national evidence from new drug diffusion”. In: *Managerial and Decision Economics* 28.4-5, pp. 231–250. DOI: <https://doi.org/10.1002/mde.1338>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/mde.1338>.
- Berry, Steven, James Levinsohn, and Ariel Pakes (1995). “Automobile Prices in Market Equilibrium”. In: *Econometrica* 63.4, pp. 841–890.
- Berry, Steven T and Philip A Haile (2014). “Identification in differentiated products markets using market level data”. In: *Econometrica* 82.5, pp. 1749–1797.
- Björnerstedt, Jonas and Frank Verboven (2016). “Does merger simulation work? Evidence from the Swedish analgesics market”. In: *American Economic Journal: Applied Economics* 8.3, pp. 125–164.
- Blundell, Richard, Xiaohong Chen, and Dennis Kristensen (2007). “Semi-nonparametric IV estimation of shape-invariant Engel curves”. In: *Econometrica* 75.6, pp. 1613–1669.
- Bokhari, Farasat AS and Franco Mariuzzo (2018). “Demand estimation and merger simulations for drugs: Logits v. AIDS”. In: *International Journal of Industrial Organization* 61, pp. 653–685.
- Bonhomme, Stéphane, Koen Jochmans, and Martin Weidner (2026). “A Neyman-Orthogonalization Approach to The Incidental Parameter Problem in Likelihood Models”. In: *Econometrica*.
- Branstetter, Lee, Chirantan Chatterjee, and Matthew J Higgins (2016). “Regulation and welfare: evidence from paragraph IV generic entry in the pharmaceutical industry”. In: *The RAND Journal of Economics* 47.4, pp. 857–890.
- Brekke, Kurt R, Ingrid Königbauer, and Odd Rune Straume (2007). “Reference pricing of pharmaceuticals”. In: *Journal of Health Economics* 26.3, pp. 613–642.
- Brot-Goldberg, Z. C., A. Chandra, B. R. Handel, and J. T. Kolstad (2017). “What does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics”. In: *The Quarterly Journal of Economics* 132.3, pp. 1261–1318. ISSN: 0033-5533. URL: <https://doi.org/10.1093/qje/qjx013>.
- Cai, Tony, Weidong Liu, and Xi Luo (2011). “A constrained ℓ_1 minimization approach to sparse precision matrix estimation”. In: *Journal of the American Statistical Association* 106.494, pp. 594–607.
- Caner, Mehmet (2009). “Lasso-type GMM estimator”. In: *Econometric Theory* 25.1, pp. 270–290.

- Caner, Mehmet and Hao Helen Zhang (2014). “Adaptive elastic net for generalized methods of moments”. In: *Journal of Business & Economic Statistics* 32.1, pp. 30–47.
- Carrera, Mariana and Sofia Villas-Boas (2023). “Generic aversion and observational learning in the over-the-counter drug market”. In: *American Economic Journal: Applied Economics* 15.3, pp. 380–410.
- Castanheira, Micael, Carmine Ornaghi, and Georges Siotis (2019). “The unexpected consequences of generic entry”. In: *Journal of Health Economics* 68, p. 102243.
- Chandra, A., J. Gruber, and R. McKnight (2014). “The impact of patient cost-sharing on low-income populations: Evidence from Massachusetts”. In: *Journal of Health Economics* 33, pp. 57–66. ISSN: 0167-6296. URL: <https://www.sciencedirect.com/science/article/pii/S0167629613001409>.
- Chaudhuri, Shubham, Pinelopi K Goldberg, and Panle Jia (2006). “Estimating the effects of global patent protection in pharmaceuticals: a case study of quinolones in India”. In: *American Economic Review* 96.5, pp. 1477–1514.
- Chen, Xiaohong, Timothy Christensen, and Sid Kankanala (2025). “Adaptive estimation and uniform confidence bands for nonparametric structural functions and elasticities”. In: *Review of Economic Studies* 92.1, pp. 162–196.
- Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins (2018). *Double/debiased machine learning for treatment and structural parameters*.
- Chernozhukov, Victor, Whitney K Newey, and Rahul Singh (2022). “Automatic debiased machine learning of causal and structural effects”. In: *Econometrica* 90.3, pp. 967–1027.
- Ching, Andrew T (2010). “Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration”. In: *International Journal of Industrial Organization* 28.6, pp. 619–638.
- Chintagunta, Pradeep K. and Ramarao Desiraju (2005). “Strategic Pricing and Detailing Behavior in International Markets”. In: *Marketing Science* 24.1, pp. 67–80. URL: <http://www.jstor.org/stable/40056939>.
- Comin, Diego, Danial Lashkari, and Marti Mestieri (2021). “Structural change with long-run income and price effects”. In: *Econometrica* 89.1, pp. 311–374.
- Coscelli, Andrea (2000). “The importance of doctors’ and patients’ preferences in the prescription decision”. In: *The Journal of Industrial Economics* 48.3, pp. 349–369.
- Costinot, Arnaud, Dave Donaldson, Margaret Kyle, and Heidi Williams (2019). “The more we die, the more we sell? a simple test of the home-market effect”. In: *The quarterly journal of economics* 134.2, pp. 843–894.
- Crawford, Gregory S and Matthew Shum (2005). “Uncertainty and learning in pharmaceutical demand”. In: *econometrica* 73.4, pp. 1137–1173.
- Dafny, Leemore, Christopher Ody, and Matt Schmitt (2017). “When discounts raise costs: the effect of copay coupons on generic utilization”. In: *American Economic Journal: Economic Policy* 9.2, pp. 91–123.
- Danzon, M. P. (2019). “Drug Pricing and Value in Oncology”. In: *Regulatory and Economic Aspects in Oncology* 213.
- Danzon, P. M. (1997). “Price Discrimination for Pharmaceuticals: Welfare Effects in the US and the EU”. In: *International Journal of the Economics of Business* 4.3.
- (2018). “Differential Pricing of Pharmaceuticals: Theory, Evidence and Emerging Issues”. In: *Pharmacoeconomics* 36.12, pp. 1395–1405.
- Danzon, Patricia M and Li-Wei Chao (2000a). “Cross-national price differences for pharmaceuticals: how large, and why?” In: *Journal of health economics* 19.2, pp. 159–195.
- (2000b). “Does regulation drive out competition in pharmaceutical markets?” In: *The Journal of Law and Economics* 43.2, pp. 311–358.

- Danzon, Patricia M and Michael F Furukawa (2003). “Prices And Availability Of Pharmaceuticals: Evidence From Nine Countries: The US market structure, with its higher prices for on-patent products and strong generic competition, appears more favorable to innovation than markets elsewhere.” In: *Health affairs* 22.Suppl1, W3–521.
- Danzon, Patricia M and Adrian Towse (2003). “Differential pricing for pharmaceuticals: reconciling access, R&D and patents”. In: *International journal of health care finance and economics* 3, pp. 183–205.
- Dubois, Pierre, Ashvin Gandhi, and Shoshana Vasserman (2022). *Bargaining and international reference pricing in the pharmaceutical industry*. Tech. rep. National Bureau of Economic Research.
- Dubois, Pierre and Laura Lasio (2018). “Identifying industry margins with price constraints: Structural estimation on pharmaceuticals”. In: *American Economic Review* 108.12, pp. 3685–3724.
- Dubois, Pierre, Yassine Lefouili, and Stéphane Straub (2021). “Pooled procurement of drugs in low and middle income countries”. In: *European Economic Review* 132, p. 103655.
- Duggan, Mark and Fiona Scott Morton (2010). “The effect of Medicare Part D on pharmaceutical prices and utilization”. In: *American Economic Review* 100.1, pp. 590–607.
- Einav, L., A. Finkelstein, and M. Polyakova (2018). “Private Provision of Social Insurance: Drug-Specific Price Elasticities and Cost Sharing in Medicare Part D”. In: *American Economic Journal: Economic Policy* 10.3, pp. 122–53. URL: <https://www.aeaweb.org/articles?id=10.1257/pol.20160355>.
- Einav, Liran, Amy Finkelstein, and Paul Schrimpf (2015). “The response of drug expenditure to nonlinear contract design: Evidence from Medicare Part D”. In: *The quarterly journal of economics* 130.2, pp. 841–899.
- Elbadawi, Ibrahim, A Ronald Gallant, and Geraldo Souza (1983). “An elasticity can be estimated consistently without a priori knowledge of functional form”. In: *Econometrica: Journal of the Econometric Society*, pp. 1731–1751.
- Ellis, R. P., B. Martins, and W. Zhu (2017). “Health care demand elasticities by type of service”. In: *Journal of Health Economics* 55, pp. 232–243.
- Ellison, Sara Fisher, Iain Cockburn, Zvi Griliches, and Jerry Hausman (1997). “Characteristics of demand for pharmaceutical products: an examination of four cephalosporins”. In: *The RAND journal of Economics*, pp. 426–446.
- Executive Office of the President (May 2025). *Executive Order 14297: Delivering Most-Favored-Nation Prescription Drug Pricing to American Patients*. <https://www.whitehouse.gov/presidential-actions/2025/05/delivering-most-favored-nation-prescription-drug-pricing-to-american-patients/>. 90 Fed. Reg. 20819 (May 15, 2025).
- Fan, Jianqing and Yuan Liao (2014). “Endogeneity in high dimensions”. In: *Annals of statistics* 42.3, p. 872.
- Filson, Darren (2012). “A Markov-perfect equilibrium model of the impacts of price controls on the performance of the pharmaceutical industry”. In: *The RAND Journal of Economics* 43.1, pp. 110–138.
- Gatwood, Justin, Teresa B Gibson, Michael E Chernew, Amanda M Farr, Emily Vogtmann, and A Mark Fendrick (2014). “Price elasticity and medication use: cost sharing across multiple clinical conditions”. In: *Journal of Managed Care Pharmacy* 20.11, pp. 1102–1107.
- Gilman, B. H. and J. Kautter (2008). “Impact of Multitiered Copayments on the Use and Cost of Prescription Drugs among Medicare Beneficiaries”. In: *Health Research and Educational Trust* 43.3, pp. 478–495.
- Goldman, D. P., G. F. Joyce, J. J. Escarce, J. E. Pace, M. D. Solomon, M. Laouri, P. B. Lansman, and S. M. Teutsch (2004). “Pharmacy Benefits and the Use of Drugs by the Chronically III”. In: *JAMA The Journal of the American Medical Association* 291.19, pp. 2344–2350.
- Goldman, Dana P, Geoffrey F Joyce, and Yuhui Zheng (2007). “Prescription drug cost sharing: associations with medication and medical utilization and spending and health”. In: *Jama* 298.1, pp. 61–69.

- Goldsmith-Pinkham, Paul, Peter Hull, and Michal Kolesár (2024). “Contamination bias in linear regressions”. In: *American Economic Review* 114.12, pp. 4015–4051.
- Granlund, David and Mats A. Bergman (2018). “Price competition in pharmaceuticals—evidence from 1303 Swedish markets”. In: *Journal of health economics* 61, pp. 1–12.
- Granlund, David and Niklas Rudholm (2011). “Consumer information and pharmaceutical prices: Theory and evidence”. In: *Oxford Bulletin of Economics and Statistics* 73.2, pp. 230–254.
- Griliches Z., Cockburn I. (Feb. 1993). “GENERICS AND NEW GOODS IN PHARMACEUTICAL PRICE INDEXES”. In: Working Paper Series 4272. URL: <https://www.nber.org/papers/w4272>.
- Gugler, Klaus, Lukas Pirnbacher, and Florian Szücs (2026). “Market power and regulation in pharmaceutical markets”. In: *Journal of the European Economic Association*, jvag010.
- Hall, Robert E and Charles I Jones (2007). “The value of life and the rise in health spending”. In: *The Quarterly Journal of Economics* 122.1, pp. 39–72.
- Hanoch, Giora (1975). “Production and demand models with direct or indirect implicit additivity”. In: *Econometrica: Journal of the Econometric Society*, pp. 395–419.
- HHS, U.S. Dept. (2025). *HHS, CMS set most-favored-nation pricing targets to end global freeloading on American patients*.
- Huang, Dongling and Christian Rojas (2013). “The Outside Good Bias in Logit Models of Demand with Aggregate Data”. In: *Economics Bulletin* 33.1, pp. 198–206.
- (2014). “Eliminating the Outside Good Bias in Logit Models of Demand with Aggregate Data.” In: *Review of Marketing Science* 12.1.
- Jack, W. and J. O. Lanjouw (2005). “Financing Pharmaceutical Innovation: How Much Should Poor Countries Contribute?” In: *The World Bank Economic Review* 19.1, pp. 45–67.
- Janssen, Aljoscha and David Granlund (2023). “The importance of the first generic substitution: Evidence from Sweden”. In: *Journal of Economic Behavior & Organization* 213, pp. 1–25.
- Jin, Hua (2026). “Estimation of BLP models with high-dimensional controls”. In: *arXiv preprint arXiv:2605.01594*.
- Joskow, Paul L (2007). “Regulation of natural monopoly”. In: *Handbook of law and economics* 2, pp. 1227–1348.
- Landsman, P. B., W. Yu, X. Liu, S. M. Teutsch, and M. L. Berger (2005). “Impact of 3-tier pharmacy benefit design and increased consumer cost-sharing on drug utilization”. In: *The American Journal of Managed Care* 11.10, pp. 621–628.
- Lewbel, Arthur and Krishna Pendakur (2009). “Tricks with Hicks: The EASI demand system”. In: *American Economic Review* 99.3, pp. 827–863.
- Lichtenberg, F. (2011). “Pharmaceutical Companies’ Variation Of Drug Prices Within And Among Countries Can Improve Long-Term Social Well-Being”. In: *Health Affairs* 30.8, pp. 1539–1544.
- Lichtenberg, Frank R (2010). “Pharmaceutical price discrimination and social welfare”. In: *Capitalism and Society* 5.1.
- Lieber, Jonas, Pietro Tebaldi, and Alexander Torgovitsky (2026). “Counterfactuals in Discrete Choice Models Under Flexible Assumptions on Random Utility”. In:
- Lundin, Douglas (2000). “Moral hazard in physician prescription behavior”. In: *Journal of health economics* 19.5, pp. 639–662.
- Maini, Luca and Fabio Pammolli (2023). “Reference pricing as a deterrent to entry: evidence from the European pharmaceutical market”. In: *American Economic Journal: Microeconomics* 15.2, pp. 345–383.
- Maleug, D. A. and M. Schwartz (1994). “Parallel imports, demand dispersion, and international price discrimination”. In: *Journal of International Economics* 34, pp. 167–195.

- Manning, W. G., J. P. Newhouse, N. Duan, E. B. Keeler, and Ar. Leibowitz (1987). "Health Insurance and the Demand for Health Care". In: *American Economic Review* 77.3, pp. 251–277.
- Maskus, K. E. (2001). "Parallel Imports in Pharmaceuticals Implications for Competition and Prices in Developing Countries". In:
- Morel, C. M., A. McGuire, and E. Mossialos (2011). "The Level Of Income Appears To Have No Consistent Bearing On Pharmaceutical Prices Across Countries". In: *Health Affairs* 30.8, pp. 1545–1552.
- Nevo, Aviv (2000). "A practitioner's guide to estimation of random-coefficients logit models of demand". In: *Journal of economics & management strategy* 9.4, pp. 513–548.
- (2001). "Measuring market power in the ready-to-eat cereal industry". In: *Econometrica* 69.2, pp. 307–342.
- Parasrampur, Sonal and Stephen Murphy (2024). "Comparing US and international market size and average pricing for prescription drugs, 2017-2022". In:
- Petrou, P. and S. Vondros (2016). "Pharmaceutical price comparisons across the European Union and relative affordability in Cyprus". In: *Health Policy and Technology* 5.4, pp. 350–356.
- Ramsey, F. P. (1927). "A Contribution to the Theory of Taxation". In: *The Economic Journal* 37.145, pp. 47–61. ISSN: 00130133, 14680297. URL: <http://www.jstor.org/stable/222721>.
- Reynaert, Mathias and Frank Verboven (2014). "Improving the performance of random coefficients demand models: The role of optimal instruments". In: *Journal of econometrics* 179.1, pp. 83–98.
- Riccaboni, Massimo, Torben Swoboda, and Walter Van Dyck (2022). "Pharmaceutical net price transparency across european markets: insights from a multi-agent simulation model". In: *Health Policy* 126.6, pp. 534–540.
- Sanderson, Eleanor and Frank Windmeijer (2016). "A weak instrument F-test in linear IV models with multiple endogenous variables". In: *Journal of econometrics* 190.2, pp. 212–221.
- Sawada, Masayuki and Kohei Kawaguchi (2020). "Estimating High-Dimensional Discrete Choice Model of Differentiated Products with Random Coefficients". In: *arXiv preprint arXiv:2004.08791*.
- Semenova, Vira, Matt Goldman, Victor Chernozhukov, and Matt Taddy (2023). "Inference on heterogeneous treatment effects in high-dimensional dynamic panels under weak dependence". In: *Quantitative Economics* 14.2, pp. 471–510.
- Sherer, F. M. and J. Watal (2002). "Post-TRIPS Options for Access to Patented Medicines in Developing Nations". In: *Journal of International Economic Law* 5.4, pp. 913–939.
- Shrank, William H, Nitesh K Choudhry, Joshua N Liberman, and Troyen A Brennan (2011). "The use of generic drugs in prevention of chronic disease is far more cost-effective than thought, and may save money". In: *Health affairs* 30.7, pp. 1351–1357.
- Srivastava, Divya and Alistair McGuire (2014). "Analysis of prices paid by low-income countries-how price sensitive is government demand for medicines?" In: *BMC Public Health* 14.1, p. 767.
- Tebaldi, Pietro, Alexander Torgovitsky, and Hanbin Yang (2023). "Nonparametric estimates of demand in the california health insurance exchange". In: *Econometrica* 91.1, pp. 107–146.
- Thaver, Ali, Mark Chee, Kristen Wroblewski, Arlene Weissman, Michelle Cook, James Zhang, Anita Samarth, Christopher Moriates, September Wallingford, Shalini Lynch, et al. (2020). "Physician and Nurse Practitioner Attitudes on Generic Prescribing of Oral Contraceptive Pills and Antidepressants: Thaver et al." In: *Journal of General Internal Medicine* 35.12, pp. 3478–3484.
- Towse, A., A. Cole, and B. Zamora (2018). "The Debate on Indication-Based Pricing in the U.S. and Five Major European Countries". In: *OHE Consulting Report*.
- Towse, A., M. Pistollato, J. Mestre-Ferrandiz, Z. Khan, S. Kaura, and L. Garrison (2015). "European Union Pharmaceutical Markets: A Case for Differential Pricing?" In: *International Journal of the Economics of Business* 22.2, pp. 263–275.

- Van de Geer, Sara, Peter Bühlmann, Ya'acov Ritov, and Ruben Dezeure (2014). "On asymptotically optimal confidence regions and tests for high-dimensional models". In:
- Vogelsang, Ingo (2003). "Price regulation of access to telecommunications networks". In: *Journal of Economic Literature* 41.3, pp. 830–862.
- Waning, B., W. Kaplan, A. C. King, D. A. Lawrence, H. G. Leufkens, and M. P. Fox (2009). "Global strategies to reduce the price of antiretroviral medicines: Evidence from transactional databases". In: *Bulletin of the World Health Organisation* 87, pp. 520–528.
- Yadav, P. (2010a). "Differential pricing for pharmaceuticals". In: *UK Department for International Development*. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/67672/diff-pcing-pharma.pdf.
- Yadav, Prashant (2010b). "Differential pricing for pharmaceuticals". In: *DFID: London*.
- Ye, Bei and Daniel R van Langenberg (2015). "Mesalazine preparations for the treatment of ulcerative colitis: are all created equal?" In: *World journal of gastrointestinal pharmacology and therapeutics* 6.4, p. 137.
- Yeung, Kai, Anirban Basu, Ryan N Hansen, and Sean D Sullivan (2018). "Price elasticities of pharmaceuticals in a value based-formulary setting". In: *Health economics* 27.11, pp. 1788–1804.