Economics and Human Biology 62 (2026) 101616

Contents lists available at ScienceDirect

Economics and Human Biology

journal homepage: www.elsevier.com/locate/ehb

Antibiotic resistance, drug prices, and entry☆

Miaoqing Jia [a], Ching-to Albert Ma [b,*]

[a] Population Health Sciences, Weill Cornell Medicine, United States of America
[b] Department of Economics, Boston University, United States of America

ARTICLE INFO

JEL classification:
I1
I11
L1

Keywords:
Antibiotic resistance
Drug prices
Pharmaceuticals
Competition
Monopoly
Entry
Cournot

ABSTRACT

Antibiotics lose power to kill microbes through excessive use, commonly known as antibiotic resistance, which is modeled by future cost increase from current use. The first best incorporates future cost externality due to current consumption. Resistance and consumer welfare deviate from first best in various market structures. Low prices under competition lead to high consumption and resistance. Competitive managed-care plans use rationing contracts, which partially internalize resistance cost externality by restricting use. A drug monopolist fully internalizes the externality, but sets a high price. We derive necessary and sufficient conditions for higher consumer surplus under monopoly than competition. Facing a potential entrant, an incumbent may consider entry deterrence and accommodation. To deter entry, the incumbent reduces current sales to lower future cost; this alleviates antibiotic resistance. Entry accommodation gives rise to two countervailing effects. First, incumbent and entrant will share the market. The incumbent has less incentive to internalize the cost effect, so raises production and exacerbates resistance. Second, the incumbent makes more profit when its future cost is low, so reduces production and mitigates resistance.

1. Introduction

Antibiotics are medicines that kill bacteria, protozoa, and fungi. Modern antibiotics have led to extraordinary successes in the fight against infections. Due to overuse and misuse, however, microbes have become resistant to antibiotics.[1] Such antibiotic resistance threats are estimated to cost more than $4.6 billion annually in the United States (Nelson et al., 2021). The increased costs due to antibiotic resistance come from an extended length of stay, additional treatments, and more intensive care required for patients.[2] Based on reports from WHO, pharmaceutical companies are reluctant to invest in antibiotic development due to the high cost of research and development and low return compared to other drugs (Klug et al., 2021).

The economic issue of antibiotic resistance is a classic "tragedy of the commons", wherein self-interested individuals free ride and deplete a shared resource, which, in this case, is antibiotics' power to kill microbes.[3] In this paper, we define *resistance* by the antibiotic becoming more costly in the future because of higher current consumption. The cost-increase definition is equivalent to antibiotics decreasing future quality due to current consumption.

An antibiotic-resistance market failure results from current users not internalizing future cost-increase effects. A two-period model is sufficient to elucidate the economic principles. We study this market failure under (1) perfect competition, (2) monopoly, and (3) Cournot competition between an incumbent and a potential entrant. We also separately study perfect competition and monopoly under conventional pricing and drug-plan contracts.

First, conventional pricing in a competitive market is marginal-cost pricing. Consumption is highest at the lowest price, leading to most

☆ This article is part of a Special issue entitled: 'Healthcare Trends & Future Perspectives' published in Economics and Human Biology.

* Corresponding author.
E-mail addresses: mij4018@med.cornell.edu (M. Jia), ma@bu.edu (C.-t.A. Ma).

[1] https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance, Overview section.

[2] The U.S. Centers for Disease Control in 2021 estimated that more than 35,000 people died because of antimicrobial-resistant infections, with more than 2.8 million infections occurring in the United States. See https://www.cdc.gov/antimicrobial-resistance/stories/partner-estimates.html. Also, the Food and Agriculture Organization of the United Nations claims that antibiotic resistance "would cost the world $412 billion a year in additional healthcare costs and $443 billion per year in lost workforce productivity". See https://www.fao.org/antimicrobial-resistance/news-and-events/news/news-details/en/c/1680317/.

[3] Other examples of failures of this sort are over-fishing, over-grazing, and environmental pollution. We focus only on human use of antibiotics. The evidence on whether animal use of antibiotics contributes to resistance in humans remains mixed; see, e.g., Laxminarayan et al. (2013), and Van Boeckel et al. (2015). Adda (2020) finds that human prescriptions are the main determinant of resistance even controlling for animal production. Agricultural considerations are beyond the scope of this work.

https://doi.org/10.1016/j.ehb.2026.101616
Received 20 October 2025; Received in revised form 22 April 2026; Accepted 21 May 2026
Available online 30 May 2026

serious antibiotic resistance. Alternatively, competition may be in the form of drug plans in which consumers pay a premium and receive medicines under a rationing rule at zero monetary cost. The rationing rule assigns the medicine to consumers according to their health status or valuation. Given the rationing rule, a drug plan sets a premium. Consumers select the (nonlinear) contract that yields the highest expected utility. This is a standard view of health-plan competition.

We characterize the competitive drug plan equilibrium. Drug plans are able to internalize some of the cost-increase (or quality-reduction) consequence of antibiotic use because plans are responsible for costs (when consumers pay nothing ex post). However, consumers have to be convinced to be rationed ex post, so some additional plan benefits have to be offered. But the gain in resistance internalization is first order, so drug-plan competition must yield higher consumer surplus than conventional price competition.

Second, we consider a monopoly market. A single firm produces antibiotics, so the monopolist fully internalizes the adverse effects of antimicrobial resistance. Price and cost must be positively related. Internalization by the monopolist may also be valued by consumers because future costs, hence prices, may not become higher under competition. Nevertheless, the monopolist's concern is profit, not consumer surplus, so this misalignment may benefit or hurt consumers. We give necessary and sufficient conditions for the monopoly allocation yielding higher consumer surplus than the competitive allocation. When a monopolist can use drug plans, it has a two-part tariff scheme to extract all consumer surplus. Antibiotic resistance can be completely internalized, so allocations are first best, but consumers retain no surplus.

Third, we consider imperfect competition. In the first period, the incumbent monopolist chooses to sell a quantity of antibiotics, which then determines its cost in the second period. Then in the second period, the entrant, having a higher quality antibiotic, can enter if its post-entry profit can cover its fixed and variable costs. We model the post-entry competition by a Cournot game with firms having drugs with different qualities.

The strategic issues become more involved. The incumbent can deter entry by lowering its own future cost, which reduces the entrant's variable profit. The entry deterrence strategy works in favor of preventing antibiotic resistance due to lower current consumption. But of course the loss is the lack of a superior drug, as well as reduced consumer surplus.

The incumbent can accommodate entry. Under this strategy, the incumbent chooses its first-period quantity, anticipating the effect on its cost in the second period when Cournot competition against the entrant will happen. There are two effects in play. First, as in standard Cournot interactions, the incumbent's future profit is decreasing in its future costs. If the incumbent manages to lower its future cost, its future profit will increase. This is an incentive to reduce antibiotic resistance. Second, and this is a counter effect, because its market share is reduced by entry, the incumbent has less to internalize. The total effect is the sum of these two, so it is ambiguous if entry accommodation will actually exacerbate or reduce antibiotic resistance.

The antibiotic market structure yields surprising effects. Whereas it is clear that conventional perfect competition may lead to antibiotic resistance in the most serious way, drug plan competition entails remedies. The monopoly situation is more subtle. The monopolist may internalize the antibiotic resistance effect, but its interest is in using the internalization to raise profit, not to enhance consumer surplus. However, consumers do benefit from lower cost, even if the monopolist's pure profit motive may benefit consumers. Finally, in the potential entry scenario, the effect on antibiotic resistance is nuanced. Entry deterrence and accommodation yield different incentives for reducing antibiotic resistance.

Results here help to reframe policy issues. We think that it is best that these issues are discussed after results here have been presented, so we devote a section on financial and non-financial policies, as well as long-term contracts between health regulators and pharmaceutical firms. In any case, our central point is that firms exist in markets, where they interact with other firms.

This paper makes four contributions. First, we develop a two-period model of antibiotic resistance in which current consumption raises future cost, yielding closed-form solutions unavailable in SIS frameworks. Second, we compare consumer welfare across competitive pricing, competitive managed-care plans, and monopoly, deriving necessary and sufficient conditions for monopoly to yield higher consumer surplus than competition. Third, we show that entry deterrence reduces resistance while entry accommodation has ambiguous effects. Fourth, we characterize subscription-type monopoly drug plans that replicate the first best while extracting all consumer surplus.

The paper is organized as follows. Section 2 begins with empirics about various market structures, followed by a literature review. Section 3 presents the demand and cost structures, and derives the first best. In Section 4, we show internalization failures under conventional pricing in a competitive market. We then show that managed-care drug plan competition can alleviate some internalization failure. Section 5 analyzes how a monopolist will internalize cost increase, and compares welfare between competitive and monopoly markets. Section 6 considers potential entry by a firm with a higher-quality drug. Cournot competition with differentiated products formalizes post-entry interactions between firms. Section 7 relates some current policy discussions to our results. Finally, Section 8 draws some conclusions. Where proofs of results are not in the main sections, they are collected in Appendix A. Calculated examples are in Appendix B.

2. Empirics and literature review

Different antibiotic market structures are now illustrated. National Sales Perspectives (NSP) at IQVIA (a leading global provider of data to the life sciences industry) aggregates dollar and unit sales at actual transaction prices from wholesalers, distributors, and pharmaceutical manufacturers, for 90% of the U.S. market.[4] We consider three common antibiotic classes: broad-spectrum penicillins, medium/narrow-spectrum penicillins, and tetracyclines & combinations. Between 2017 and 2022, we find the top four companies by total dollar sales in each of these three classes. Market shares are computed as sales divided by the total sales in each quarter of each year.

In each of the following three figures, the horizontal axis lists quarters over the years whereas the vertical axis measures market shares. Fig. 1 plots broad-spectrum penicillins market shares; by 2022, no firm had more than 1/3 of total sales. The broad-spectrum market is competitive. Fig. 2 plots medium/narrow-spectrum penicillins market shares. Pfizer maintained a very high market share throughout the data period. The medium/narrow-spectrum market is concentrated. Finally, Fig. 3 plots the tetracyclines and combinations market shares. There, market entry and exit are apparent. Almirall S.A. initially held more than 30% market share but exited the market by early 2020, while Paratek Pharmaceuticals entered the market in Q1 2019, increased its share gradually to take the lead in 2020, and eventually reached over 25% by 2022. Finally, we compute Herfindahl Indexes for the three classes, and they are shown in Fig. 4. The NSP data confirm the empirical relevance of the markets we study.[5]

Common-pool resource and Pigouvian approaches. Outterson (2009) and Outterson et al. (2015) frame antibiotic effectiveness as a common-pool resource that is steadily depleted through use. An individual consumer, making a short-run antibiotic use decision, fails to consider the long-run depletion effect of this resource. Outterson et al. (2015) argue that the current rate of innovation is insufficient to meet the growing challenge of antibiotic resistance, a concern echoed by the U.S. Centers for Disease Control and Prevention in 2024.

[4] https://www.iqvia.com/insights/the-iqvia-institute/available-iqvia-data

[5] Other antibiotic examples are available from the authors.

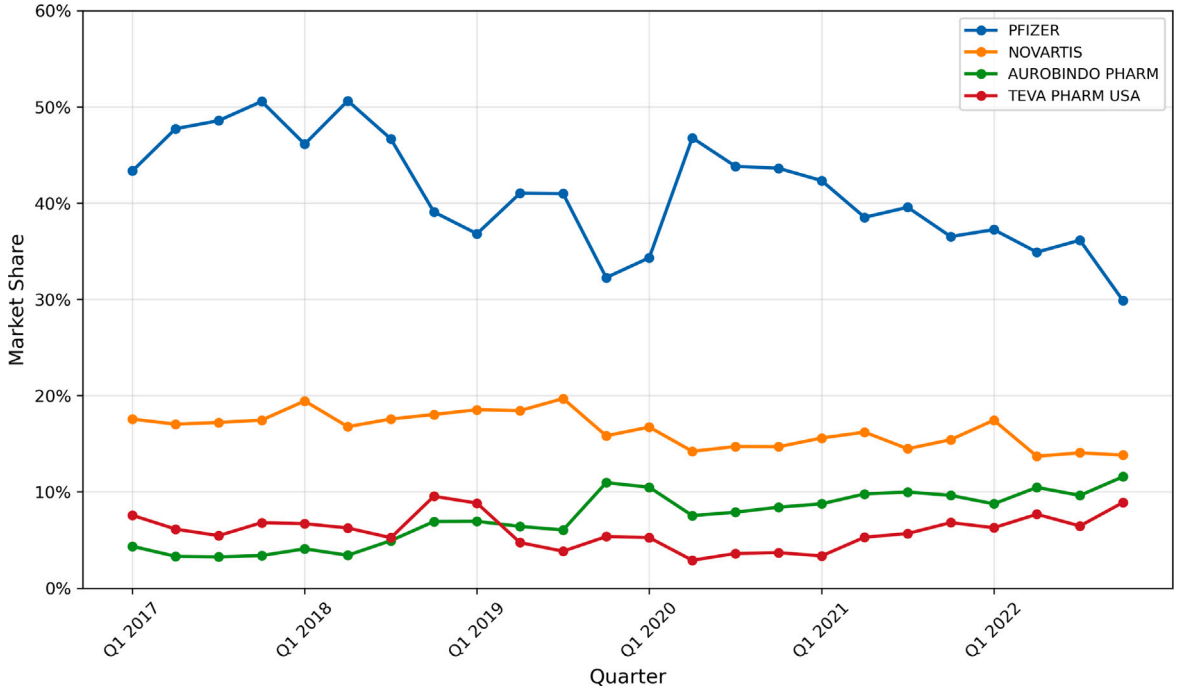

Fig. 1. Broad-spectrum penicillins market shares.

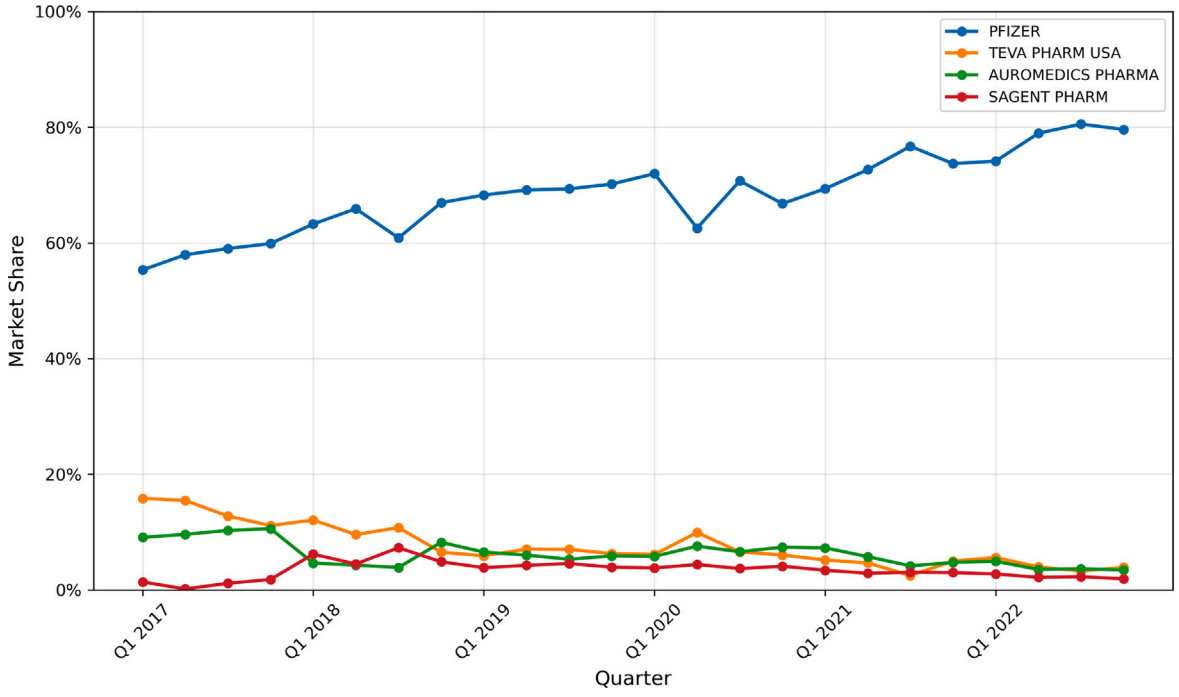

Fig. 2. Medium/narrow-spectrum penicillins market shares.

The tragedy of the commons problem naturally calls for a Pigouvian remedy. Giubilini (2019) advocates for a tax on antibiotic uses because humans should have a moral obligation to protect the common good. While such taxes may create inequities and unequal access to healthcare, Giubilini (2019) argues that the antibiotic emergency may justify such drastic measures. Kuhn et al. (2011) extend this approach by incorporating a life-cycle framework for evaluating healthcare distribution. They highlight the spillover effects of antibiotic use across the life cycle, necessitating a tax-transfer system to align individual behaviors with social objectives. We are unaware of any literature that investigates market effects on resistance.

Epidemiological and Industrial Organization models. Laxminarayan and Brown (2001) use a Susceptible–Infected–Susceptible (SIS) model, adapted from the standard Susceptible–Infected–Recovered (SIR) framework in epidemiology, to describe multiple antibiotic uses. Their model derives conditions for consumers switching between antibiotics based on cost and drug efficacies. Laxminarayan and Weitzman (2002) show that resistance is compounded by uniform treatment strategies

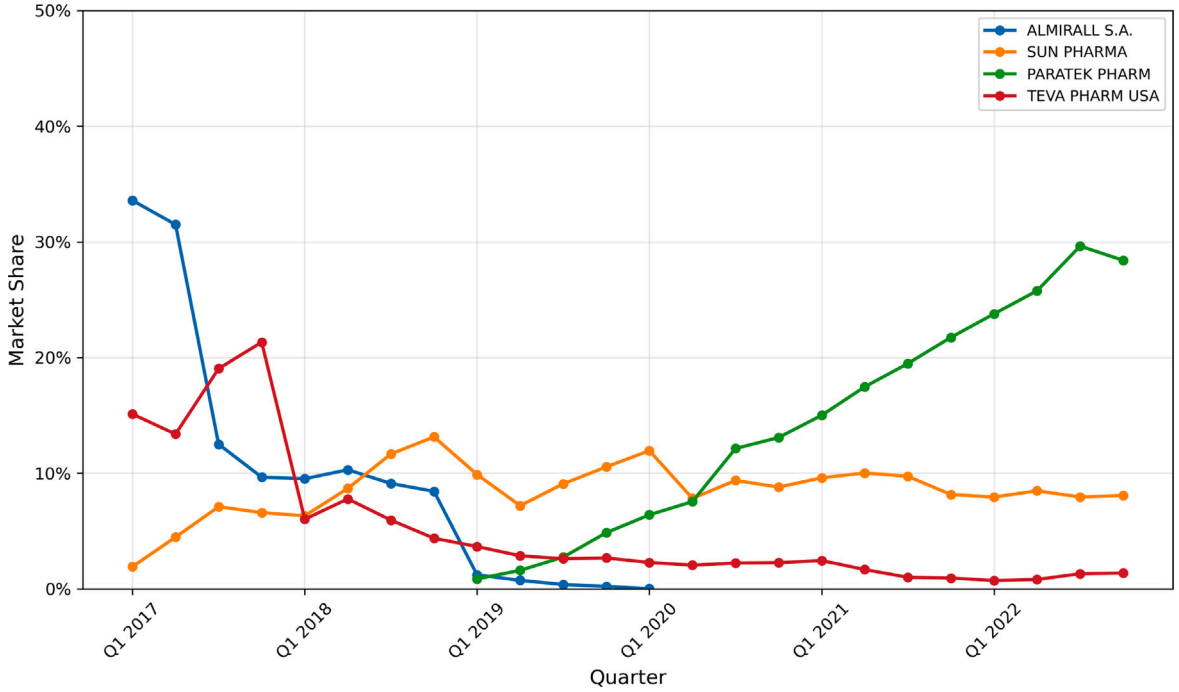

Fig. 3. Tetracyclines and combinations market shares.

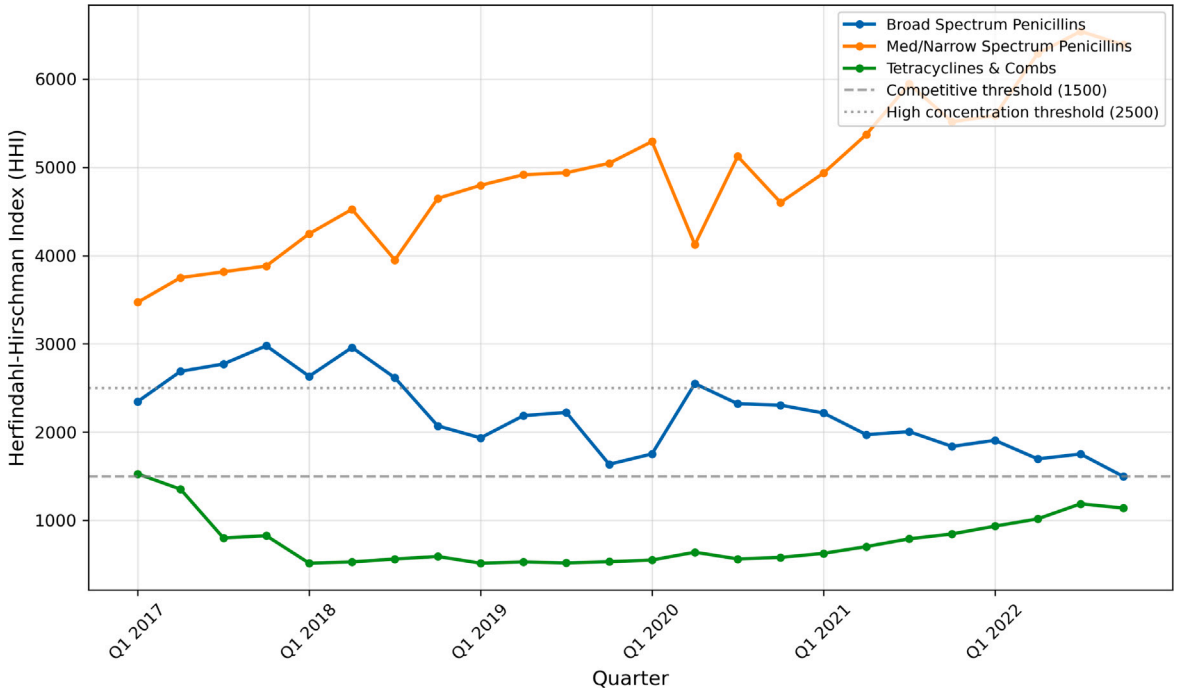

Fig. 4. Herfindahl indexes.

and thus favor a flexible strategy of treatment. Herrmann and Gaudet (2009) and Herrmann and Nkuiya (2017) explore antibiotic dynamic efficacy under open-access conditions, showing how taxes and subsidies can be employed to restore efficiency in antibiotic usage. Herrmann (2010) explores pricing policy under patent protection with an SIS model, whereas Herrmann et al. (2013) study innovation under incentives provided by dynamic tax-subsidy policies. In Albert (2021), the SIS model is combined with economic components to study the evolution of resistance under imperfect competition, with perfect competition and monopoly as special cases. The SIS model involves a set of differential equations but closed-form solutions are unavailable, so simulations and calibrations are often required to draw conclusions. This paper uses demand and cost functions as primitives. The patient-physician interaction in Albert (2021) is simplified as an overall demand, and we model imperfect competition explicitly by Cournot.

Empirical work. Filippini et al. (2006, 2009) and Filippini and Masiero (2012) investigate habit and persistence in antibiotic use and regional variations in consumption patterns. Their studies reveal that past consumption of antibiotics will indeed lead to its future consumption; even in a small country such as Switzerland, regional variations

can be significant, too large to be attributed to price variations. Bokhari et al. (2024) use a nested logit model to demonstrate how taxes can shift demand from broad-spectrum to narrow-spectrum antibiotics, a shift that could help mitigate resistance.

3. Model

3.1. Consumers

There is a unit mass of consumers, indexed by $w \in [0,1]$. Each consumer may benefit from taking a medicine, which has a quality normalized at 1. There is a downward sloping (and twice differentiable) inverse demand function $D : [0,1] \to \mathfrak{R}_+$. A type-$w$ consumer values the medicine at $D(w)$; those with types less than w have higher valuations than $D(w)$. On paying a price p for the medicine, consumer w gets a utility $D(w)-p$; in other words, at price p, the mass of consumers who purchase is w where $D(w) = p$. Antibiotics are often covered by insurance, so any price p can be regarded as the insured price; we abstract from insurance considerations.

3.2. Drug resistance and cost

Microbe drug resistance is studied in a two-period model. Consumers consider using the medicine in each of two periods; consumer benefits are independently and identically drawn from the same distribution in each period. We also assume that there is no discounting.[6] Denote the masses of consumers using medicine in periods 1 and 2 by w_1 and w_2, respectively.

The marginal cost of medicine in period 1 is $c_0 \geq 0$, and its marginal cost in period 2 depends on total consumption in period 1. If w consumers use medicine in period 1, the marginal cost in period 2 becomes $C(w)$, which is an increasing function $C : [0,1] \to \mathfrak{R}_+$. We set $C(0) = c_0$ and $C(1) = \overline{c}$. In other words, if there is no consumption in period 1, the marginal cost remains at c_0, and if all consumers use medicine, the marginal cost in period 2 is at its highest at $\overline{c}$. We further normalize c_0 to 0. This normalization is without loss of generality. If $c_0 > 0$, then consumers with $D(w) < c_0$ would never consume medicine, so we might just redefine the mass of consumers with benefits higher than c_0 as 1. Similarly, we set $D(1) = 0$.

We model drug resistance by a cost increase: when more consumers use the medicine (a higher value of w), its future marginal cost goes up. We do not mean that current consumption raises future production cost. The interpretation is that medicine and resource are required for a treatment. Resistance may mean that more medicine and resource are required for the same treatment, hence the higher cost. The cost increase may also be due to changes in treatment regimen if the antibiotic is part of a set of drugs.

We can model antibiotic resistance as a quality or valuation reduction, and perhaps that is a more natural interpretation of resistance. Suppose that consumption by w_1 consumers in period 1 reduces valuation by $R(w_1)$ in period 2. Let consumer w obtain a benefit $D(w) - R(w_1)$ in period 2. This is a parallel shift downward of the (period 2) demand function $D(w)$ by $R(w_1)$. If we set the cost increase in period 2 from consumption w_1 in period 1 as $C(w_1) = R(w_1)$, we obtain the same model. All results work identically under the quality reduction formalism. In Section 6, a new firm may enter with a higher quality drug, and cost comparative statics are more convenient in Cournot analysis.

The increased-cost or reduced-quality formulations are reduced form representation of resistance dynamics. In Susceptible Infected Susceptible (SIS) models such as Laxminarayan and Brown (2001) and Albert (2021), greater antibiotic use shifts the bacterial population toward resistant strains, raising treatment cost. The function $C(w_1)$ captures this channel: higher period-1 consumption leads to a larger resistant fraction and more costly treatment in the second period. The reduced form permits closed form solutions and comparative statics, which the SIS framework typically does not. The qualitative features that resistance is increasing in consumption and that current use imposes a negative externality on future users are preserved.

3.3. First best

We write down the first-best benchmark. An allocation is defined by the sets of consumers who get the medicine in each period. Obviously, if a type-w consumer gets allocated the medicine, so should consumers with lower types (who have higher valuations than consumer w). We define an allocation by (w_1, w_2), where types $w \leq w_i$ will be allocated the medicine in period i, $i = 1, 2$.

In allocation (w_1, w_2), consumers' total benefit in the two periods is $\int_0^{w_1} D(w)\mathrm{d}w + \int_0^{w_2} D(w)\mathrm{d}w$. The total cost is $\int_0^{w_2} C(w_1)\mathrm{d}w$, where we have used the normalization that medicine cost in period 1 is zero, as well as the relationship between period 1 consumption and period 2 cost.

A first-best allocation is (w_1^*, w_2^*) that maximizes the total benefit less total cost:

$$U(w_1, w_2) \equiv \int_0^{w_1} D(w)\mathrm{d}w + \int_0^{w_2} [D(w) - C(w_1)]\mathrm{d}w. \tag{1}$$

We assume that the social welfare function (1) is concave, so U in (1) has a negative definite second-order derivative matrix.[7] Using standard first-order conditions, we have the following (with the proof omitted):

Proposition 1. *The first-best allocation (w_1^*, w_2^*) is given by*

$$D(w_1^*) - w_2^* C'(w_1^*) = 0 \tag{2}$$

$$D(w_2^*) - C(w_1^*) = 0. \tag{3}$$

Eq. (3) in Proposition 1 states that in period 2 consumers with benefits higher than cost should use the medicine. Eq. (2) in the Proposition shows an intertemporal tradeoff. Lowering the threshold w_1 reduces benefit by $D(w_1)$ in period 1 but this consumption reduction lowers period-two cost by $w_2 C'(w_1)$. At the first best, these two effects must be equal. If there was no cost increase (so that $C' \equiv 0$), all consumers would get the medicine in both periods. Proposition 1 lays out the rationale for medicine use restraint to avoid higher costs in the future.

4. Competition and internalization failure

Individual consumers do not internalize the cost increase because a single consumer cannot affect aggregate consumption. In a competitive market, the medicine's price is the marginal cost $c_0 = 0$. At this price, all consumers will purchase medicine in period 1. The (marginal) cost of medicine in period 2 becomes the highest $C(1) = \overline{c}$, and consumers with types below w_2, where $D(w_2) = \overline{c}$ will purchase medicine. Consumer surplus becomes

$$\int_0^1 D(x)\mathrm{d}x + \int_0^{w_2} [D(x) - C(1)]\mathrm{d}x < \int_0^{w_1^*} D(x)\mathrm{d}x + \int_0^{w_2^*} [D(x) - C(w_1^*)]\mathrm{d}x. \tag{4}$$

[6] Introducing a discount factor $\delta \in (0,1)$ on period 2 payoffs would scale the future cost term by δ in both the social welfare and firms' profit. The first best would become $D(w_1^*) = \delta w_2^* C'(w_1^*)$ and $D(w_2^*) = C(w_1^*)$. Comparative results across market structures are qualitatively unchanged. We set $\delta = 1$ for uncluttered notation.

[7] Straightforward calculation yields: $\partial^2 U/\partial w_1^2 = D'(w_1) - w_2 C''(w_1)$, $\partial^2 U/\partial w_2^2 = D'(w_2)$, $\partial^2 U/\partial w_1 \partial w_2 = -C'(w_1)$. Hence we assume that $D'(w_1) - w_2 C''(w_1) < 0$, and $[D'(w_1) - w_2 C''(w_1)]D'(w_2) - C'(w_1)^2 > 0$.

The inequality follows because the first best allocation (w_1^*, w_2^*) gives higher social surplus than any other feasible allocation. In other words, allocation $(1, \underline{w}_2)$ where $D(\underline{w}_2) = C(1) \equiv \overline{c}$, is feasible in (1) but only the first-best allocation (w_1^*, w_2^*) yields the highest surplus. In a competitive market, prices are always at marginal costs, so the first-period price is below the social cost of consumption. We summarize the "tragedy of the commons" result by the following (with proof omitted):

Proposition 2. *In a competitive market where medicine costs are always equal to marginal costs, compared to the first best, consumers use medicine excessively in period 1, which drives up cost in period 2. The total social surplus is below first best.*

Proposition 2 demonstrates a market failure due to competitive firms unable to internalize the resistance externality. Low prices are ineffective for internalization. Can an alternative form of competition perform better?

4.1. Drug plans and internalization

A managed-care organization designs a drug plan, a contract consisting of a premium and provision thresholds (π, w_1, w_2). In exchange for a premium π, paid to the plan before consumers learn their types, ex post a consumer of type w is given medicine, at no cost, in periods 1 and 2 if w is less than w_i in period i, $i = 1, 2$. Consumers' type information is assumed available to the drug plan for contract implementation. Given quantity w_i available in period i, those consumers with valuations or benefits higher than $D(w_i)$ will be allocated the medicine. This interpretation of managed care as delegation and efficient rationing is standard. In recent theoretical literature, Frank et al. (2000), Glazer and McGuire (2000, 2006) explicitly use rationing allocations to characterize health plans.

We continue to assume perfect competition. There are many drug plans. There are also many non-managed-care insurance plans that offer prescriptions at marginal cost. Consumers cannot enroll in multiple insurance plans. A consumer either enrolls in a drug plan and follows the plan rationing rule, or enrolls in a non-managed-care plan and obtains prescriptions to buy medicine at marginal cost.

A drug plan allows consumers to gain some benefits (besides the zero out-of-pocket drug expenses). These are network organization features such as health and treatment advantages. We denote this benefit by $B > 0$; the plan benefit is needed to motivate consumers to surrender their decisions to a managed-care organization. Those consumers who have not enrolled make independent choices about medicine at cost but do not get the plan benefit B. We assume that B is exogenous, but will show how its magnitude will affect allocations. The parameter B can represent either the value of ancillary services that managed-care plans provide (care coordination, referral networks) or a policy lever: subsidies for public enrollment or stewardship programs raise B. Higher values of B lead to greater enrollment and more internalization. The case $B \geq \overline{B}$ corresponds to benefits sufficiently large that all consumers enroll, achieving the first best.

We consider a representative drug plan; the analysis is the same if there are multiple plans and consumers randomly choose among identical plans. Suppose a drug plan sets premium and periods 1 and 2 thresholds (π, w_1, w_2). Let s, $0 \leq s \leq 1$, denote the mass of consumers who have enrolled in the plan; so $1 - s$ have not and these consumers always choose to use medicine in period 1. The total consumption in period 1, given the plan contract, is $1-s+sw_1$ consisting of independent consumers' and w_1 of enrollees' uses. The cost of medicine in period 2 becomes $C(1 - s + sw_1)$.

Given the mass of enrolled consumers s, a consumer's expected utility from enrolling in the plan is

$$\int_0^{w_1} D(x)\mathrm{d}x + \int_0^{w_2} D(x)\mathrm{d}x - \pi + B. \tag{5}$$

According to the thresholds, those consumers with types below w_i will be given medicine in period i, $i = 1, 2$. The valuation of consumer w is $D(w)$, so the two integrals are the expected benefit from the plan's rationing rule. The consumer pays the premium π but enjoys the plan benefit B.

An independent consumer (one without drug-plan enrollment) will always purchase the drug in period 1. Given the mass of enrolled consumers s and the rationing rule w_1 in period 1, the medicine's marginal cost will become $C(1-s+sw_1)$ in period 2. Those independent consumers with benefit $D(w) > C(1-s+sw_1)$ will purchase the medicine in period 2. A consumer's expected utility from staying unenrolled in a drug plan is

$$\int_0^1 D(x)\mathrm{d}x + \int_0^{\widehat{w}} [D(x) - C(1 - s + sw_1)]\ \mathrm{d}x, \tag{6}$$

where $D(\widehat{w}) = C(1 - s + sw_1)$. The first integral is the benefit in period 1 from consuming medicine at cost 0; the second integral is the benefit of the consumer optimally choosing to purchase at cost when benefit is higher than cost.

The drug plan's profit per enrollee is

$$\pi - w_2 C(1 - s + sw_1). \tag{7}$$

It receives a premium π per enrollee; with total medicine consumption $1 - s + sw_1$ in period 1, in period 2 the cost for providing the medicine to consumers with types below w_2 is $w_2 C(1 - s + sw_1)$.

We have focused on one representative drug plan. This might have created the impression that there was no competition. The single-plan focus is for expositional convenience. The fraction s of drug-plan enrollees is interpreted as the aggregate enrollment fraction. One particular drug plan may only enroll a fraction of s but the cost in period 2 remains the same, namely $C(1 - s + sw_1)$ when each plan sets the threshold w_1 in period 1. Now we define an equilibrium.

Definition 1. A Drug Plan Competitive Equilibrium is a drug plan contract (π^e, w_1^e, w_2^e) and an enrollment share s^e, $0 \leq s^e \leq 1$, such that
(1) given $s = s^e$, (π^e, w_1^e, w_2^e) maximizes the enrollees' expected utility in (5), subject to the plan making a zero profit in (7);
(2) consumers maximize their utilities choosing between staying independent and enrolling in the drug plan, and a fraction s^e optimally choose to enroll.

Given the enrollment share, a competitive drug plan must choose a premium and rationing thresholds to maximize ex ante consumer surplus, and make a zero profit. This is the first requirement in the definition. Ex ante consumers are identical, so in order for a fraction of consumers to enroll, all consumers must be indifferent between enrolling in the plan and staying independent. This is the implication of the second requirement. We now characterize drug plan competitive equilibria.

Consider a given share s of consumers enrolling in a drug plan. In a Drug Plan Competitive Equilibrium, the plan chooses w_1 and w_2 to maximize consumer's surplus and charge a zero-profit premium. Using (5) and setting (7) to zero to obtain the premium $\pi = w_2 C(1 - s + sw_1)$, for any given s, we have the optimal drug-plan thresholds as follows:

$$(\widetilde{w}_1(s), \widetilde{w}_2(s)) = \arg\max_{w_1, w_2} \left\{ \int_0^{w_1} D(x)\mathrm{d}x + \int_0^{w_2} D(x)\mathrm{d}x - w_2 C(1 - s + sw_1) \right\}, \tag{8}$$

where we have emphasized that the optimal thresholds depend on s. It is obvious that at $s = 0$ (zero drug plan enrollment), $(\widetilde{w}_1, \widetilde{w}_2) = (1, \underline{w}_2)$, the competitive equilibrium allocation, and that at $s = 1$ (full drug plan enrollment), $(\widetilde{w}_1, \widetilde{w}_2) = (w_1^*, w_2^*)$, the first best.

As s increases from 0, a drug plan internalizes more of the cost increase in period 2 due to period 1 consumption. Hence, it optimally reduces consumption of the medicine in period 1. Because of the more favorable cost in period 2, the drug plan also raises consumption in

period 2. We provide mild conditions to confirm these monotone effects in the following Lemma (proof in Appendix A).

Lemma 1. *Suppose that C is linear or the values of its second-order derivative are small enough. As s increases from 0 towards 1, $\widetilde{w}_1(s)$ monotone decreases from 1 to w_1^*, and $\widetilde{w}_2(s)$ monotone increases from $\underline{w}_2$ to w_2^*.*

Next we define a cutoff value on B. Let $\overline{B} \equiv \int_{w_1^*}^1 D(x)\mathrm{d}x$. The integral $\int_{w_1^*}^1 D(x)\mathrm{d}x$ is a consumer's surplus loss forgoing consumption at $0 < D(w) < D(w_1^*)$. As we will see, any $B > \overline{B}$ will induce all consumers to join the drug plan. The proof of the following Proposition is in Appendix A.

Proposition 3. *At each B, $0 \le B < \overline{B}$, there is a unique Drug Plan Competitive Equilibrium $(\pi^e, w_1^e, w_2^e; s^e)$ given by the unique solution of the following*

$$\int_{w_1^e}^1 D(x)dx = B \tag{9}$$

$$D(w_1^e) - w_2^e s^e C'(1 - s^e + s^e w_1^e) = 0 \tag{10}$$

$$D(w_2^e) - C(1 - s^e + s^e w_1^e) = 0 \tag{11}$$

$$w_2^e C(1 - s^e + s^e w_1^e) = \pi^e. \tag{12}$$

The equilibrium threshold in period 1 is strictly less than 1 but above the first best: $w_1^* < w_1^e < 1$. At $B \ge \overline{B}$, the unique Drug Plan Competitive Equilibrium is first best, with all consumers enrolling.

In period 2, there is no concern for further cost increases, so independent consumers and managed-care enrollees have the same consumption allocation.[8] Upon joining a drug plan, a consumer experiences a loss of expected surplus in period 1: the consumer will not be allocated medicine if the valuation is less than $D(w_1^e)$. Hence, the drug plan benefit must compensate for the loss, and this is the requirement in (9).

Next, drug plans maximize consumer surplus to compete. For a given share of consumers joining a drug plan, thresholds w_1^e and w_2^e maximize enrollees' surplus, so must satisfy the first-order conditions in (10) and (11). However, w_1^e is already pinned down by (9). Hence, the drug plan enrollee share s^e in (10) and (11) must adjust accordingly. The premium must make the drug plan earn zero profit, which is (12). Finally, (9) and (11) say that consumers are ex ante indifferent between enrolling and staying independent, so a share s^e joining is (part of) an equilibrium.

A drug plan bears its enrollees' costs. When a drug plan has enrolled a fraction s of consumers, it internalizes the cost increase in period 2 due to members' period 1 consumption. However, it cannot stop nonmembers from purchasing in period 1. Suppose that the plan reduces members' period 1 consumption at w_1 by 1 unit, the cost increase in period 2 is still $C'(1 - s + sw_1) > C'(w_1)$. Consumers accept drug-plan rationing because of the plan benefit B. When B is small, the equilibrium rationing in period 1 cannot be very strong, so w_1^e is still close to 1, maximum consumption. When the benefit B increases, more consumers are willing to join, and the drug plan can internalize more of the cost consequence of excessive consumption in period 1. As B increases towards $\overline{B}$, the internalization becomes stronger and approaches the first best.

A drug plan with quantity rationing offers a way to alleviate the competitive market's failure to internalize the cost increase. It does have a caveat: there is, indeed, no free lunch. Consumers are "bribed" to give up their control and use of the competitive prices. The importance of Proposition 3 is that it explains the channel through which the bribe B leads to rationing to reduce period 1 consumption—even when drug plans operate in a competitive environment.

A drug plan can be likened to a notion of antibiotic stewardship, a guideline offered by public health policy makers for practitioners. The CDC has developed the Outpatient Antibiotic Stewardship program, which lays out best prescription practices to guide healthcare providers.[9] Our results here can be used as a framework to establish the stewardship guidelines, but we also warn that the implementation of stewardship may not be costless.

5. Monopoly and internalization

Monopoly is in sharp contrast with competition: a single pharmaceutical firm sells to consumers. Let the monopolist set medicine prices at p_1 and p_2 in periods 1 and 2, respectively. These prices are set up at the beginning of period 1. Sequentially setting p_i at the beginning of period i, $i = 1, 2$, turns out to be exactly the same. With complete information, the monopolist can anticipate demand responses, so the timing of the monopolist's pricing decisions is inconsequential.

Given (p_1, p_2), let $D(w_1) = p_1$ and $D(w_2) = p_2$; consumer w_i is indifferent between buying and forgoing consumption in period i, $i = 1, 2$. The quantity sold in period i is w_i. The marginal cost in period 2 is $C(w_1)$. The monopolist's profit is

$$\pi(w_1, w_2) \equiv w_1 D(w_1) + w_2[D(w_2) - C(w_1)]. \tag{13}$$

The monopolist recognizes that the sales amount in period 1, w_1, affects its cost in period 2, which becomes $C(w_1)$. The monopolist does internalize the cost increase, but exploits its market power.

We assume that the profit function in (13) is concave. The first-order derivatives of (13) with respect to w_1 and w_2 are:

$$\frac{\partial \pi}{\partial w_1} = w_1 D'(w_1) + D(w_1) - w_2 C'(w_1) \quad \text{and} \quad \frac{\partial \pi}{\partial w_2} = w_2 D'(w_2) + D(w_2) - C(w_1). \tag{14}$$

The first two terms in each of the partial derivatives are marginal revenues in the two periods, each being less than the demands, $D(w_1)$ and $D(w_2)$. The term $w_2 C'(w_1)$ is the incremental cost for serving w_2 consumers in period 2 when w_1 consumers are served in period 1, so that is the marginal cost for selling to consumers in period 1. The marginal cost in period 2 is $C(w_1)$. Setting the first-order derivatives to zero yields the necessary and sufficient conditions for maximum profit.

In period 2, the optimal price $D(w_2^M)$ (or quantity w_2^M) is characterized by the familiar marginal-revenue-marginal-cost condition:

$$\frac{D(w_2^M) - C(w_1^M)}{D(w_2^M)} = -\frac{w_2^M D'(w_2^M)}{D(w_2^M)}, \tag{15}$$

which says that the price-cost margin is inversely proportional to the demand elasticity.[10] This is the usual under-production result in a monopoly.

In period 1, the optimal price $D(w_1^M)$ is given by $D(w_1^M) + w_1^M D'(w_1^M) = w_2^M C'(w_1^M)$, which can be rewritten as

$$\frac{D(w_1^M) - w_2^M C'(w_1^M)}{D(w_1^M)} = -\frac{w_1^M D'(w_1^M)}{D(w_1^M)}, \tag{16}$$

[8] Enrollees do not pay ex post, but independent consumers do. Managed-care plans, however, are able to enforce their rationing rule. In each case, those consumers with valuations higher than the marginal cost will be allocated the medicine.

[9] https://www.cdc.gov/antibiotic-use/media/pdfs/Core-Elements-Outpatient-508.pdf

[10] Quantity is w_2, so the proportional change in quantity is simply $\frac{1}{w_2}$; price is $D(w_2)$, so the proportional change in price is $\frac{D'(w_2)}{D(w_2)}$.

which also has an inverse-elasticity interpretation, except that the marginal cost pertains to the incremental cost in period 2 due to period 1 production. We summarize the above as follows (with its proof having been laid out above):

Proposition 4. *Profit-maximizing monopoly prices,* $(D(w_1^M), D(w_2^M))$*, are in* (15) *and* (16)*.*

Obviously, the consumer surplus from the monopolist's profit-maximizing prices will be below the first best. However, the comparison of consumer surpluses between monopoly and perfect competition is ambiguous in general. Under competition, the cost in period 2 is at its maximum due to free riding. Even when consumers purchase at cost, their surplus may be small. Under monopoly, the usual price-cost margins will mean lower consumptions in both periods. The lower consumption in period 1 means a lower cost in period 2, which also may benefit consumers.

5.1. Consumer surplus comparison between monopoly and competition

In a competitive market, prices are at marginal cost, but the resistance externality is not internalized, and cost in period 2 is high. A monopolist internalizes the resistance externality, so cost in period 2 is low, but the monopolist exploits the lower cost to maximize profits. The next proposition compares consumer welfare across the two markets.

Proposition 5. *(i) Suppose that period 1 consumer surplus under competition is more than two times the consumer surplus in period 1 under monopoly if the monopolist chooses the period 1 profit-maximizing price ignoring any cost increase in period 2; that is*

$$\int_0^1 D(x)dx \geq 2\int_0^{\widetilde{w}} [D(x) - D(\widetilde{w})]dx, \tag{17}$$

where $\widetilde{w} = \arg\max_w wD(w)$*. Consumer surplus under competition is at least as high as monopoly.*

(ii) Conversely, if

$$\int_0^1 D(x)dx < 2\int_0^{\widetilde{w}} [D(x) - D(\widetilde{w})]dx, \tag{18}$$

there exist cost functions C *such that consumer surplus under competition is lower than under monopoly.*

Under competition, period 1 consumer surplus, $\int_0^1 D(x)\mathrm{d}x$, is a lower bound of total consumer surplus for the two periods. The worst that can happen is the complete lack of consumption in period 2. This will be the case when, after all the consumers purchase in period 1, the cost in period 2, $C(1)$, becomes higher than the maximum consumer valuation. Under competition, consumers must be able to do better than the left-hand side of the condition in (17).

The monopolist will internalize the cost increase, and the profit function in (13), rewritten here, is $w_1 D(w_1) + w_2[D(w_2) - C(w_1)]$. Generally the monopolist will price more than $D(\widetilde{w})$, where $\widetilde{w} = \arg\max_w wD(w)$, so period 1 consumer surplus cannot exceed $\int_0^{\widetilde{w}}[D(x)-D(\widetilde{w})]\mathrm{d}x$. Now in period 2, the cost cannot decrease (and may actually rise), so again the monopolist must not sell more than $\widetilde{w}$. For periods 1 and 2 together, consumer surplus cannot exceed $2\int_0^{\widetilde{w}}[D(x)-D(\widetilde{w})]\mathrm{d}x$ under monopoly. This part of Proposition 5 only refers to the demand function D, so it is valid for all cost functions.

For the second part of Proposition 5, we present cost functions for which monopoly may improve welfare. Consider a class of (increasing and convex) cost functions C indexed by $w^\dagger$, defined as follows

$$\left\{\begin{array}{lll} \text{at } 0 \leq w < w^\dagger, & C'(w) < \varepsilon & \text{where } \varepsilon \text{ is close to } 0 \\ \text{at } w^\dagger < w < 1, & C'(w) > \delta & \text{where } \delta \text{ is very large.} \end{array}\right\}$$

The cost function C has arbitrarily small derivatives for w smaller than $w^\dagger$, but very large derivatives once w exceeds $w^\dagger$. Let $w^\dagger$ be close to 1. The graph of this cost function example is (mirror-imaged) L-shaped.

Under competition, all consumers buy in period 1, so $C(1)$ becomes very high. Few, if any, consumers purchase in period 2. Therefore, the consumer surplus lower bound under competition $\int_0^1 D(x)\mathrm{d}x$ is (approximately) the actual consumer surplus over the two periods.

When $C'(w)$ is very small for $w < w^\dagger$ (which is close to 1), the monopolist will choose a period 1 quantity close to $\widetilde{w}$ to avoid raising its own cost. Because the monopolist's cost only increases slightly, the period 2 quantity is also close to $\widetilde{w}$. Hence, consumer surplus will be approximately $2\int_0^{\widetilde{w}}[D(x) - D(\widetilde{w})]\mathrm{d}x$. According to (18), the consumer surplus in competition is less than under monopoly. Clearly, there are many such examples with higher monopoly welfare than competition.

5.2. Monopoly drug plan

Earlier we have considered competitive drug plans. However, full efficiency cannot be maintained unless such plans confer high benefits aside from drug coverage. We now consider drug plans offered by a monopolist. Such a drug plan contract is defined by a premium and two thresholds: (π, w_1, w_2). In exchange for a premium π, consumers receive the medicine in period i if the type is less than w_i in period i, $i = 1, 2$. The price π is paid upfront, before consumers learn their types. Consumers do not have to pay any more expenses, but they are rationed according to the contract.[11]

Under the contract (π, w_1, w_2), a consumer's expected utility is $\int_0^{w_1} D(x)\mathrm{d}x + \int_0^{w_2} D(x)\mathrm{d}x - \pi$. The monopolist's profit is $\pi - w_2 C(w_1)$. Let the consumer's utility of refusing the contract be 0. Then given w_1 and w_2, the highest price Π that will be accepted is

$$\Pi = \int_0^{w_1} D(x)\mathrm{d}x + \int_0^{w_2} D(x)\mathrm{d}x.$$

At this price the monopolist's profit becomes

$$\int_0^{w_1} D(x)\mathrm{d}x + \int_0^{w_2} D(x)\mathrm{d}x - w_2 C(w_1),$$

which is social welfare in (1). The monopolist, internalizing cost increase and extracting all consumer surplus, chooses the first best (w_1^*, w_2^*) to maximize profit. The next result is obvious and its proof is omitted.

Proposition 6. *If the monopolist offers a drug plan, it chooses the first best, and sets a premium to extract all consumer surplus.*

A drug plan is the same as a two-part tariff. For a price Π, a monopolist sells consumers the rights to purchase medicine at price $D(w_1)$ in the first period and price $D(w_2)$ in the second period. The price Π is collected before consumers learn their types. Under this contract $(\Pi, D(w_1), D(w_2))$, the consumer's expected utility is

$$\int_0^{w_1} [D(x) - D(w_1)]\mathrm{d}x + \int_0^{w_2} [D(x) - D(w_2)]\mathrm{d}x - \Pi.$$

From contract $(\Pi, D(w_1), D(w_2))$, the monopolist's profit is

$$\Pi + w_1 D(w_1) + w_2[D(w_2) - C(w_1)].$$

The monopolist internalizes resistance effect, extracts all surplus, and implements the first best.

6. Entry in period 2: Cournot competition with differentiated products

Next, we consider a more complex market structure. We let there be a potential entrant. Firm 1 is the incumbent monopolist in period 1;

[11] There is no need to include a drug plan benefit here because consumers have no other market options.

Firm 2, the potential entrant in period 2. In period 1, Firm 1 produces Drug 1, at quality 1. Firm 1's cost structure remains the same; if it sells w_1 units of Drug 1 in period 1, its marginal cost becomes $C(w_1)$ in period 2. For convenience, we now denote Firm 1's period 2 cost by $c_1 = C(w_1)$, due to its sales of w_1 in period 1.

The potential entrant Firm 2 produces Drug 2, a superior medicine, say at quality $q > 1$, at constant marginal cost $c_2 > 0$. The entrant has to incur a fixed cost, $FC > 0$, to enter the market in period 2. Consumer w's valuation of Drug 2 is $qD(w)$. We assume that if the medicines are sold at cost, high-valuation consumers prefer Drug 2 and low-valuation consumers prefer Drug 1. That is, there is some type $w^\dagger$, $0 < w^\dagger < 1$, such that $D(w^\dagger) = qD(w^\dagger) - c_2$, so consumers with types lower than $w^\dagger$ prefer Drug 2, whereas those higher prefer Drug 1.

We further assume that Drug 2's quality q is not affected by resistance. A new antibiotic may belong to a different chemical class and is not subject to resistance built up against the incumbent's drug. If both drugs shared the same resistance mechanism, the entrant's cost would also depend on w_1. Independent resistance is a neutral benchmark.

If Firm 2 has entered in period 2, the two firms compete in the fashion of Cournot by choosing quantities. Suppose in period 2, Firm 1 produces w_ℓ units of the lower quality Drug 1, and Firm 2 produces w_h units of the higher quality Drug 2. The Cournot construct postulates that there are two prices, p_ℓ for Drug 1 and p_h for Drug 2, that will clear the market. When consumers are free to choose which product to purchase, the demand for Drug 1 will be w_ℓ at price p_ℓ and the demand for Drug 2 will be w_h at price p_h.

The prices that clear the market satisfy the following:

$$p_\ell = D(w_h + w_\ell) \quad \text{and} \quad D(w_h)q - p_h = D(w_h) - p_\ell. \tag{19}$$

According to (19) consumer w_h is indifferent between buying Drug 2 at quality q and price p_h and buying Drug 1 at quality 1 and price p_ℓ, so consumers with types below w_h prefer to purchase Drug 2. The consumer with type $w_\ell + w_h$ is indifferent between buying Drug 1 and nothing; consumers with types between w_h and $w_\ell + w_h$ prefer to purchase Drug 1. Prices in (19) clear the market.

After substitution, we write the market-clearing prices as

$$p_\ell = D(w_h + w_\ell) \quad \text{and} \quad p_h = D(w_h)(q-1) + D(w_h + w_\ell).$$

In period 2, variable profits for Firm 1 and Firm 2 are, respectively, $[p_\ell - c_1]w_\ell$ and $[p_h - c_2]w_h$, where again $c_1 = C(w_1)$ is Firm 1's cost in period 2 after it has produced w_1 in period 1. Taking into account Firm 2's fixed cost, we write profits as

$$\pi_1(w_\ell, w_h) = w_\ell[D(w_h + w_\ell) - c_1] \tag{20}$$

$$\pi_2(w_\ell, w_h) = w_h[D(w_h)(q-1) + D(w_h + w_\ell) - c_2] - FC. \tag{21}$$

Definition 2 (*Cournot Equilibrium*). A Cournot equilibrium is (w_ℓ^e, w_h^e) that are mutual best responses:

$$w_\ell^e = \arg\max_{w_\ell} w_\ell[D(w_h^e + w_\ell) - c_1] \tag{22}$$

$$w_h^e = \arg\max_{w_h} w_h[D(w_h)(q-1) + D(w_h + w_\ell^e) - c_2] - FC \tag{23}$$

provided that Firm 2's equilibrium profit is positive. If Firm 2's Cournot equilibrium profit is negative, it does not enter and Firm 1 continues to be a monopolist in period 2.

We state some comparative statics of firms' quantity optimizations: as Firm 1's cost increases, its profit falls, and its profit-maximizing quantity also decreases; Firm 2's profit falls as Firm 1's quantity increases. The proof of the following lemma is in Appendix A.

Lemma 2. *Given Firm 2's quantity w_h, Firm 1's maximum profit,* $\max_{w_\ell} w_\ell[D(w_h + w_\ell) - c_1]$, *and profit-maximizing quantity,* $\arg\max_{w_\ell} w_\ell[D(w_h + w_\ell) - c_1]$, *are decreasing in cost c_1.*

Firm 2's maximum profit, $\max_{w_h} \{w_h[D(w_h)(q-1) + D(w_h + w_\ell) - c_2] - FC\}$, *is decreasing in Firm 1's quantity w_ℓ.*

Lemma 2 states some natural results: maximum profit and profit-maximizing quantities are decreasing in its own cost. Firm 1's lower cost also harms Firm 2: when its cost falls, Firm 1 produces more, which reduces Firm 2's residual demand.

In period 2, Firm 1's cost is $c_1 = C(w_1)$ due to w_1 units of Drug 1 production in period 1. To emphasize the subgame defined by $c_1 = C(w_1)$, we write explicitly the Cournot equilibrium as $(w_\ell^e(c_1), w_h^e(c_1))$ in (22) and (23).

The Cournot equilibrium quantities (w_ℓ^e, w_h^e) satisfy the respective first-order conditions of the maximization of (22) with respect to w_ℓ, and the maximization of (23) with respect to w_h:

$$D(w_h^e + w_\ell^e) - c_1 + w_\ell^e D'(w_h^e + w_\ell^e) = 0 \tag{24}$$

$$D(w_h^e)(q-1) + D(w_h^e + w_\ell^e) - c_2 + w_h^e[D'(w_h^e)(q-1) + D'(w_h^e + w_\ell^e)] = 0. \tag{25}$$

We assume that these reaction functions uniquely determine the Cournot equilibrium. Then we write the Cournot equilibrium profits as

$$\pi_1^e(c_1) \equiv \pi_1(w_\ell^e(c_1), w_h^e(c_1)) = w_\ell^e(c_1)[D(w_h^e(c_1) + w_\ell^e(c_1)) - c_1] \tag{26}$$

$$\pi_2^e(c_1) \equiv \pi_2(w_\ell^e(c_1), w_h^e(c_1)) = w_h^e(c_1)[D(w_h^e(c_1))(q-1) + D(w_h^e(c_1) + w_\ell^e(c_1)) - c_2] - FC. \tag{27}$$

Firm 2's equilibrium profit is its maximized profit from choosing w_h optimally against Firm 1's quantity $w_\ell^e(c_1)$. Hence, we apply the envelope theorem to obtain

$$\frac{\mathrm{d}\pi_2^e(c_1)}{\mathrm{d}c_1} = \frac{\partial \pi_2(w_\ell^e(c_1), w_h^e(c_1))}{\partial w_\ell} \frac{\mathrm{d}w_\ell^e(c_1)}{\mathrm{d}c_1} = w_h^e(c_1)D'(w_h^e(c_1) + w_\ell^e(c_1)) \frac{\mathrm{d}w_\ell^e(c_1)}{\mathrm{d}c_1},$$

which is positive if Firm 1's Cournot equilibrium quantity is decreasing in Firm 1's own cost, namely $\dfrac{\mathrm{d}w_\ell^e(c_1)}{\mathrm{d}c_1} < 0$. The following presents comparative statics on Cournot equilibria with multiple goods, and its proof is in Appendix A.

Lemma 3. *Suppose that the demand function D is concave. Firm 1's Cournot equilibrium quantity $w_\ell^e(c_1)$ is decreasing in its own cost. Firm 2's Cournot equilibrium quantity $w_h^e(c_1)$ is increasing in the rival's cost.*

6.1. Entry deterrence and accommodation

How can incumbent Firm 1's choices in period 1 affect entry? We begin with:

Definition 3 (*Entry Deterrence*). Given any FC, let Firm 1 choose a quantity in period 1 so low that in period 2, Firm 2 cannot earn a positive profit in the Cournot equilibrium. Firm 1's quantity in period 1 is said to be an entry deterrence quantity.

Formally, define w_1^D, with $c_1^D = C(w_1^D)$, by

$$\pi_2^e(c_1^D) = w_h^e(c_1^D)[D(w_h^e(c_1^D))(q-1) + D(w_h^e(c_1^D) + w_\ell^e(c_1^D)) - c_2] - FC = 0, \tag{28}$$

where $\pi_2^e(c_1^D) \equiv \pi_2(w_\ell^e(c_1^D), w_h^e(c_1^D))$ is Firm 2's Cournot equilibrium profit post entry. Firm 1's quantity in period 1, w_1, is said to deter entry if $w_1 \le w_1^D$.

The quantity w_1^D is the critical value for Firm 1 to deter entry. Any quantity w_1 smaller than w_1^D will make Firm 2's Cournot equilibrium profit vanish.[12] Inverting (28), we obtain $c_1 = \phi(FC)$. Now, at Firm 2's fixed cost FC, the value $\phi(FC)$ gives Firm 1's period 1 (highest) cost to deter entry. (In case Firm 2's fixed cost is so low that entry cannot be prevented, then we just set $\phi(FC)$ to be an arbitrary negative number.)

If the monopoly quantity w_1^M in Proposition 4 results in a cost $C(w_1^M) \le \phi(FC)$, then entry is deterred automatically. If the monopoly

[12] By Lemmas 2 and 3, the value of w_1^D is well defined.

quantity is higher, $C(w_1^M) > \phi(FC)$, Firm 1 deters entry by suppressing production. So suppose that Firm 1 produces $w_1 \leq w_1^D < w_1^M$. The profit in period 1 becomes $w_1 D(w_1)$. Then in period 2 its cost is $c_1 = C(w_1)$. Because entry has been deterred, Firm 1's period 2 profit is $w[D(w)-C(w_1)]$ from producing w in period 2. Given entry deterrence, Firm 1's period 2 maximum profit is $\max_w w[D(w)-C(w_1)]$. The optimal quantity to deter entry is

$$\arg \max_{w_1 \leq w_1^D} \{w_1 D(w_1) + \max_w w[D(w) - C(w_1)]\}. \tag{29}$$

The threat of entry makes the incumbent produce less, reducing resistance! We present the following (with its proof in Appendix A):

Proposition 7. *If Firm 1 deters entry by producing w_1 with $C(w_1) \leq \phi(FC)$, it reduces antibiotic resistance compared to the monopoly level in Proposition 4. That is, in an equilibrium with entry deterred, in period 1 Firm 1 produces less than w_1^M in Proposition 4, Firm 1's profit falls from the monopoly level, and the cost increase in period 2 is lower than under the monopoly regime.*

Next, we consider the entry-accommodation strategy. Here, in period 1 Firm 1 produces w_1 with $C(w_1) > \phi(FC)$. This means a higher profit in period 1 than deterrence, but now Firm 2 enters, and the Cournot equilibrium results. Firm 1's profits in the two periods sum to

$$w_1 D(w_1) + \max_{w_\ell} w_\ell [D(w_h^e + w_\ell) - C(w_1)] = w_1 D(w_1) + w_\ell^e [D(w_h^e + w_\ell^e) - C(w_1)],$$

where (w_ℓ^e, w_h^e) are the Cournot equilibrium quantities. Firm 1's period 1 equilibrium quantity is

$$\arg \max_{w_1 \geq w_1^D} \{w_1 D(w_1) + w_\ell^e(w_1)[D(w_h^e(c_1) + w_\ell^e(c_1)) - c_1]\}, \tag{30}$$

where $c_1 = C(w_1)$, and $(w_\ell^e(c_1), w_h^e(c_1))$ are the Cournot equilibrium quantities in period 2. We present the following (with its proof in Appendix A).

Proposition 8. *If Firm 1 accommodates entry by producing w_1 where $C(w_1) > \phi(FC)$, it may or may not exacerbate antibiotic resistance compared to the monopoly level in Proposition 4, so the cost in period 2 may be higher or lower than under monopoly.*

When entry is accommodated, the first-order derivative of Firm 1's profit with respect to quantity w_1 is

$$w_1 D'(w_1) + D(w_1) + w_\ell^e(w_1)[D'(w_h^e(w_1) + w_\ell^e(w_1))] \frac{\mathrm{d}w_h^e(w_1)}{\mathrm{d}w_1} - w_\ell^e(w_1) C'(w_1).$$

This is to be compared with the optimal monopoly choice of period 1 quantity, characterized by

$$w_1 D'(w_1) + D(w_1) - w_2 C'(w_1),$$

where w_1 and w_2 are optimal quantities in the two periods under monopoly. The extra term under entry accommodation in Cournot competition is

$$w_\ell^e(w_1)[D'(w_h^e(w_1) + w_\ell^e(w_1))] \frac{\mathrm{d}w_h^e(w_1)}{\mathrm{d}w_1}. \tag{31}$$

By Lemma 3, Firm 2's quantity is increasing in Firm 1's period 1 cost, so the derivative $\dfrac{\mathrm{d}w_h^e(w)}{\mathrm{d}w}$ is positive.

There are two opposing effects in entry accommodation. First, in period 2, Firm 1 produces only w_ℓ of the total market quantity $w_\ell + w_h$. In period 1, Firm 1 does not internalize the full resistance externality; this is an incentive to raise quantity. Second, there is a strategic consideration, and this is (partly) reflected by the extra term in (31). Firm 1 earns more period 2 profit when its cost is lower; this lower cost actually makes Firm 2 reduce its quantity (w_h). To lower cost, Firm 1 has an incentive to lower quantity in period 1. Which effect dominates depends on the demand and cost functions, so the net effect is ambiguous.

A numerical example of the entry deterrence and accommodation strategies is presented in Appendix B. By varying the entrant's fixed cost, we illustrate how deterrence and accommodation may affect resistance.

7. Policy discussions

How do our theoretical models relate to policies? The current policies can roughly be divided into financial and non-financial approaches. The former are antibiotics being paid for by subscriptions, push–pull financial incentives; the latter is a set of stewardship programs.

The UK's 2024 Antimicrobial Products Subscription Model is the first subscription-based payment. The National Health Service pays pharmaceutical companies fixed annual fees based on therapeutic values. In return, companies no longer receive revenue from NHS-ordered prescriptions (NHS England, 2024). Similarly, the PASTEUR Act in the U.S. would authorize subscription contracts worth up to $6 billion over ten years (Lienhard, 2020). It is thought that delinking revenues from actual antibiotic prescription volumes would dull over-prescription incentives.[13]

In a competitive market, profit-maximizing firms do sell more, as we show in Section 4; subscription payment works well against resistance. However, if the firm is a monopoly or enjoys significant market power, the firm does internalize the cost-quality degradation due to overuse, as we show in Section 5. Monopoly and significant market power also mean higher prices. The subscription payment removes the monopoly pricing distortion, but it also eliminates the monopolist's internalization of the cost-quality externality.[14] In practice, the NHS charges hospitals at a nationally agreed invoice price through normal distribution channels. Whether the invoice price can be set correctly remains an open question. In effect, the NHS has taken over the internalization responsibility.

Besides subscription, current policy discussions are about subsidizing fixed and variable costs, respectively push and pull incentives, which concern new drug developments. A push policy directly encourages entry, which reduces an incumbent firm's incentive to internalize resistance. A pull incentive, by contrast, subsidizes the entrant's marginal cost, so alters the post-entry equilibrium, but ultimately also encourages entry. As we have shown in Section 6, the incumbent loses some internalization incentives when entry is accommodated, an unintended consequence of push and pull policies. The incumbent deterring entry leads to resistance alleviation, but consumers will not have a new drug.[15]

Long-term subscription policies theoretically may replace push–pull incentives because firms are effectively selling their entire production line to the government. Even entry incentives may be managed by subscription, because a potential entrant is effectively innovating on behalf of the government. It is an open question if the government can solve a fundamental market failure better than other market-like instruments, such as competitive drug plans, and entry deterrence and accommodation strategic effects against resistance.

Financial policies on firms must be complemented by non-price subscription policies because firms no longer sell to consumers directly. In fact, the U.S. CDC has developed the Core Elements of Outpatient

[13] Outterson et al. (2015) proposed such delinking schemes.

[14] By extension, in an imperfectly competitive market, firms do internalize some of the resistance externality.

[15] Patent protection may help research and development. However, patent protection applies to all innovations, not special to antibiotics. Fine tuning of patent length to account for resistance externality is likely needed here, but this is beyond the scope of the current analysis.

Antibiotic Stewardship, which is a set of rationing guidelines.[16] In effect, these guidelines are non-price rationing schemes. In theory, an NHS regulator can ration uses to mimic the first best. In practice, this is unlikely. First, the effectiveness vs resistance tradeoff has to be patient dependent, and a broad policy may not contain enough flexibility. Second, prescriptions have to be decentralized, so another free-riding problem crops up. Individual providers may overprescribe when pressured by patients. Again, the fundamental problem remains that prescribing providers do not actually internalize the resistance externality.

Beyond the market structure interventions analyzed in this paper, other approaches may be explored. Educating consumers and persuading them to use antibiotics more carefully may be likened to shrinking the demand function. The ultimate technological advance is to develop antibiotics that avoid resistance entirely.

8. Conclusion

We have presented a theory of antibiotic resistance, the excessive use of an antibiotic degrading its power to kill microbes. We model resistance as an increase in cost or a decrease in quality. The degradation is like the depletion of a common resource, a tragedy-of-the-commons free riding problem. Whereas the extant literature has recognized the resistance externality, we analyze equilibrium outcomes in perfect competition, managed-care competition, monopoly, and imperfect competition under potential entry.

Perfect competition leads to maximum resistance. Competitive drug plans achieve partial internalization. Monopoly fully internalizes but extracts consumer surplus through pricing. We derive conditions for monopoly to yield higher consumer surplus than competition. A monopolist offering a subscription-type drug plan replicates the first best while extracting all consumer surplus. When there is a potential entrant, the incumbent deterring entry reduces resistance, but accommodating entry may have ambiguous effects on resistance.

The primitives in our model are a demand function and a cost function, which can be estimated from data. Bokhari et al. (2024) estimate demand for antibiotics and simulate the effects of taxes. Filippini et al. (2006, 2009) and Filippini and Masiero (2012) document consumption patterns and persistence. Nelson et al. (2021) estimate the costs of resistance. Combined with these empirical foundations, the model can inform the calibration of policies.

Our analysis complements policy discussions by reframing the resistance externality in a market context. Indeed, we take seriously that pharmaceutical firms operate in markets. Neither do we assume that resistance is to be completely eliminated. We perform welfare analysis with respect to various market outcomes.

CRediT authorship contribution statement

Miaoqing Jia: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Ching-to Albert Ma:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization.

Acknowledgments

Our thanks to Rena Conti and Kevin Outterson for their early feedback and comments. Abigail Dow and Rochelle Rayan provided excellent research assistance. Our thanks for seminar participants' comments at Academia Sinica, Taipei, Taiwan; Bocconi University, Milan; Boston University; Chinese University of Hong Kong; Hong Kong Baptist University; Monash University, Melbourne; National Sun Yat-Sen University, Kaohsiung, Taiwan; National Taiwan University; Taiwan Society of Health Economics; University of Central Florida, Orlando; University of Hamburg; University of Sydney; and at the International Industrial Organization Conference in Boston, 2026. An editor and two referees gave us valuable suggestions.

[16] https://www.cdc.gov/antibiotic-use/hcp/core-elements/outpatient-antibiotic-stewardship.html

Appendix A. Proofs

Proof of Lemma 1

The first-order conditions with respect to w_1 and w_2 are (10) and (11) in Proposition 3. Eqs. (10) and (11) define w_1 and w_2 as implicit functions of s. We differentiate (10) to obtain

$$D'(w_1)\frac{\mathrm{d}w_1}{\mathrm{d}s} - w_2\left[sC''(1-s+sw_1)(-1+w_1+s\frac{\mathrm{d}w_1}{\mathrm{d}s}) + C'(1-s+sw_1)\right] - sC'(1-s+sw_1)\frac{\mathrm{d}w_2}{\mathrm{d}s} = 0.$$

Next, we differentiate (11) to obtain

$$D'(w_2)\frac{\mathrm{d}w_2}{\mathrm{d}s} - C'(1-s+sw_1)(-1+w_1) - sC'(1-s+sw_1)\frac{\mathrm{d}w_1}{\mathrm{d}s} = 0.$$

Collecting terms and simplifying, we obtain

$$\begin{aligned}&[D'(w_1) - s^2w_2C''(1-s+sw_1)]\frac{\mathrm{d}w_1}{\mathrm{d}s} - sC'(1-s+sw_1)\frac{\mathrm{d}w_2}{\mathrm{d}s}\\ &\quad = w_2\left[sC''(1-s+sw_1)(-1+w_1) + C'(1-s+sw_1)\right]\\ &-sC'(1-s+sw_1)\frac{\mathrm{d}w_1}{\mathrm{d}s} + D'(w_2)\frac{\mathrm{d}w_2}{\mathrm{d}s} = C'(1-s+sw_1)(-1+w_1).\end{aligned}$$

We rewrite the two simultaneous equations in matrix form

$$\begin{bmatrix} D'(w_1) - s^2w_2C''(1-s+sw_1) & -sC'(1-s+sw_1) \\ -sC'(1-s+sw_1) & D'(w_2) \end{bmatrix}\begin{bmatrix} \frac{\mathrm{d}w_1}{\mathrm{d}s} \\ \frac{\mathrm{d}w_2}{\mathrm{d}s} \end{bmatrix} = \begin{bmatrix} w_2\left[sC''(1-s+sw_1)(-1+w_1) + C'(1-s+sw_1)\right] \\ C'(1-s+sw_1)(-1+w_1) \end{bmatrix}.$$

The determinant of the left-hand side parameter matrix is positive due to the second-order conditions for the maximization of (8); we call this determinant $|A|$. We then apply Cramer's rule to solve for the derivatives $\frac{dw_1}{ds}$ and $\frac{\mathrm{d}w_2}{\mathrm{d}s}$. After simplification we have

$$\frac{\mathrm{d}w_1}{\mathrm{d}s} = \frac{w_2D'(w_2)\left[C'(1-s+sw_1) - sC''(1-s+sw_1)(1-w_1)\right] - sC'(1-s+sw_1)^2(1-w_1)}{|A|}.$$

Next, we have

$$\frac{\mathrm{d}w_2}{\mathrm{d}s} = \frac{-\left[D'(w_1) - s^2w_2C''(1-s+sw_1)\right]C'(1-s+sw_1)(1-w_1)}{|A|} + \frac{sC'(1-s+sw_1)w_2\left[C'(1-s+sw_1) - sC''(1-s+sw_1)(1-w_1)\right]}{|A|}.$$

Hence, if C'' is sufficiently small, $\frac{\mathrm{d}w_1}{\mathrm{d}s}$ is negative, and $\frac{\mathrm{d}w_2}{\mathrm{d}s}$ is positive.

Proof of Proposition 3

We construct the unique Drug Plan Competitive Equilibrium in the following way. First, for any given $B < \overline{B}$, find w_1^e to satisfy (9). Because $B < \overline{B}$, such a w_1^e satisfies $w_1^* < w_1^e < 1$. Second, given w_1^e, choose s^e and w_2^e to satisfy (10) and (11). Finally, set premium π^e by (12).

Now we confirm that we have a Drug Plan Competitive Equilibrium. Because $B < \overline{B}$, such a w_1^e from (9) has $w_1^* < w_1^e < 1$. Eqs. (10) and (11) are the first-order conditions for $(\widetilde{w}_1(s), \widetilde{w}_2(s))$ in (8). The concavity of the social welfare function (1) implies that the objective in (8) is

concave. Hence, given s, $(\widetilde{w}_1(s), \widetilde{w}_2(s))$ is unique. By Lemma 1, as s varies between 0 and 1, $\widetilde{w}_1(s)$ monotone decreases from 1 to w_1^*. Hence, there is an s^e such that (10) and (11) would hold for that $s = s^e$. Then obviously, $w_2^e = \widetilde{w}_2(s^e)$. By construction, consumers, ex ante, are indifferent between enrolling and staying independent, so it is optimal for s^e of consumers to join. We have constructed the unique Drug Plan Competitive Equilibrium for $B < \overline{B}$.

Finally, consider $B > \overline{B}$. At $B = \overline{B}$, all consumers optimally choosing the drug plan is an equilibrium. Any higher B will not change the equilibrium.

Proof of Lemma 2

By the envelope theorem, Firm 1's maximum profit has a derivative with respect to cost c_1 equal to the partial derivative of the profit function with respect to cost c_1, and this partial derivative is $-w_\ell < 0$. Next, the cross partial of Firm 1's profit with respect to cost c_1 and quantity w_ℓ is

$$\frac{\partial^2 w_\ell [D(w_h + w_\ell) - c_1]}{\partial c_1 \partial w_\ell} = -1.$$

The derivative of Firm 1's profit-maximizing quantity with respect to cost has the same sign as this cross partial, so it is also negative.

By the envelope theorem again, Firm 2's maximum profit has a derivative with respect to Firm 1's quantity equal to the partial derivative with respect to w_ℓ, which is $w_h D'(w_h + w_\ell) < 0$ because the demand D is a decreasing function.

Proof of Lemma 3

Totally differentiating the first-order conditions respectively for the maximizations of (20) and (21) (which then define a Cournot equilibrium in (22) and (23)) we have

$$\frac{\partial^2 \pi_1(w_\ell^e, w_h^e)}{\partial w_\ell^2} \mathrm{d}w_\ell + \frac{\partial^2 \pi_1(w_\ell^e, w_h^e)}{\partial w_\ell \partial w_h} \mathrm{d}w_h = -\frac{\partial^2 \pi_1(w_\ell^e, w_h^e)}{\partial w_\ell \partial c_1} \mathrm{d}c_1$$

$$\frac{\partial^2 \pi_2(w_\ell^e, w_h^e)}{\partial w_h \partial w_\ell} \mathrm{d}w_\ell + \frac{\partial^2 \pi_2(w_\ell^e, w_h^e)}{\partial w_h^2} \mathrm{d}w_h = -\frac{\partial^2 \pi_2(w_\ell^e, w_h^e)}{\partial w_h \partial c_2} \mathrm{d}c_2,$$

and from (24) and (25) they turn out to be

$$\left[2D'(w_h^e + w_\ell^e) + D''(w_h^e + w_\ell^e)\right] \mathrm{d}w_\ell + \left[D'(w_h^e + w_\ell^e) + w_\ell^e D''(w_h^e + w_\ell^e)\right] \mathrm{d}w_h = \mathrm{d}c_1$$

$$\left[D'(w_h^e + w_\ell^e) + w_h^e D''(w_h^e + w_\ell^e)\right] \mathrm{d}w_\ell + \{2D'(w_h^e + w_\ell^e) + w_h^e D''(w_h^e + w_\ell^e) + (q-1)[2D'(w_h^e) + D''(w_h^e)]\} \mathrm{d}w_h = 0.$$

We express this system in matrix form

$$\begin{bmatrix} 2D'(w_h^e + w_\ell^e) + D''(w_h^e + w_\ell^e) & D'(w_h^e + w_\ell^e) + w_\ell^e D''(w_h^e + w_\ell^e) \\ D'(w_h^e + w_\ell^e) + w_h^e D''(w_h^e + w_\ell^e) & 2D'(w_h^e + w_\ell^e) + w_h^e D''(w_h^e + w_\ell^e) + (q-1)[2D'(w_h^e) + D''(w_h^e)] \end{bmatrix} \begin{bmatrix} \mathrm{d}w_\ell \\ \mathrm{d}w_h \end{bmatrix} = \begin{bmatrix} \mathrm{d}c_1 \\ 0 \end{bmatrix}.$$

We use Cramer's Rule to compute

$$\frac{\mathrm{d}w_\ell^e}{\mathrm{d}c_1} = \frac{\det(A_\ell)}{\det(A)},$$

where $\det(A)$ is the determinant of the coefficient matrix A and $\det(A_\ell)$ is the determinant of the matrix obtained by replacing the first column of A with the transpose of $(1, 0)$.

The determinant of A is

$$\begin{aligned} \det(A) = {} & \left[2D'(w_h^e + w_\ell^e) + D''(w_h^e + w_\ell^e)\right] \times \{2D'(w_h^e + w_\ell^e) \\ & + w_h^e D''(w_h^e + w_\ell^e) + (q-1)[2D'(w_h^e) + D''(w_h^e)]\} \\ & - \left[D'(w_h^e + w_\ell^e) + w_\ell^e D''(w_h^e + w_\ell^e)\right] \\ & \times \left[D'(w_h^e + w_\ell^e) + w_h^e D''(w_h^e + w_\ell^e)\right]. \end{aligned}$$

The determinant of A_ℓ is

$$\begin{aligned} \det(A_\ell) = {} & \left[\mathrm{d}c_1\right] \times \{2D'(w_h^e + w_\ell^e) + w_h^e D''(w_h^e + w_\ell^e) \\ & + (q-1)[2D'(w_h^e) + D''(w_h^e)]\} \\ & - \left[D'(w_h^e + w_\ell^e) + w_\ell^e D''(w_h^e + w_\ell^e)\right] \times 0 \\ = {} & \mathrm{d}c_1 \times \{2D'(w_h^e + w_\ell^e) + w_h^e D''(w_h^e + w_\ell^e) \\ & + (q-1)[2D'(w_h^e) + D''(w_h^e)]\}. \end{aligned}$$

By the concavity of the demand function, we have $D''(w) \le 0$, so $\det(A)$ is positive and $\det(A_\ell)$ is negative. Therefore,

$$\frac{\mathrm{d}w_\ell^e}{\mathrm{d}c_1} = \frac{\det(A_\ell)}{\det(A)} < 0.$$

Next, $\det(A_h)$ is the determinant of the matrix obtained by replacing the second column of A with $(\mathrm{d}c_1, 0)^T$ and therefore

$$\begin{aligned} \det(A_h) &= \left[2D'(w_h^e + w_\ell^e) + D''(w_h^e + w_\ell^e)\right] \cdot 0 - \left[D'(w_h^e + w_\ell^e) + w_h^e D''(w_h^e + w_\ell^e)\right] \cdot \mathrm{d}c_1 \\ &= -\left[D'(w_h^e + w_\ell^e) + w_h^e D''(w_h^e + w_\ell^e)\right] \cdot \mathrm{d}c_1. \end{aligned}$$

By the concavity of the demand function, we have $(D''w) \le 0$, so $\det(A_h)$ is positive. Recall that $\det(A)$ is positive; thus we get

$$\frac{\mathrm{d}w_h^e}{\mathrm{d}c_1} = \frac{\det(A_h)}{\det(A)} > 0.$$

Proof of Proposition 7

Let $c_1^M = C(w_1^M)$, where w_1^M is the monopoly period 1 quantity in Proposition 4. Either the monopoly period 1 quantity yields a cost $c_1^M = C(w_1^M)$ below or above the deterrence threshold. That is, either $c_1^M \le \phi(FC)$, or $c_1^M > \phi(FC)$. If $c_1^M \le \phi(FC) = C(w_1^D)$, Firm 1's monopoly quantity w_1^M already deters entry, so the cost increase in period 2 will remain the same. If $c_1^M > \phi(FC)$, then Firm 1 must produce a period 1 quantity less than w_1^D. The constraint $w_1 \le w_1^D$ in (29) binds. Because $c_1^M > \phi(FC) = C(w_1^D)$, we have $w_1 < w_1^M$. Firm 1 produces less than the monopoly level. Firm 1's total profit must decrease because of the binding constraint $w_1 \le w_1^D < w_1^M$.

Proof of Proposition 8

Recall that $w_1^M > w_1^D$ so that the period 1 monopoly quantity is larger than the entry deterrence threshold. To accommodate entry, Firm 1 chooses quantities according to (30). The monopolist chooses (w_1, w_2) to maximize profit in (13). The two maximization problems are respectively:

$$\max_{w_1 \ge w_1^D} \{w_1 D(w_1) + w_\ell^e(w_1)[D(w_h^e(w_1) + w_\ell^e(w_1)) - C(w_1)]\} \quad \text{and}$$

$$\max_{w_1, w_2} w_1 D(w_1) + w_2[D(w_2) - C(w_1)].$$

The first-order derivative of Firm 1's entry-accommodating equilibrium profit with respect to w_1 is

$$w_1 D'(w_1) + D(w_1) - w_\ell^e(w_1) C'(w_1) + w_\ell^e(w_1)[D'(w_h^e(w_1) + w_\ell^e(w_1))] \frac{\mathrm{d}w_h^e(w_1)}{\mathrm{d}w_1},$$

where we have ignored the derivatives with respect to $w_\ell^e(w_1)$, by the envelope theorem. Notice that $w_\ell^e(w_1)$ must be less than what Firm 1 would have chosen if it was the only firm in period 2. Also, the last term, $w_\ell^e(w_1) D'(w_h^e(w_1) + w_\ell^e(w_1)) \frac{\mathrm{d}w_h^e(w_1)}{\mathrm{d}w_1}$, is negative because D' is negative and $\frac{\mathrm{d}w_h^e(w_1)}{\mathrm{d}w_1} > 0$ by Lemma 3. It is possible that w_1 may be higher or lower than w_1^M.

Appendix B. Numerical example of entry deterrence and accommodation

Using a numerical example, we illustrate entry deterrence and accommodation in the analysis in Section 6. We use linear demand and resistance cost functions (same notation): $D(w) = 1 - w$, and $C(w_1) = 0.5w_1$. The entrant's Drug 2 has a higher quality at $q = 1.5$ and marginal cost $c_2 = 0.1$.

B.1. Period 2 subgames

Cournot subgame. If Firm 2 enters in period 2, the two firms compete in quantities. Given the incumbent's period 2 cost $c_1 = C(w_1)$, and the firms' quantities w_ℓ and w_h, the market-clearing prices are

$$p_\ell = (1 - w_h - w_\ell) \quad \text{and} \quad p_h = (1 - w_h)(0.5) + (1 - w_h - w_\ell).$$

We write out the profit functions and use the first-order conditions to solve for the Cournot equilibrium quantities.[17] The equilibrium quantities are

$$w_\ell^e(c_1) = \frac{1.6 - 3c_1}{5}, \qquad w_h^e(c_1) = \frac{1.8 + c_1}{5}, \tag{32}$$

and firms' equilibrium profits are

$$\pi_1^e(c_1) = \frac{(1.6 - 3c_1)^2}{25}, \qquad \pi_2^e(c_1) = \frac{1.5(1.8 + c_1)^2}{25} - FC. \tag{33}$$

Monopoly subgame. If the entrant does not enter, Firm 1 remains a monopolist in period 2 with cost c_1. The optimal quantity and profit are

$$w_2^m(c_1) = \frac{1 - c_1}{2}, \qquad \pi_1^m(c_1) = \frac{(1 - c_1)^2}{4}. \tag{34}$$

B.2. Entry and fixed cost

Firm 2 enters if and only if $\pi_2^e(c_1) \geq 0$. From (33), the entrant's profit is strictly increasing in c_1: a higher incumbent cost allows the entrant to earn more profits. Define the cost threshold c_1^D by

$$\pi_2^e(c_1^D) = \frac{1.5(1.8 + c_1^D)^2}{25} - FC = 0; \tag{35}$$

for any given FC, Firm 2 enters if and only if $c_1 \geq c_1^D$.

Now the possible range of c_1 is between 0 and 0.5, corresponding to Firm 1, in period 1, producing nothing and supplying the full market. So it suffices to consider cost thresholds between 0 and 0.5. From (35), setting $c_1^D = 0$ yields $FC = 0.194$, and setting $c_1^D = 0.5$ yields $FC = 0.317$, which give the relevant range of fixed costs to study deterrence and accommodation.

We choose a relatively high value of FC to illustrate equilibrium deterrence, and then a small value of FC to illustrate accommodation. In each case, we consider the subgame perfect equilibrium profits under deterrence and accommodation. Incumbent Firm 1 will choose the deterrence period 1 quantity when FC is large; the opposite is true when FC is low.

B.3. Case 1: Equilibrium entry deterrence ($FC = 0.228$)

We set $FC = 0.228$, so the corresponding cost threshold, from (35), is $c_1^D = 0.149$, and deterrence requires that Firm 1 chooses $w_1 \leq w_1^D = 0.299$.

Deterrence profit. If the incumbent was the only firm, then it would choose the period 1 quantity w_1 to maximize the total profit, which is $w_1(1 - w_1) + \frac{(1 - 0.5w_1)^2}{4}$ without any constraint on w_1. It is straightforward to verify that this optimal quantity in period 1 would be 0.4. However, this is higher than value of $w_1^D = 0.299$, so the incumbent must cut back on supply to deter entry.

In fact, Firm 1 deters entry by choosing $w_1 = w_1^D = 0.299$. The total profit is

$$\Pi^D = w_1^D(1 - w_1^D) + \frac{(1 - c_1^D)^2}{4}, \tag{36}$$

where the first term is period 1 profit and the second term is period 2 monopoly profit from (34). We obtain the profit from deterrence

$$\Pi^D = 0.299 \times 0.701 + \frac{(1 - 0.149)^2}{4} = 0.390$$

by substituting $w_1^D = 0.299$ and $c_1^D = 0.149$ into the profit function just above.[18]

Accommodation profit. If Firm 1 accommodates entry, it may choose a period 1 quantity higher than 0.299. We let Firm 1 choose w_1 to maximize total profit under the constraint of $w_1 \geq 0.299$. The constraint is not binding, and the first-order condition yields $w_1^A = 0.444$ as the optimal accommodation quantity.[19]

Total profit under accommodation is

$$\Pi^A = w_1^A(1 - w_1^A) + \frac{(1.6 - 3c_1^A)^2}{25}, \tag{37}$$

where $c_1^A = 0.5 \times 0.444 = 0.222$. We obtain $\Pi^A = 0.282$ by substituting $w_1^A = 0.444$ and $c_1^A = 0.5w_1^A$ into the profit expression.[20]

Equilibrium. Since $\Pi^D = 0.390 > \Pi^A = 0.282$, the incumbent chooses entry deterrence. Equilibrium quantities and period 2 costs are as follows

$$w_1^* = 0.299, \quad c_1^* = 0.149, \quad w_2^* = 0.426 \text{ (monopoly)}.$$

Entry is deterred, and the incumbent restricts period 1 sales to keep its future cost low. This illustrates Proposition 7: entry deterrence reduces antibiotic resistance compared to the monopoly level without entry threat.

Fig. 5 plots the incumbent's total profit as a function of w_1. The profit function has a kink at $w_1^D = 0.299$: to the left, the incumbent deters entry and earns monopoly profit in period 2; to the right, entry occurs and Cournot competition ensues. The global maximum occurs at deterrence.

B.4. Case 2: Equilibrium entry accommodation ($FC = 0.196$)

Next, we set $FC = 0.196$, which is quite a bit smaller than the value for illustrating deterrence. The entry threshold, from (35), is $c_1^D = 0.007$, so deterrence requires $w_1 \leq w_1^D = 0.015$—almost no production in period 1.

Deterrence profit. The profit from deterrence is

$$\Pi^D = 0.015 \times 0.985 + \frac{(1 - 0.007)^2}{4} = 0.261, \tag{38}$$

which is obtained by substituting $w_1^D = 0.015$ and $c_1^D = 0.007$ into (36), the deterrence profit expression.[21]

[17] The profit functions are $\pi_1 = (1 - w_h - w_\ell - c_1)w_\ell$ and $\pi_2 = (1.4 - 1.5w_h - w_\ell)w_h - FC$. We use the first-order conditions $1 - w_h - 2w_\ell - c_1 = 0$ and $1.4 - 3w_h - w_\ell = 0$ to obtain (32).

[18] Period 1 profit is $0.299 \times 0.701 = 0.210$. Period 2 monopoly profit is $(1 - 0.149)^2/4 = 0.181$. Total becomes $0.210 + 0.181 = 0.390$.

[19] The first-order condition for maximizing $w_1(1 - w_1) + (1.6 - 1.5w_1)^2/25$ is $1 - 2w_1 - \frac{3(1.6 - 1.5w_1)}{25} = 0$, yielding $w_1^A = 20.2/45.5 \approx 0.444$.

[20] Period 1 profit: $0.444 \times 0.556 = 0.247$. Period 2 Cournot profit: $(1.6 - 0.666)^2/25 = 0.035$. Total: $0.247 + 0.035 = 0.282$.

[21] Period 1 profit: $0.015 \times 0.985 = 0.015$. Period 2 monopoly profit: $(1 - 0.007)^2/4 = 0.246$. Total: $0.015 + 0.246 = 0.261$.

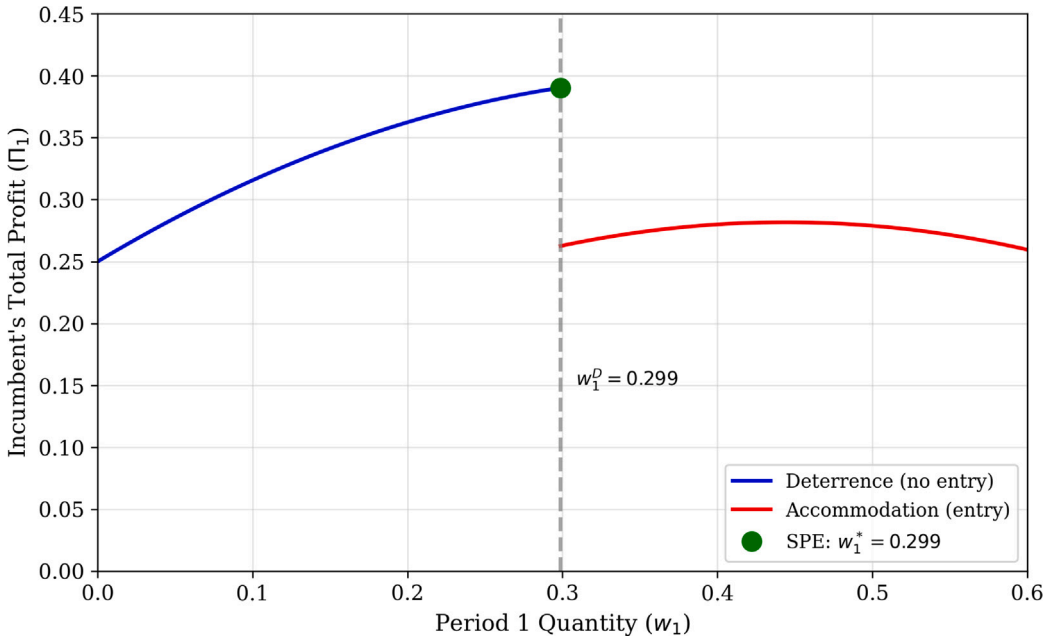

Fig. 5. Incumbent's total profit: Entry deterrence ($FC = 0.228$).
Notes: The blue curve shows profit when entry is deterred; the red curve shows profit when entry occurs. The vertical dashed line marks the entry threshold $w_1^D = 0.299$. The equilibrium is at $w_1^* = 0.299$.

Accommodation profit. The profit from accommodation is

$$\Pi^A = 0.444 \times 0.556 + \frac{(1.6 - 0.666)^2}{25} = 0.282, \tag{39}$$

which is obtained by substituting $w_1^A = 0.444$ and $c_1^A = 0.222$ into (37), the accommodation profit expression.[22]

Equilibrium. Because $\Pi^A = 0.282 > \Pi^D = 0.261$, the incumbent chooses entry accommodation even though deterrence is possible. Equilibrium quantities and period 2 cost are as follows:

$$w_1^* = 0.444, \quad c_1^* = 0.222, \quad w_h^e = 0.404, \quad w_\ell^e = 0.187.$$

Entry occurs because deterrence would require sacrificing too much period 1 profit. This illustrates Proposition 8: under accommodation, the incumbent has weaker incentives to internalize the resistance externality, leading to higher production and greater resistance.

Fig. 6 plots the incumbent's total profit for this case. The deterrence region ($w_1 \leq 0.015$) is extremely narrow, and the global maximum occurs in the accommodation region at $w_1^* = 0.444$. The horizontal dashed lines show that the accommodation profit ($\Pi^A = 0.282$) exceeds the deterrence profit ($\Pi^D = 0.261$).

B.5. Comparison

Table 1 summarizes the equilibrium outcomes for the two cases.

The comparison shows two points. First, the fixed cost parameter FC determines whether deterrence is profitable: higher fixed costs make entry less attractive, enabling the incumbent to deter at higher period 1 quantity. Second, when accommodation happens in equilibrium, the incumbent produces more in period 1, leading to higher resistance. We confirm Proposition 8—that market sharing weakens the incumbent's incentive to internalize the resistance externality.

Table 1
Equilibrium outcomes under entry deterrence and accommodation.

	Case 1: Deterrence ($FC = 0.228$)	Case 2: Accommodation ($FC = 0.196$)
Entry threshold c_1^D	0.149	0.007
Deterrence requires $w_1 \leq$	0.299	0.015
Π^D (deterrence profit)	0.390	0.261
Π^A (accommodation profit)	0.282	0.282
Equilibrium strategy	Deterrence	Accommodation
Period 1 quantity w_1^*	0.299	0.444
Period 2 cost c_1^*	0.149	0.222
Entry occurs?	No	Yes
Period 2 market structure	Monopoly	Cournot duopoly
Incumbent's total profit	0.390	0.282
Antibiotic resistance (compared to monopoly)	Alleviated	Exacerbated

Data availability

No data was used for the research described in the article.

[22] The accommodation profit $\Pi^A = 0.282$ is unchanged from Case 1, as it does not depend on FC.

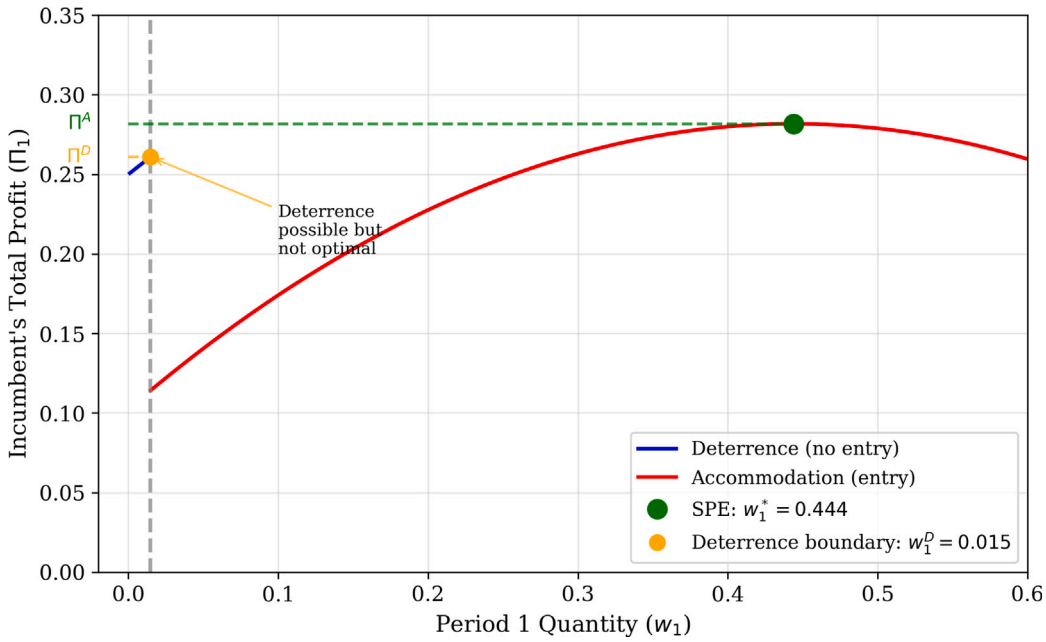

Fig. 6. Incumbent's total profit: Entry accommodation ($FC = 0.196$).
Notes: The deterrence region is narrow ($w_1 \leq 0.015$), marked by the vertical dashed line. Although deterrence is feasible (orange point), the equilibrium period 1 supply is in the accommodation region at $w_1^* = 0.444$ (green point); accommodation profit is higher than deterrence, $\Pi^A > \Pi^D$.

References

Adda, J., 2020. The economics of health epidemics: Preventing the spread of antibiotic resistance. Am. Econ. Rev. Pap. Proc. 110, 255–259.

Albert, J., 2021. Strategic dynamics of antibiotic use and the evolution of antibiotic-resistant infections. Int. J. Ind. Organ. 77, 102759.

Bokhari, F.A.S., Mariuzzo, F., Yan, W., 2024. Antibacterial resistance and the cost of affecting demand: The case of UK antibiotics. Int. J. Ind. Organ. 95, 103082.

Filippini, M., Masiero, G., 2012. An empirical analysis of habit and addiction to antibiotics. Empir. Econ. 42, 471–486.

Filippini, M., Masiero, G., Moschetti, K., 2006. Socioeconomic determinants of regional differences in outpatient antibiotic consumption: Evidence from Switzerland. Health Policy 78 (1), 77–92.

Filippini, M., Masiero, G., Moschetti, K., 2009. Regional consumption of antibiotics: A demand system approach. Econ. Model. 26 (6), 1389–1397.

Frank, R.G., Glazer, J., McGuire, T.G., 2000. Measuring adverse selection in managed health care. J. Health Econ. 19 (6), 829–854.

Giubilini, A., 2019. Antibiotic resistance as a tragedy of the commons: An ethical argument for a tax on antibiotic use in humans. Bioethics 33 (7), 776–784.

Glazer, J., McGuire, T.G., 2000. Optimal risk adjustment in markets with adverse selection: An application to managed care. Am. Econ. Rev. 90 (4), 1055–1071.

Glazer, J., McGuire, T.G., 2006. Optimal quality reporting in markets for health plans. J. Health Econ. 25 (2), 295–310.

Herrmann, M., 2010. Monopoly pricing of an antibiotic subject to bacterial resistance. J. Health Econ. 29, 137–150.

Herrmann, M., Gaudet, G., 2009. The economic dynamics of antibiotic efficacy under open access. J. Environ. Econ. Manag. 57 (3), 334–350.

Herrmann, M., Nkuiya, B., 2017. Inducing optimal substitution between antibiotics under open access to the resource of antibiotic susceptibility. Health Econ. 26 (6), 703–723.

Herrmann, M., Nkuiya, B., Dussault, A.-R., 2013. Innovation and antibiotic use within antibiotic classes: Market incentives and economic instruments. Resour. Energy Econ. 35, 582–598.

Klug, D.M., Idiris, F.I., Blaskovich, M.A., von Delft, F., Dowson, C.G., Kirchhelle, C., et al., 2021. There is no market for new antibiotics: this allows an open approach to research and development. Wellcome Open Res. 6, 146.

Kuhn, M., Prskawetz, A., Wrzaczek, S., Feichtinger, G., 2011. Externalities in a life cycle model with endogenous survival. J. Math. Econom. 47 (4–5), 627–641.

Laxminarayan, R., Brown, G.M., 2001. Economics of antibiotic resistance: A theory of optimal use. J. Environ. Econ. Manag. 42 (2), 183–206.

Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A.K.M., Wertheim, H.F.L., Sumpradit, N., et al., 2013. Antibiotic resistance—the need for global solutions. Lancet Infect. Dis. 13 (12), 1057–1098.

Laxminarayan, R., Weitzman, M.L., 2002. On the implications of endogenous resistance to medications. J. Health Econ. 21 (4), 709–718.

Lienhard, K., 2020. Bennet, Young bill would create new antibiotic subscription model. InsideHealthPolicy Com's FDA Week 26 (40), 1–11.

Nelson, R.E., Hatfield, K.M., Wolford, H., Samore, M.H., Scott, R.D. II, Reddy, S.C., Olubajo, B., Paul, P., Jernigan, J.A., Baggs, J., 2021. National estimates of healthcare costs associated with multidrug-resistant bacterial infections among hospitalized patients in the United States. Clin. Infect. Dis. 72 (Supplement 1), S17–S26.

NHS England, 2024. Antimicrobial products subscription model: Guidance on commercial arrangements. https://www.england.nhs.uk/long-read/antimicrobial-products-subscription-model-guidance-on-commercial-arrangements/.

Outterson, K., 2009. The legal ecology of resistance: The role of antibiotic resistance in pharmaceutical innovation. Cardozo Law Rev. 31 (3), 613–678.

Outterson, K., Powers, J.H., Daniel, G.W., McClellan, M.B., 2015. Repairing the broken market for antibiotic innovation. Health Aff. 34 (2), 277–285.

Van Boeckel, T.P., Brower, C., Gilbert, M., Grenfell, B.T., Levin, S.A., Robinson, T.P., Teillant, A., Laxminarayan, R., 2015. Global trends in antimicrobial use in food animals. Proc. Natl. Acad. Sci. 112 (18), 5649–5654.