

Volume 43, Issue 3

A cost-benefit analysis of the medicines patent pool

Lucy Xiaolu Wang

University of Massachusetts Amherst; Max Planck Institute for Innovation & Competition; CCHE Canada

Abstract

Understanding the cost and benefit of global public health institutions is important but challenging. This study provides a cost-benefit analysis of the first public health-oriented patent pooling and licensing institution, the Medicines Patent Pool (MPP), which is devoted to improving generic drug licensing and supply in low- and middle-income countries. A simple structural model of demand and supply is estimated on a dataset that covers 103 LMIC and 29 HIV drugs with data on sales, MPP licenses, patents, country-year level diseases and demographics, and institutional factors during 2007-2017. Counterfactuals are simulated in the absence of the MPP or with further expansions. The estimated benefits to consumers and firms far exceed the operating costs.

Citation: Lucy Xiaolu Wang, (2023) "A cost-benefit analysis of the medicines patent pool", *Economics Bulletin*, Volume 43, Issue 3, pages 1298-1319

Contact: Lucy Xiaolu Wang - xiaoluwang@umass.edu.

Submitted: January 30, 2023. **Published:** September 30, 2023.

1. Introduction

Understanding global public health institutions is increasingly important given the lack of rapid and affordable drug access in Low and Middle-Income Countries (LMIC) during recent pandemics (e.g., AIDS and Covid-19). Among many advances, the Medicines Patent Pool (MPP) has made much progress in licensing Covid-19 technologies and, most recently, cancer drugs.¹ As the first-ever public health-oriented voluntary intellectual property (IP) licensing institution, the MPP was founded in 2010 to target HIV/AIDS drugs and has since expanded into many disease areas. The MPP negotiates with branded pharmaceutical firms to voluntarily pool patents and allow streamlined, low-cost generic licensing for sales in LMIC. Patent pools have been widely used in many industries to ease licensing hurdles and spur technology diffusion (Shapiro, 2001; Lerner and Tirole, 2004; Lampe and Moser, 2013), but a cost-benefit analysis is not always feasible given data limitations. As a modern pool with open data, the MPP limits sales territories to LMIC and increases drug access without reducing R&D (Wang, 2022), so a welfare analysis for LMIC drug access would be informative regarding the overall welfare impact.

Understanding the welfare impact of an institution in the context of LMIC drug access is challenging. In the case of HIV/AIDS, the disease cannot be cured, but daily take of HIV drug cocktails can turn HIV from fatal to a chronic condition. It is worth noting that HIV drug supply is critical, with almost 30 million patients living with HIV relying on daily drug cocktails. The daily drug cocktail reliance and chronic nature make it difficult to directly estimate lives saved, especially when many factors in the supply chain, distribution, and patient behavior can affect mortality. All else equal, improved generic drug supply can increase drug access for patients in need and extend their productive life span, which can be better captured by consumer surplus in the absence of more accurate measures. The MPP has successfully engaged with branded drug providers: four out of the total nine branded HIV drug providers available by 2017 joined the MPP.² An economic analysis of social welfare and organizational costs would be valuable to understand the return to the MPP, and the framework can be extended for future scenarios.

This paper uses a structural model to estimate the welfare effects of the MPP and conduct a cost-benefit analysis using its operating costs. Specifically, I estimate a nested discrete choice demand model to capture the feature that compounds are more substitutable within the same drug class and HIV drug choices are sequential (Figure A1). Physicians first choose which drug class (based on whether a patient is treatment-naïve and any known resistance), then select which drugs within a chosen class can be used (e.g., based on different side effects). Note that the pool does not directly enter patients' preferences towards a drug but will make it cheaper to produce. Two extreme cases are estimated for the supply side, competitive pricing and Bertrand-Nash oligopolistic pricing, and I then simulate counterfactuals without the MPP or with more aggressive expansion. I then use estimates from the demand and supply models to simulate counterfactuals: what if the MPP did not exist at all? What if the MPP has implemented an expansion to the broadest geographic coverage possible? I obtain the changes in consumer and producer surplus accordingly and compare them with the operating costs of the pool afterward.

¹ Sources: 1) all licenses: <https://medicinespatentpool.org/progress-achievements/licences>; 2) Covid-related: <https://medicinespatentpool.org/covid-19>; 3) latest MPP license: <https://medicinespatentpool.org/news-publications-post/mpp-signs-licence-agreement-to-increase-access-to-nilotinib-for-the-treatment-of-chronic-myeloid-leukaemia>.

² The four branded entities including the two market leaders, Gilead Science and ViiV Healthcare (i.e., a joint venture of GlaxoSmithKline, Pfizer, and Shionogi), and in addition, AbbVie Inc. and Bristol-Myers Squibb.

The model is estimated on a comprehensive dataset that covers 103 LMIC and 29 HIV drugs with data on sales, MPP licenses, patents, country-year level diseases and demographics, and institutional factors between 2007-2017. The dataset is constructed based on various resources, including drug profiles and dosing information from AIDSinfo.gov, time-varying MPP coverage information from its official website and licensing contracts, drug sales data from the Global Fund's Price and Quality Reporting, international drug patent data from the MPP patent portal (MedsPAL) and commercial source Drug Patent Watch, country-year level control variables from the World Bank and the Institute for Health Metrics and Evaluation. Wang (2022) documented more detailed information on the data construction. The MPP is estimated to increase welfare substantially compared to no-MPP cases. Across analyses, consumer surplus increases by \$0.7-1.4 billion (8.6-18.9%), and producer surplus can also increase by up to \$181 million (4.5%). The welfare gains far exceed the \$33 million operating cost in the same period.

This paper contributes to the literature on drug diffusion to LMIC with a new angle using simple structural models to facilitate cost-benefit analyses with application to a health institution. Maintaining an affordable and stable drug supply has been particularly challenging in developing countries (Kremer, 2002), which can be facilitated by procurement institutions (Danzon et al., 2015; Dubois et al., 2021; Wang and Zahur, 2021). Institutions facilitating resource pooling can improve welfare, but a prior study of the US context shows that the welfare effect is ambiguous in merge simulation with drug cocktails (Song *et al.*, 2017). Extensive patenting makes licensing difficult (Hemphill and Sampat, 2012). Prior studies find that the MPP increases drug diffusion, but do not carry out a full cost-benefit analysis of the institution (Wang, 2022; Galasso and Schankerman, 2022). This paper offers the first systematic cost-benefit analysis based on estimation of supply and demand in a large sample, which is important to inform future policies regarding new health institution building and expansion.

This article is organized as follows. Section 2 presents the model and estimation. Section 3 reports the results and counterfactual simulations. Section 4 concludes. The appendix provides more technical details and empirical results.

2. Model and Estimation

Quantifying the cost-benefit of a patent pool is difficult, as most products are sold globally despite relatively concentrated innovation activities. The MPP separates the market into sales territories and increases access and R&D (Wang, 2022), so a welfare analysis for LMIC in the diffusion sample is informative yet conservative. Another typical challenge is to characterize *ex ante* substitution patterns across technologies, which is feasible in my setting using drug classes. The decision of HIV drug use follows sequential choices: first, the choice of a drug class, and then a choice of drugs within the chosen class. The MPP would not directly enter patients' preferences towards a drug (as most patients in LMIC would not be familiar with the MPP) but would make the drug cheaper to produce. Including the MPP as a cost shifter is tied with prior findings and the idea that the pool lowers licensing and royalty costs (per-unit based).

I observe drug purchases but not how patients use them, which limits the richness of the demand model I can estimate. Additionally, my sample period (2007-2017) is insufficiently long to model new drug launches, resulting in conservative welfare estimations. I also abstract away from branded-generic interactions with the drug-country-year level aggregate data, as it is limited by data and not a core part of the research question. This exercise is best viewed as an empirical illustration of drug cocktails in the context of a patent pool, which is useful to understand the welfare impact of a biomedical patent pool and examine its cost versus benefit.

2.1 Demand-Side: Nested Discrete Choice Model

Patients in each country choose whether to take drugs and choose among available choices following physicians' suggestions. The choice sets of drugs vary across countries and over time. D_{ct} denotes the set of HIV drugs available from my dataset in country c at year t , and index drug by j . Each patient-physician pair in a country-year decides whether to take a drug in the choice set $D_{ct} = \{1, 2, \dots, d_{ct}\}$ or to use an outside option: either taking no drugs, or those from sources outside my sample. Each drug j belongs to drug class $g(j)$ based on its mechanism of action. Specifically, each patient-physician pair i in country c at year t chooses drug j from the $D_{ct} + 1$ options to maximize the conditional indirect utility function as below:

$$u_{ijct} = \underbrace{X_{jct}\beta - \alpha p_{jct} + \xi_{jct}}_{\delta_{jct}} + \zeta_{ig(j)ct} + (1 - \sigma)\varepsilon_{ijct}. \quad (1)$$

Here, δ_{jct} is the mean utility of drug j in country-year market ct . X_{jct} is a vector of observables, including the number of compounds in a drug, number of distinct products for a drug in each country-year, HIV prevalence and death rates, income, log(population), institutional factors, and fixed effects at country and year levels. p_{jct} is the average price of drug j in country-year ct . ξ_{jct} is the unobserved quality of drug j for patients in country c in year t . ε_{ijct} is an independent taste shock following the extreme value distribution, $\zeta_{ig(j)ct}$ is a group-specific taste of the patient-physician pair, and $(1 - \sigma)$ measures the relative weight of idiosyncratic and group preferences.

Given the nested logit functional form, the market share of each product for any set of product qualities δ_j can be calculated à la Berry (1994).³ The estimation equation is as follows:

$$\ln(s_j) - \ln(s_0) = \sigma \ln(s_{j|g(j)}) + \underbrace{X_j\beta - \alpha p_j + \xi_j}_{\delta_j - \delta_0}, \quad (2)$$

where $s_{j|g(j)}$ is drug j 's share of its nest in a given country-year market ct . I estimate the model using two-stage least square with instruments for conditional market share $\ln(s_{j|g(j)})$ and price p_j . Three instruments are included: patent status, the number of manufacturers for a drug in a country-year, and the number of distinct competing products (i.e., drug product and manufacturer combinations for other drugs in the same drug class). Patent status IV is relevant as a patented drug (in a country in a given year) is usually priced higher, and the patent status would not affect consumer preference directly conditional on using the same drug, as most patients would not generate specific utility on the patent status itself. The latter two IVs are the standard "BLP instruments" following Berry *et al.* (1995).⁴ Standard errors are clustered at the drug-country level. Using parameters from the demand estimation, the nested logit model allows researchers to calculate the expected consumer surplus in each market using equation (3) below:

³ A plain logit model is also implemented to test the robustness of demand specifications. The market size is defined as 30% of the population. The results are robust to a set of other definitions (Table A3).

⁴ The BLP instruments are valid as long as the drug characteristics are exogenous. The justification is that drug characteristics are determined at the time of trials; once approved, firms cannot change the characteristics in the short-run. The instruments are relevant, as a drug that has close competitors (drugs within the same drug class work through the same mechanism of action) is likely to be priced lower.

$$E(CS_{ct}) = \frac{1}{\alpha} \ln \left\{ \sum_g \left[\sum_{j \in g} \exp \left(\frac{\delta_j}{1-\sigma} \right) \right]^{1-\sigma} \right\}. \quad (3)$$

Here, the choices among nests (i.e., drug classes) depend only on their mean utilities $\delta_j = X_j\beta - \alpha p_j + \xi_j$, as there are no nest-specific observables entering the utility function. The total consumer surplus during the sample period can be calculated as the summation of market-size-weighted expected consumer surplus values across markets. Therefore, the baseline consumer surplus with the MPP can be calculated independently of the supply-side estimation.

2.2 Supply-Side: Bertrand-Nash Game

In the absence of the universe of data and a proper model on oligopolistic pricing with a competitive fringe and threats of compulsory licensing in over a hundred LMIC, two simple and conservative cases are modeled. The first one models supply using constant marginal cost pricing as in Bertrand competition. The second uses a Bertrand-Nash game in an oligopolistic setting. In both cases, the MPP affects pricing by lowering marginal costs as the MPP lowers royalty rates in joint licensing are also typically calculated on a per-unit basis.⁵

In the first case, a marginal cost pricing scheme is estimated to capture the feature that generic firms produce at marginal costs, and generic market shares are high in LMIC. The results from previous parts support competitive pricing in that generic shares are high in base levels and increase with the MPP. I use j to denote drug-country pair and estimate the following equation.

$$p_j = mc_j = X_j\gamma + \beta MPP_j + \omega_j \quad (4)$$

Here, MPP indicates whether a drug-country is in the MPP in a year. X_j is a vector of drug-country-year level controls including whether a drug is effectively patented in a country-year, number of drug products and competitors for a drug in a country-year, country-year level HIV prevalence and age-adjusted death rates, population, income, institutional factors, country and year fixed effects. Quantity is not included in this baseline case because firms do not face capacity constraints in a competitive market with many generic firms. Once the pricing equation is estimated, counterfactual prices can be simulated by changing the MPP indicator.

In the second case, oligopolistic pricing is estimated in a static Bertrand–Nash game with differentiated products. Firms can have market power in the short run before reaching the long-run competitive equilibrium. Each firm acts as either a single-product or multi-product drug firm where each drug belongs to a nest (i.e., drug class). This design captures the feature that the brand ownership of a drug does not matter as much for generic drug makers, but the “business stealing effect” comes from drugs with similar characteristics (i.e., drug class). Each country-year is a market, and a firm maximizes profit from all drugs they produce using the equation:

$$\Pi = \sum_j (p_j - mc_j) \times s_j(p_j) \times M. \quad (5)$$

⁵ Due to the lack of complete HIV drug sales, I cannot back out the fixed costs and separately identify additional MPP-induced cost reduction. This suggests cost results are likely conservative. In common with prior work (Song *et al.*, 2017), I model drug pricing as a static game without sufficiently rich data.

Here, j indexes drug j in a country-year market, and M represents the same market size measure used in the demand estimation. Note that the substitution patterns between drugs j and k are affected by their corresponding drug classes g_j and g_k . The first-order condition yields:

$$s_j(p_j) + \sum_k (p_k - mc_k) \times \frac{\partial s_k(p)}{\partial p_j} = 0. \quad (6)$$

The price derivative matrix is defined as $\Delta_{jk} = -\frac{\partial s_k}{\partial p_j}$, a la Berry *et al.* (1995). Here, f_j indicates that drug j is owned by firm f , and two ownership structures are estimated. First, a single-product oligopoly is estimated, where each representative firm owns one product. Second, a multi-product oligopoly with drug-level ownership is estimated, where a cross-firm cocktail owner is assigned to a separate entity. Appendix B provides the derivation and additional results.

$$\Delta_{jk} = \begin{cases} \alpha s_j \left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma} s_{j|g} - s_j \right) & , j = k \\ -\alpha s_j \left(\frac{\sigma}{1-\sigma} s_{k|g} + s_k \right) & , j \neq k, g_j = g_k, f_j = f_k \\ -\alpha s_j s_k & , j \neq k, g_j \neq g_k, f_j = f_k \\ 0 & , o.w. (i.e., f_j \neq f_k) \end{cases} \quad (7)$$

Rewriting the F.O.C as $s - \Delta[p - mc] = 0$; it follows that $p = mc + \Delta^{-1}s$. The new estimation equation is as below, suppressing the country-year subscript and indexing drug with j :

$$mc_j = p_j - \Delta^{-1}s_j = X_j\gamma + \beta MPP_j + \eta q_j + \omega_j. \quad (8)$$

The covariate X_j is specified the same as in equation (4), except that quantity q is included in the regression to capture firms' capacity constraints in oligopolistic settings. q is instrumented using the number of competing products in the same class (Berry *et al.*, 1995; Song *et al.*, 2017).

3 Results and Counterfactuals

Demand estimation results are consistent with model assumptions (Table A1). In the nested logit model, the coefficient estimate of within-group market share is 0.863 and significant at the 1% level. This result supports the assumption that drugs are closer substitutes within a group than between groups. A comparison of price coefficients in OLS and the nested logit model reveals a positive correlation between price and demand shocks and that the instrumental variables mitigate the problem. The price coefficient estimates increase in absolute value from -0.137 in the OLS model to -1.946 in the nested logit model, and the estimates are always statistically significant at the 1% level.⁶

Supply-side estimates suggest that the MPP reduced drug costs, and the result is robust to various assumptions concerning market structure. Results under constant marginal cost pricing are similar to those with oligopolistic pricing and suggest significant MPP-induced cost reductions at the 1% level (Table A2). Specifically, the MPP reduces cost by \$0.6 per patient-day

⁶ The first-stage Kleibergen–Paap F-statistic for both endogenous variables is 19.30, and the Sanderson–Windmeijer multi-variate first-stage F-statistics for each endogenous variable are much larger.

under competitive pricing and by \$1.9 under oligopolistic pricing across ownership assumptions. The estimated quantity coefficients are positive yet small and are consistent with supply curves being upward sloping, as oligopolistic firms face some capacity constraint as quantity goes up. In cases where marginal costs are not independent of quantity, patent protection increases costs, and the number of competitors in the market reduces marginal costs. These results are robust to the inclusion of country-year level observable controls and to statistical tests on the instruments.

Parameters from demand and supply estimation are then used to simulate counterfactuals in which the MPP does not exist or expand. For example, the MPP parameters are shut down in the estimated oligopolistic marginal cost curve. Counterfactual equilibrium prices and quantities are computed by solving the new first-order condition and updating for each country-year market. New consumer and producer surpluses are then computed (see Appendix B for details).

The welfare and counterfactual results quantify how much the MPP increases consumer and producer surpluses (Table 1). Consumers benefit from MPP-induced cost savings via joint licensing, and producers can benefit from market expansion in LMIC and increased sales via the generic network. The estimated current consumer surplus across all markets (country-year) is \$8.7 billion, on the same order of magnitude as the total revenue during the sample period (\$4.3 billion) and consistent with demand being elastic. Compared to the counterfactual cases without a pool, the MPP results in \$0.69-1.39 billion consumer surplus gains across market structure specifications. These are equivalent to 8.60-18.94% increases from the counterfactual consumer surpluses. The estimated producer surplus gains are up to \$181 million and are consistent with theory in that producers can also benefit from cost-reducing technologies.

Table 1: Welfare Estimation: Consumer & Producer Surpluses

welfare estimates (\$ M)	MC	Bertrand-Nash Oligopoly	
	pricing flat MC	single-prod. firm	multi-prod. firm
\widehat{CS}_0	7,354.5	8,055.1	8,025.8
\widehat{CS}	8,747.7	8,747.7	8,747.7
\widehat{CS}_1	8,883.3	8,821.5	8,816.4
$\Delta\$: \widehat{CS}_0$	1,393.2	692.6	721.9
$\Delta\% : \widehat{CS}_0$	18.94%	8.60%	8.99%
$\Delta\$: \widehat{CS}_1$	135.6	73.8	68.7
$\Delta\% : \widehat{CS}_1$	1.55%	0.84%	0.78%
\widehat{PS}_0	0	3,998.1	4,194.8
\widehat{PS}	0	4,179.5	4,309.6
\widehat{PS}_1	0	4,320.7	4,462.0
$\Delta\$: \widehat{PS}_0$	0	181.4	114.8
$\Delta\% : \widehat{PS}_0$	0	4.54%	2.74%
$\Delta\$: \widehat{PS}_1$	0	141.2	152.4
$\Delta\% : \widehat{PS}_1$	0	3.38%	3.54%

Notes: This table reports estimating the gains in consumer and producer surpluses with the MPP. Here, \widehat{CS} denotes the total consumer surplus cross-market. \widehat{CS}_0 is the counterfactual consumer surplus without the MPP. Similarly, \widehat{PS} and \widehat{PS}_0 denote the actual and counterfactual producer surpluses, respectively. \widehat{CS}_1 and \widehat{PS}_1

denote the counterfactual consumer and producer surpluses with a fully expanded MPP (covering all developing countries in the sample for each compound in the pool), respectively.

The costs of operating the patent pool can be incorporated using a back-of-the-envelope cost-benefit type analysis using audited financial statements from the MPP (Table A4). The total operating cost of the MPP is about \$33 million from its establishment in July 2010 to the end of 2017. This amount is small compared to the welfare gains estimated above. These counterfactual results, together with the diffusion and innovation analyses, suggest that the MPP is an effective business model in the LMIC setting with elastic demand for drugs treating infectious diseases. Lower prices from generic access generate larger responses in quantity increases, which can further increase profits. This calculation is conservative as both branded and generic firms also benefit from reduced administrative costs that are not incorporated here.

4 Conclusion

Based on a structural model of demand and supply in a sample of 103 LMIC over 2007-2017, I find that the MPP increases consumer surplus by about \$0.7 billion (9%) and can increase producer surplus by up to \$181 million (4.5%), far exceeding the \$33 million operating costs in the same period. The benefit-cost ratio is 25.4-26.5 based on social welfare, and 21 and 3.5-5.5 based on consumer and producer surplus, respectively. Overall, the MPP increases welfare by lowering the cost of licensing and offering new models of marketing in underdeveloped LMIC markets in a cost-effective manner. My welfare estimations are larger than those reported in the “back of envelope calculation” by Galasso and Schankerman (2022), because their calculation is based on a smaller sample of 32 countries and with roughly unit elastic demand. There are some limitations. My sample period is insufficiently long to model new drug launches, resulting in conservative welfare calculations. Thus, this exercise is mainly as an empirical cost-benefit evaluation of the Medicines Patent Pool on LMIC drug access.

Although this paper focuses on HIV drugs due to data availability, the idea can be applied more broadly. It is worth noting that HIV/AIDS and its comorbidities are estimated to yield larger disease burdens than Covid-19 in many LMIC given the care disruption caused by Covid-19 (Bell and Hansen, 2021). In fact, many institutions and investments for the AIDS pandemic are critical resources to fight Covid-19 in LMIC.⁷ In addition, federal agencies and practitioners have shown growing interests in building more biomedical patent pools (Clark et al., 2000; Van Overwalle, 2016), and understanding the cost and benefit is very important for resource allocation. A structural model-based cost-benefit estimation contributes to the endeavor with monetary measures of both cost and benefit without the controversies of comparing money to lives. Overall, this paper contributes to understanding novel institutions to improve drug supply.

⁷ Source: https://www.unaids.org/sites/default/files/media_asset/20200909_Lessons-HIV-COVID19.pdf.

References

- Bell, D., & Hansen, K. S. (2021). Relative burdens of the COVID-19, malaria, tuberculosis, and HIV/AIDS epidemics in sub-Saharan Africa. *The American Journal of Tropical Medicine and Hygiene*, 105(6), 1510.
- Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242-262.
- Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841-890.
- Clark, J., Piccolo, J., Stanton, B., Tyson, K., Critharis, M., & Kunin, S. (2000). *Patent pools: a solution to the problem of access in biotechnology patents?*. mimeo, USPTO.
- Danzon, P. M., Mulcahy, A. W., & Towse, A. K. (2015). Pharmaceutical pricing in emerging markets: effects of income, competition, and procurement. *Health economics*, 24(2), 238-252.
- Dubois, P., Lefouili, Y., & Straub, S. (2021). Pooled procurement of drugs in low and middle income countries. *European Economic Review*, 132, 103655.
- Galasso, A., & Schankerman, M. (2022). *Licensing life-saving drugs for developing countries: Evidence from the medicines patent pool* (No. w28545). National Bureau of Economic Research.
- Hemphill, C. S., & Sampat, B. N. (2012). Evergreening, patent challenges, and effective market life in pharmaceuticals. *Journal of health economics*, 31(2), 327-339.
- Kremer, M. (2002). Pharmaceuticals and the developing world. *The Journal of Economic Perspectives*, 16(4), 67-90.
- Lampe, R., & Moser, P. (2013). Patent pools and innovation in substitute technologies—evidence from the 19th-century sewing machine industry. *The RAND Journal of Economics*, 44(4), 757-778.
- Lerner, J., & Tirole, J. (2004). Efficient patent pools. *The American Economic Review*, 94(3), 691-711.
- Shapiro, C. (2001). Setting compatibility standards: cooperation or collusion. *Expanding the bounds of intellectual property*. University Press, Oxford, 8-0.
- Song, M., Nicholson, S., & Lucarelli, C. (2017). Mergers with interfirm bundling: a case of pharmaceutical cocktails. *The RAND Journal of Economics*, 48(3), 810-834.
- Van Overwalle, G. (2016). Patent Pools and Clearinghouses in the Life Sciences: Back to the Future. *Research Handbook on IP and the Life Sciences*, D. Matthews & H. Zech (eds.), Edward Elgar
- Wang, L. X. (2022). Global drug diffusion and innovation with the medicines patent pool. *Journal of Health Economics*, 85, 102671.
<https://doi.org/10.1016/j.jhealeco.2022.102671>
- Wang, L. X., & Zahur, N. (2021). Procurement Institutions and Essential Drug Supply in Low and Middle-Income Countries. Available at SSRN: <https://ssrn.com/abstract=3926761>

Online Appendix: Not for Publication

A Cost-Benefit Analysis of the Medicines Patent Pool

Appendix A: Figures and Tables

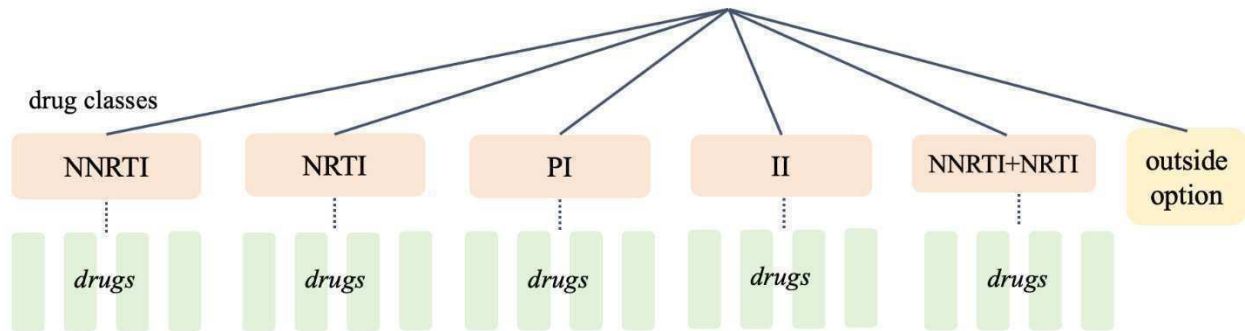


Figure A1: Two-level Nested Logit Decision Structure

Notes: This figure illustrates the sequential decision structure for the nested logit demand. The acronyms for the drug classes (based on mechanisms of action) are: NNRTI (Non-Nucleoside Reverse Transcriptase Inhibitors), NRTI (Nucleoside Reverse Transcriptase Inhibitors), PI (Protease Inhibitors), II (Integrase Inhibitors). Detailed information on drug within each drug class is available as in Table A1 and Appendix D (Medical Appendix) in Wang (2022). The outside option represents sales not captured by the sample.

Table A1: Results from the Demand Estimation

	(1)	(2)	(3)
	OLS	Nested logit	Logit
$\ln(s_{j g(j)})$	0.702*** (0.0144)	0.862*** (0.0814)	
p_j	-0.137*** (0.0227)	-1.946*** (0.243)	-3.483*** (0.441)
drug age (U.S. appr.)	0.0119* (0.00637)	-0.196*** (0.0404)	-0.449*** (0.0838)
prod. variety	0.345*** (0.0335)	-0.00503 (0.122)	0.434** (0.179)
regulatory quality	0.00194 (0.00558)	-0.0646*** (0.0208)	-0.121*** (0.0378)
rule of law	0.0226*** (0.00546)	0.0507*** (0.0162)	0.0532* (0.0291)
control of corruption	-0.00783* (0.00446)	0.0361** (0.0148)	0.0785*** (0.0272)
Kleibergen-Paap F statistic		19.50	
1 st stage ($s_{j g}$)		104.42	
1 st stage (p_j)		46.91	54.56
	country FE, year FE, and X_j controls		
Observations	7,084	7,084	7,084

Note: This table presents results of estimating the nested logit demand model as in Equation (2) and compares it with OLS and a plain logit. The instruments for conditional market share and price are: (1) whether a drug is effectively patented in the country-year, (2) the number of manufacturers for the same drug in a country-year, and (3) the number of competing products, i.e., drug product-firm combinations for other drugs in the same drug class. IVs for the plain logit do not include the second instrument to avoid over-identification. Only main parameters of interests are reported for simplicity. Observable controls, X_j , include within-drug product variety in a country-year, number of compounds within a drug, number of years since a drug's U.S. approval, country-year level HIV prevalence and age-adjusted death rates, institutional factors (i.e., the six world governance indicators), log(population) and GDP per capita. The excluded instruments are at drug-country-year level: patent status, number of competitors and number of close competitors in the same drug class. The first-stage statistics displayed immediately under coefficients-of-interests are the Kleibergen-Paap F statistic that are robust to heteroskedasticity. The first-stage F statistics for each endogenous variable is the Sanderson-Windmeijer multivariate F test of excluded instruments. Standard errors are clustered at drug-country level. Robust p-value: *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Estimations of Pricing Equations

Dept. var:	(1)	(2)	(3)
marginal cost (\$)	<u>MC</u> <u>pricing</u> flat MC	<u>Bertrand-Nash Oligopoly</u> single-prod. firm	multi-prod. firm
MPP_j	-0.642*** (0.112)	-1.908*** (0.524)	-1.952*** (0.539)
Q_j		3.60e-07*** (1.27e-07)	3.83e-07*** (1.31e-07)
#variety	-0.209*** (0.0616)	0.445* (0.234)	0.495** (0.244)
#firms _{det}	-0.310*** (0.0398)	-1.584*** (0.480)	-1.662*** (0.494)
Patent _{det}	-0.173 (0.192)	0.210 (0.255)	0.198 (0.262)
year FE	Y	Y	Y
country FE	Y	Y	Y
X _j controls	Y	Y	Y
Kleibergen-Paap rk Wald F-stat		16.66	16.66
Observations	7,084	7,084	7,084

Notes: This table reports the results from estimating competitive marginal cost pricing and oligopolistic pricing on the drug-country-year diffusion sample using Equations (4) and (8), respectively. Only main parameters of interests are reported for simplicity. X_j is a vector of drug-country-year level controls including whether a drug is effectively patented in a country-year, number of drug products and competitors for a drug in a country-year, country-year level HIV prevalence and age-adjusted death rates, population, GDP per capita, and institutional factors. Country and year fixed effects are always included. Quantity variable is instrumented by the number of competing products in the same drug class within a market (country-year). Standard errors are clustered at drug-country level. The first-stage F-statistics reported are adjusted for heteroskedasticity clustering. Robust p-value: *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Sensitivity Analysis of Demand Estimation to Market Size

	(1)	(2)	(3)	(4)	(5)
market size	10%	30%	50%	70%	pop*pr. HIV
measures	population	population	population	population	death _{15to59}
$\ln(s_{j g})$	0.862*** (0.0826)	0.861*** (0.0812)	0.861*** (0.0810)	0.861*** (0.0809)	0.863*** (0.0811)
p_j	-1.968*** (0.247)	-1.942*** (0.243)	-1.938*** (0.242)	-1.937*** (0.242)	-1.941*** (0.243)
1 st stage joint	19.50	19.50	19.50	19.50	19.50
1 st stage $(s_{j g})$	104.42	104.42	104.42	104.42	104.42
1 st stage (p_j)	46.91	46.91	46.91	46.91	46.91
country FE	Y	Y	Y	Y	Y
year FE	Y	Y	Y	Y	Y
X_j controls	Y	Y	Y	Y	Y
Observations	7,084	7,084	7,084	7,084	7,084

Note: This table presents results of estimating the nested logit demand model as in Equation (2), and each column demonstrates robustness of the estimation to alternative market size measures. Observable controls, X_j include within drug product variety in a country-year, number of compounds within a drug, number of years since a drug's US approval, country-year level HIV prevalence and age-adjusted death rates, institutional factors (i.e., the six world governance indicators), log(population) and GDP per capita. The excluded instruments are at drug-country-year level: patent status, number of competitors and number of close competitors in the same drug class. The first-stage statistics displayed immediately under coefficients-of-interests are the Kleibergen-Paap F statistic that are robust to heteroskedasticity. The first-stage F statistics for each endogenous variable is the Sanderson-Windmeijer multivariate F test of excluded instruments. Each j denotes drug-country (dc) for simplicity in notations. Robust p-value: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A4: Pool Operating Expenses from Financial Statement

time period	use of the funds (raw \$/SFr.)	CHF/USD (annual)	MPP costs (\$ current)
2010.7-2011.12	\$ -4,254,666	NA	-4,254,666
2012.1-2012.12	SFr. -4,086,052	0.9377	-4,357,526
2013.1-2013.13	SFr. -4,271,467	0.9269	-4,608,336
2014.1-2014.12	SFr. -4,332,580	0.9147	-4,736,613
2015.1-2015.12	SFr. -4,759,073	0.9628	-4,942,951
2016.1-2016.12	SFr. -4,568,395	0.9848	-4,638,906
2017.1-2017.12	SFr. -4,974,406	0.9842	-5,054,263

Notes: The MPP operating costs are obtained from the financial statements in the “Annual Reports” from the MPP. Specifically, “use of the funds” within the “statement of changes in capital” is used to measure the costs of this pool. This calculation is similar to manually summing up the personnel and administrative costs (the two main categories of MPP expenditure). The annual foreign exchange rate of Swiss Francs to one U.S. Dollar is provided by the Federal Reserve Bank of St. Louis.

Appendix B: Mathematical Appendix

B.1 Deriving the price substitution matrix

I derive the substitution matrix by taking partial derivatives of market share k w.r.t price j . Here, I derive the general expression for the price derivatives from the demand side. With information from the supply-side, the relevant elements from the matrix are the products owned by the same branded firm in a given market (i.e., subset products owned by the same firm).

Given that

$$\hat{s}_j = \frac{e^{\frac{\delta_j}{1-\sigma}} \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{-\sigma}}{\sum_{g=0}^G \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma}}, \quad \hat{s}_k = \frac{e^{\frac{\delta_k}{1-\sigma}} \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma}}{\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma}}, \quad \text{and} \quad \hat{s}_{k|g} = \frac{e^{\frac{\delta_k}{1-\sigma}}}{\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}}} = \frac{\hat{s}_k}{\hat{s}_g}$$

$$\begin{aligned} \frac{ds_k}{dp_j} &= \frac{d}{dp_j} \left[\frac{e^{\frac{\delta_k}{1-\sigma}} \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma}}{\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma}} \right] \\ &= \frac{\left(e^{\frac{\delta_k}{1-\sigma}} \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} \right)' \left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right] - e^{\frac{\delta_k}{1-\sigma}} \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} \left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]'}{\left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]^2} \\ &= \frac{\overbrace{\left(e^{\frac{\delta_k}{1-\sigma}} \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} \right)'}^{\equiv A}}{\left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]} - \frac{e^{\frac{\delta_k}{1-\sigma}} \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} \overbrace{\left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]'}^{\equiv B}}{\left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]^2} \\ &= \frac{A}{\left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]} - \hat{s}_k \times \frac{B}{\left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]}, \quad \text{eqn. (B1)} \end{aligned}$$

$$\text{where } A \equiv \left(e^{\frac{\delta_k}{1-\sigma}} \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} \right)' = \left(e^{\frac{\delta_k}{1-\sigma}} \right)' \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} + e^{\frac{\delta_k}{1-\sigma}} \left(\left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} \right)'$$

$$\text{and } B = \left[\left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]' \text{ if } g(k) = g(j); = \left[\left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma} \right]' \text{ if } g(k) \neq g(j).$$

In the following part, I derive the analytic forms of the price derivatives for three cases. In each case, I first derive the expressions for A and B and then plug them back into eqn. (B1).

Simplification note: that $\left(e^{\frac{\delta_k}{1-\sigma}} \right)' = 0$ if $j \neq k$ and $\left(\left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} \right)' = 0$ if $g(j) \neq g(k)$.

Case 1: $j=k$ (diagonal elements)

$$\begin{aligned}
A &= \left(e^{\frac{\delta_j}{1-\sigma}} \right)' \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{-\sigma} + e^{\frac{\delta_j}{1-\sigma}} \left(\left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{-\sigma} \right)' \\
&= -\frac{\alpha}{1-\sigma} e^{\frac{\delta_j}{1-\sigma}} \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{-\sigma} + e^{\frac{\delta_j}{1-\sigma}} \times \frac{\alpha\sigma}{(1-\sigma)} \times \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{-\sigma} \times \frac{\overbrace{\frac{\delta_j}{e^{\frac{\delta_j}{1-\sigma}}}}^{=\hat{s}_{j|g}}}{\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}}} \\
&= -\frac{\alpha}{1-\sigma} e^{\frac{\delta_j}{1-\sigma}} \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{-\sigma} (1 - \sigma \hat{s}_{j|g}) \\
&= -\frac{\alpha}{1-\sigma} \hat{s}_j (1 - \sigma \hat{s}_{j|g}) \times \left[\sum_{g=0}^G \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma} \right] \\
B &= \left[\left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma} \right]' = (1 - \sigma) \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{-\sigma} \left(e^{\frac{\delta_j}{1-\sigma}} \right)' \\
&= -\alpha e^{\frac{\delta_j}{1-\sigma}} \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{-\sigma} = -\alpha \hat{s}_j \times \left[\sum_{g=0}^G \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma} \right]
\end{aligned}$$

Plug back to equation (B1),

$$\begin{aligned}
\frac{ds_j}{dp_j} &= \frac{-\frac{\alpha}{1-\sigma} \hat{s}_j (1 - \sigma \hat{s}_{j|g}) \times \left[\sum_{g=0}^G \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma} \right]}{\left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]} - \hat{s}_j \times \frac{-\alpha \hat{s}_j \times \left[\sum_{g=0}^G \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma} \right]}{\left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]} \\
&= -\frac{\alpha}{1-\sigma} \hat{s}_j (1 - \sigma \hat{s}_{j|g}) - \alpha \hat{s}_j \hat{s}_j \\
&= -\alpha \hat{s}_j \left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma} \hat{s}_{j|g} + \hat{s}_j \right)
\end{aligned}$$

Case 2: $j \neq k, g(j) = g(k)$ (different alternatives within the same nest)

$$\begin{aligned}
A &= \overbrace{\left(e^{\frac{\delta_k}{1-\sigma}} \right)'}^{=0} \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} + e^{\frac{\delta_k}{1-\sigma}} \left(\left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} \right)' \\
&= e^{\frac{\delta_k}{1-\sigma}} (-\sigma) \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma-1} e^{\frac{\delta_j}{1-\sigma}} \times \frac{-\alpha}{1-\sigma} \\
&= \frac{\alpha\sigma}{1-\sigma} \times \frac{e^{\frac{\delta_k}{1-\sigma}}}{\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}}} \times \frac{\delta_j}{e^{\frac{\delta_j}{1-\sigma}}} \times \underbrace{\left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma}}_{\hat{s}_j \times \sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma}} = \frac{\alpha\sigma}{1-\sigma} \hat{s}_{k|g} \hat{s}_j \times \sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma}
\end{aligned}$$

$$B = \left[\left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]' = (1-\sigma) \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} \left(e^{\frac{\delta_j}{1-\sigma}} \right)'$$

$$= -\alpha \hat{s}_j \times \left[\sum_{g=0}^G \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma} \right]$$

Plug back to equation (B1),

$$\frac{ds_k}{dp_j} = \frac{\frac{\alpha\sigma}{1-\sigma} \hat{s}_{k|g} \hat{s}_j \times \sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma}}{\left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]} - \hat{s}_k \times \frac{-\alpha \hat{s}_j \times \left[\sum_{g=0}^G \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma} \right]}{\left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]} = \alpha \hat{s}_j \left(\frac{\sigma}{1-\sigma} \hat{s}_{k|g} + \hat{s}_k \right)$$

Case 3: $j \neq k, g(j) \neq g(k)$ (different alternatives in different nests)

$$A = \overbrace{\left(e^{\frac{\delta_k}{1-\sigma}} \right)'}^{=0} \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} + e^{\frac{\delta_k}{1-\sigma}} \overbrace{\left(\left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{-\sigma} \right)'}^{=0} = 0$$

$$B = \left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]' = \left[\left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma} \right]' = -\alpha \hat{s}_j \times \left[\sum_{g=0}^G \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma} \right]$$

Plug back to equation (B1),

$$\frac{ds_k}{dp_j} = -\hat{s}_k \times \frac{-\alpha \hat{s}_j \times \left[\sum_{g=0}^G \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma} \right]}{\left[\sum_{g=0}^G \left(\sum_{k \in g} e^{\frac{\delta_k}{1-\sigma}} \right)^{1-\sigma} \right]} = \alpha \hat{s}_j \hat{s}_k$$

Summary: Finally, I summarize the three cases together.

$$\frac{ds_k}{dp_j} = \begin{cases} -\alpha \hat{s}_j \left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma} \hat{s}_{j|g} + \hat{s}_j \right) & j = k \\ \alpha \hat{s}_j \left(\frac{\sigma}{1-\sigma} \hat{s}_{k|g} + \hat{s}_k \right) & j \neq k, g(j) = g(k) \\ \alpha \hat{s}_j \hat{s}_k & j \neq k, g(j) \neq g(k) \end{cases}$$

Notes: (1) here α is the absolute value of the price coefficient. (2) when calculating the $\hat{\Delta}_{jk} = -\frac{ds_k}{dp_j}$, one also needs to put an extra condition $f_j = f_k$ in each case to index for drug ownership.

B.2 Counterfactual estimation procedures

Two counterfactual situations are evaluated: 1) without a patent pool; 2) with a fully expanded patent pool (once a compound enters, no geographic segmentation within my sample period). The goal is to use estimated demand and supply parameters to simulate counterfactual prices and quantities in each scenario (under different market structure assumptions) and compute changes in consumer and producer surpluses.

I investigate two broad cases in the supply-side market structure: marginal cost pricing and a Bertrand-Nash game. In the first case of marginal cost pricing, I assume marginal cost curves to be flat and independent of quantity. Counterfactuals regarding this case are fairly straightforward as counterfactual prices can be simulated by adjusting the counterfactual values of the MPP variable. In this case, consumers obtain all the social surplus.

In the case of a Bertrand-Nash game, I simulate counterfactual prices, quantities, and marginal costs by optimization in each country-year market. This case is then broken down to three sub-cases: single product oligopoly and multi-product oligopoly. The major difference across the three cases lies in how I define the ownership matrix. In the single product case, only the diagonal elements in the substitution matrix are relevant to a firm's pricing decision. In the multi-product case, I assign ownership based on branded-firm's drug ownership and treat cross-firm cocktails as owned by a separate firm.

Numerical optimization: oligopolistic pricing, with single/multi-product firms

$$\hat{p}_j = \underset{p_j}{\operatorname{argmin}} \left\| \hat{p}_j - \widehat{mc}_j - \underbrace{\widehat{\Delta}_{jk}^{-1} \times \hat{s}_j}_{\text{makeup}_j} \right\|^2 \quad (1)$$

$$\hat{q}_j = \operatorname{Pr}_j(\hat{p}_j) \times M = \hat{s}_j(\hat{p}_j) \times M \quad (2)$$

$$mc_j(\hat{q}_j) = \beta MPP_j^{cf} + X_j \gamma + \eta \hat{q}_j + \omega_j \quad (3)$$

$$\hat{s}_j = \frac{e^{\frac{\delta_j}{1-\sigma}} \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{-\sigma}}{\sum_{g=0}^G \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{-\sigma}}, \text{ where } \delta_j(\hat{p}_j) = X_j \beta + \xi_j - \alpha \hat{p}_j \quad (4)$$

$$\widehat{\Delta}_{jk} = \begin{cases} \alpha \hat{s}_j \left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma} \hat{s}_{j|g} - \hat{s}_j \right) & , j = k \\ -\alpha \hat{s}_j \left(\frac{\sigma}{1-\sigma} \hat{s}_{k|g} + \hat{s}_k \right) & , j \neq k, g_j = g_k, f_j = f_k \\ -\alpha \hat{s}_j \hat{s}_k & , j \neq k, g_j \neq g_k, f_j = f_k \\ 0 & , o. w. (i. e., f_j \neq f_k) \end{cases} \quad (5)$$

$$\hat{s}_{j|g} = \frac{\frac{\delta_j}{e^{1-\sigma}}}{\sum_{j \in g} \frac{\delta_j}{e^{1-\sigma}}}, \text{ where } \delta_j(\hat{p}_j) = X_j\beta + \xi_j - \alpha\hat{p}_j \quad (6)$$

Here, for each market, a profit maximization decision is built within (1) to ensure that the counterfactual new price is generated from the Bertrand-Nash game by minimizing the squared distance between the price and the sum of marginal cost and markup. The latter two are simultaneously updated using (2) and (3) within the fmincon minimization within (1). I impose mild conditions that prices are positive and less than twice the actual non-counterfactual prices to ensure that price search is within a realistic range.

More specifically, the algorithm starts with an initial guess of \hat{p}_j for each country-year market. It then calculates the objective function using \hat{p}_j , relevant demand and supply parameters, and the counterfactual marginal cost. The optimal new prices (from the first-order condition) are obtained using fmincon. The quantity and marginal cost are then updated with the above equation system.

Additional notes on an alternative estimation approach

Alternatively, one can get the quantity equation (1) using a simulation approach (less efficient).

One can obtain equation (1) by simulating demand from the nested logit utility function.

$$u_{ijct} = \frac{X_{jct}\beta - \alpha p_{jct} + \xi_{jct}}{\delta_{jct} = \ln(s_j) - \ln(s_0) - \sigma \ln(s_{j|g})} + \zeta_{ig(j)ct} + (1 - \sigma)\varepsilon_{ijct}$$

Therefore, the utility from counterfactual prices for a given ct can be expressed as below. Where the $\zeta_{ig(j)} + (1 - \sigma)\varepsilon_{ij}$ cannot be simulated with the independent GEV simulator in MATLAB, it shall be simulated using the inverse CDF approach based on the nested logit CDF (Train book, p. 79, equation (4.1)).

$$u_{ij}(\hat{p}_j) = \underbrace{X_j\beta + \xi_j}_{\delta_j - \alpha p_j} + \alpha \hat{p}_j + \zeta_{ig(j)} + (1 - \sigma)\varepsilon_{ij}$$

To simulate the utility for $N_{sim} = 100,000$ consumers across drugs in a given market (country-year), draw $N_{sim} \times J$ nested logit errors from the Generalized Extreme Value (GEV) distribution. Here $j \in \{0, 1, \dots, J\}$ indicates distinct drugs within a market, including the outside option 0. For each simulated consumer i , (1) calculate the $u_{ijct}, \forall j$, (2) find the j that maximizes utility for i , and (3) define the realized choices for person i as $z_j(i) = 1$ if i chooses j .

With the realized choices, one can calculate $\hat{s}_j = \frac{\sum_i^{N_{sim}} z_j(i)}{N_{sim}}$ and $\hat{q}_j = M \times s_j$ for a single market.

Then, repeat the process for each country-year market and save the results into a vector for (2).

B.3 Graphic representation of model fit, and additional results

To test the performance of the optimization and fixed-point algorithm, I use actual data instead of counterfactual values to test whether I can reproduce actual prices and quantities. In addition, I run the algorithm multiple times and confirm the results are not sensitive to the initial guess. The numeric precision is about 99% in all of these placebo tests. I report graphic representation below. In all cases, the placebo prices and quantities fit well with the 45-degree lines.¹

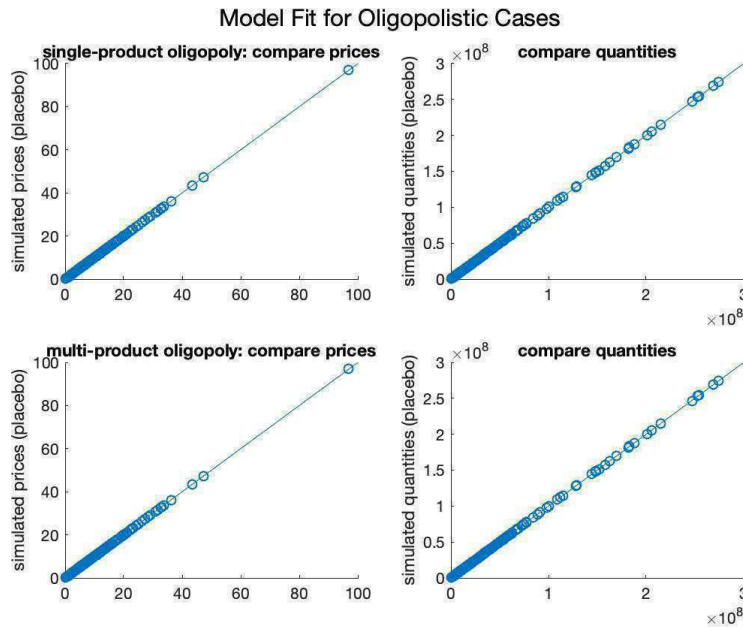


Figure B1: Graphical Demonstration of Model Fits

¹ I also produced corresponding graphs for “actual vs. counterfactual” and they are available upon request.

B.4 Additional results with alternative marginal cost assumptions

In the main analyses, I use flat marginal cost for competitive pricing and increasing marginal cost for oligopolistic pricing. These assumptions are good choices to capture the differences that capacity constraints matter in the two cases I study. This also provides more conservative estimates of the welfare gains from the MPP to consumers and producers.

As an additional exercise and comparison, I also produce the opponent counterfactuals: specifically, for price-taking firms with increasing marginal cost curves and oligopolistic pricing with flat marginal cost curves. From a realistic standpoint, the former case is more interesting as a transition stage of the two cases I discuss in the main text. However, it is worth noting that “price-takers with increasing marginal cost curve” is a typical case where the counterfactual producer surplus can explode. The reason is that when extrapolating the $MC(q=0)$ point that is needed for calculating the producer surplus triangle, the $MC(q)$ curve would generate many negative prices with small quantities that would not be observed in the data. In other words, the in-sample fit can be fine, but $0.5 \cdot \beta(q) \cdot q^2$ provides overly large estimations for producer surplus (grey numbers), despite the still reasonable estimates of relative changes. The “oligopolistic pricing with flat MC” case uses strong assumptions that firms actively optimize and extract profit in LMIC without any capacity constraints, which contradicts reality and thus generates larger divisions in counterfactual cases. I report them below for a comparison.

Table B1: Welfare Re-Estimation: Alternative MC Assumptions

welfare estimates (\$ M)	MC	Oligopolistic Pricing (w/ flat MC)	
	pricing MC(q)	single-prod. firm	multi-prod. firm
$E(\widehat{CS}_0)$	8,112.5	6,409.2	6,246.4
$E(CS)$	8,747.7	8,747.7	8,747.7
$E(\widehat{CS}_1)$	8,836.3	8,821.1	8,811.5
$\Delta\$: CS_0$	635.2	2,338.5	2,501.3
$\Delta\% : CS_0$	7.83%	36.49%	40.04%
$\Delta\$: CS_1$	88.6	73.4	63.8
$\Delta\% : CS_1$	1.01%	0.84%	0.73%
$E(\widehat{PS}_0)$	252\$B	3,071.3	3,315.2
$E(PS)$	266\$B	4,179.5	4,309.6
$E(\widehat{PS}_1)$	267\$B	4,271.8	4,392.1
$\Delta\$: PS_0$	14.1\$B	1,108.2	994.4
$\Delta\% : PS_0$	5.58%	36.08%	30.00%
$\Delta\$: PS_1$	338.4	92.3	82.5
$\Delta\% : PS_1$	0.13%	2.21%	1.91%

When assuming marginal cost increases in quantity, a shift (down) in the supply curve will also affect equilibrium quantity. Regarding this case of competitive pricing with upward sloping marginal cost curve, I simulate counterfactuals using fixed point iterations. In the following part,

I described more details regarding how to use fixed point algorithm or optimization to solve for the equilibrium values in relevant scenarios.

Fixed point iteration: competitive pricing with upward sloping MC curve

$$\hat{q}_j = \text{Pr}_j(\hat{p}_j) \times M = \hat{s}_j(\hat{p}_j) \times M \quad (1)$$

$$\hat{p}_j = mc_j(\hat{q}_j) = \beta MPP_j^{cf} + X_j\gamma + \eta\hat{q}_j + \omega_j \quad (2)$$

To fix ideas, I use $MPP_j^{cf} = 0, \forall j$ (counterfactual: without the MPP) to elaborate below. Note that the MPP only enters through supply side but not via the demand side. The analytical form for \hat{s}_j in equation (1) is as below.²

$$\hat{s}_j = \frac{e^{\frac{\delta_j}{1-\sigma}} \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{-\sigma}}{\sum_{g=0}^G \left(\sum_{j \in g} e^{\frac{\delta_j}{1-\sigma}} \right)^{1-\sigma}}, \text{ where } \delta_j(\hat{p}_j) = X_j\beta + \xi_j - \alpha\hat{p}_j$$

Now, obtaining *counterfactual* equilibrium price and quantity using *fixed point algorithm*:

For each market (country-year), find (\hat{p}_j, \hat{q}_j) s. t. (1) and (2) hold. Start with a guess p_j^0 close to the true value with a random component, e.g., $p_j^0 = p_j(0.95 + 0.1 * \text{uniform}(1))$.

Iteration #1:

$$\begin{aligned} \hat{q}_j^1 &= \hat{s}_j(\hat{p}_j) \times M \\ \hat{p}_j^1 &= \underbrace{X_j\gamma + \omega_j}_{p_j - \eta q_j} + \eta\hat{q}_j^1 \end{aligned}$$

... iteration # $l + 1$:

$$\begin{aligned} \hat{q}_j^{l+1} &= \hat{s}_j(\hat{p}_j) \times M \\ \hat{p}_j^{l+1} &= X_j\gamma + \omega_j + \eta\hat{q}_j^{l+1} \end{aligned}$$

Continue until $\|p_j^{l+1} - p_j^l\| < \varepsilon$

² More details are available from the book by Kenneth Train (2009) “Discrete Choice Methods with Simulation.”