Paying for pharmaceuticals: uniform pricing versus two-part tariffs*

Kurt R. Brekke† Dag Morten Dalen‡ Odd Rune Straume§

February 2022

Abstract

Two-part pricing (the Netflix model) has recently been proposed instead of uniform pricing for pharmaceuticals. Under two-part pricing the health plan pays a fixed fee for access to a drug at unit prices equal to marginal costs. Despite two-part pricing being socially efficient, we show that the health plan is worse off when the drug producer is a monopolist, as all surplus is extracted. This result is reversed with competition, as two-part pricing yields higher patient utility and lower drug costs for the health plan. However, if we allow for exclusive contracts, uniform pricing is preferred by the health plan. The choice of payment scheme is also shown to influence on the incentives to spend resources on drastic innovations relative to incremental, me-too innovations.

Keywords: Pharmaceuticals; Health Plans; Payment schemes

JEL Classification: I11; I18; L13; L65

*We thank two anonymous referees for very helpful suggestions. This paper is partly financed by National Funds of the FCT – Portuguese Foundation for Science and Technology within the project UIDB/03182/2020.

†Norwegian School of Economics (NHH), Department of Economics, Centre for Business Economics, Helleveien 30, N-5045 Bergen, Norway. E-mail: kurt.brekke@nhh.no

‡Corresponding author. BI Norwegian Business School, N-0442 Oslo, Norway. E-mail: dag.m.dalen@bi.no

§Department of Economics/NIPE, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal; and Department of Economics, University of Bergen. E-mail: o.r.straume@eeg.uminho.pt

1 Introduction

Pharmaceutical firms need to make large, up-front investments in drug discovery and clinical trials before new drugs can be approved and launched in the market. To stimulate such investments, pharmaceutical firms are granted patent protection, which in many cases gives rise to market power. Although market power is deliberately granted, and intended to allow pharmaceutical firms to charge higher prices over a given period of time, controversies over excessive prices are common. New life-extending drugs for cancer treatment, orphan drugs that target rare but severe diseases, and new HIV drugs are examples of drugs that have come under scrutiny for very high prices.[1]

Inducing innovation by allowing firms to charge high prices for new drugs runs the risk of reducing access to new treatments in spite of relatively low variable costs of production. This illustrates the familiar trade-off between static and dynamic efficiency. Public health plans in different countries try to mitigate supply-side market power by employing various price control mechanisms, such as direct regulation of drug prices (price caps) or regulation of the reimbursement level (reference pricing). To what extent price controls curtail market power, though, depends on the availability of close substitutes and documented treatment effects. Drugs that offer substantial health improvements will often be able to charge higher prices within such regulatory frameworks.

Recently, new drug pricing mechanisms, referred to as the Netflix model applied to pharmaceuticals, have been proposed to decouple the utilisation of a new drug from payments that allow recoupment of the firm's development costs. Instead of paying a price per package, the health plan negotiates a fixed amount (access or subscription fee) in exchange for unlimited prescription volume. Australia has adopted this approach to providing antivirals (DAAs) to patients with hepatitis C virus. According to Moon and Erikson (2019), the Australian authority negotiated an agreement to spend approximately 1 billion Australian dollars over 5 years to get unlimited access to the drugs. Two states in the US, Louisiana and Washington, have recently adopted similar purchasing strategies for the same class of drugs.[2]

As a two-part tariff, the so-called Netflix model has the intuitive advantages of giving access to the individual customers at very low marginal cost for the health plan without preventing suppliers from profiting from valuable innovations. Translating these effects into pharmaceutical markets, however, requires attention to specific institutional characteristics, including decisions made by

[1] See Mailankody and Prasad (2016).

[2] See https://www.gov.louisiana.gov/index.cfm/newsroom/detail/2031 and https://www.hca.wa.gov/assets/program/hep-c-elimination-gov-directive-18-13-final.pdf

health plans and competition between suppliers of drugs that target the same group of patients. In this paper, we compare the performance of two-part tariffs and uniform prices, both from the health plan's and drug producers' point of view. We model a market for on-patent prescription drugs that is served either by a monopolist or by two different producers supplying therapeutically substitutable drugs. A drug producer can only gain access to the market if the health plan is willing to sign a contract with this producer, and these contractual decisions determine which drugs can be prescribed by physicians affiliated with the health plan.

Although two-part tariffs have the advantage of lowering marginal prices for usage, they do not *per se* prevent firms from extracting surplus from health plans and procurers of pharmaceuticals. To what extent these new pricing schemes will reduce overall drugs costs, fixed fees included, depends on the firms' market power and the potential for therapeutic competition between two drug producers. When the quality difference between the two drugs is sufficiently large, turning the high-quality supplier into a *de facto* monopolist, we confirm the well-known property of two-part tariffs: The price charged per prescription equals marginal costs and the fixed fee extracts the entire surplus of the health plan. Although two-part tariffs ensure efficient access to the drug, the health plan prefers uniform pricing if therapeutic competition is out of reach.

Once we introduce therapeutic competition between the two suppliers, there is a striking change in the health plan's preference ranking of the two payment schemes. The reason for this is two-fold. First, with therapeutic competition, the credible threat of excluding one of the drugs from the health plan protects the plan against the aggressive surplus extraction when the drug producer is a monopolist. Second, two-part tariffs eliminate an allocative inefficiency that is present under monopoly pricing. Although different patented drugs belong to the same therapeutic class, they are not perfect substitutes. Marginal cost pricing associated with two-part tariffs ensures that the individual patients are allocated to the best drug, taking into account both treatment effects and costs. With uniform pricing, the allocation of patients between the two drugs is distorted since the firm with highest quality can exploit its demand advantage by setting a higher price.

The two-part tariff scheme (so-called Netflix model) therefore improves access to high-quality drugs, in line with its advocacy among policymakers. The distribution of surplus, however, crucially depends on the competition regime. Drug suppliers tend to prefer uniform pricing under therapeutic competition, and two-part tariffs under monopoly, which is the exact opposite of the ranking of payment regimes by the health plan. Thus, if we consider a game in which the health plan is

given the power to determine the type of payment scheme at the outset, the subgame perfect Nash equilibrium would have two-part tariffs only in the presence of (sufficiently strong) therapeutic competition. This conclusion relies on the assumption that the health plan will include all drugs that contributes to increasing its surplus. However, in an extension to the main analysis we show that if we include the possibility of exclusive contracts, where only one of the drugs is included in the health plan, then uniform pricing is always weakly preferred by the purchaser.

Our study builds on, and bridges, two different strands of literature. On the one side, we employ a standard theoretical framework applied to pharmaceutical markets for studying price competition and regulatory policies under imperfect competition. Due to heterogeneity among patients and differences in quality and the workings of different drugs, models combining horizontal and vertical differentiation have been useful in capturing important features of pharmaceutical markets, both demand-side and supply-side characteristics (see for example Brekke et al. (2007), Miraldo (2009), Bardey et al. (2010), Bardey et al. (2016), Brekke et al. (2016) and Gonzàles et al. (2016)). Among these, the general set-up in our paper relates most closely to the spatial formulation in Brekke et al. (2007) and Miraldo (2009).

The theoretical framework allows us to analyse two-part tariffs both under monopoly and with imperfect competition between producers of imperfect substitutes. The same therapeutic class will often contain several drugs with different active ingredients. These are not perfect substitutes, but with sufficiently overlapping treatment effects, prices can influence on prescription choices within the therapeutic class. Kanavos et al. (2007) analysed the existence of competition between branded statins in European markets prior to patent expiry. Their results is consistent with potential price sensitivity in the branded market for statins. Danzon and Epstein (2012) found that prices of new drugs are influenced by prices of other products in the same class. Lu and Comanor (1998) analysed therapeutic competition and found that launch prices of drugs that are closer substitutes to existing brands are typically priced at comparable levels. In addition, they found that the number of branded substitutes has a substantial negative effect on launch prices. In our model we allow intensity of therapeutic competition to depend on both drug characteristics and the price sensitivity of physicians' prescription choices.

The novel contribution of our paper in relation to this strand of literature is two-fold. First, we introduce two-part pricing into this framework, whereas previous studies focus solely on uniform pricing. As we will show, two-part pricing not only changes the drug producers' pricing decisions

(including unit prices), but indeed also the intensity of competition among the producers. Second, we allow for the health plan to decide on market access of one or both of the drugs, implying that drug producers compete both for the market and on the market. Both elements implies a radical change in equilibrium outcomes compared to the standard models.

On the other side, our paper builds on the strand of literature that focuses on two-part tariffs. Since the seminal paper by Oi (1971), two-part pricing contracts are known for allowing a monopolist to sell goods at marginal cost, but to still extract consumer surplus in the form of an up-front payment. There is a large literature on price discrimination, including two-part tariffs, in oligopoly markets, with early contributions going back to Stole (1996), Armstrong and Vickers (2001), and Yin (2004). These models focus on two-part pricing at consumer level and do not deal with the specificities of pharmaceutical markets.[3] Lakdawalla and Sood (2009, 2013) consider two-part pricing in health care markets. Health insurance, both public and private, implies that consumers pay up-front premiums in exchange for lower unit prices (co-payments) in the event of illness. Assuming that insured consumers' demand for drugs is a function of the co-payments (in addition to their health status and income), they show that health insurance eliminates the deadweight loss from market power in health care provision, including pharmaceuticals. Our paper focuses on a different issue, namely two-part pricing at health plan level instead of consumer level. This topic is conceptually different, and allows us to study the game between drug producers and health plans in terms of market access and inclusion of drugs in the health plan. To the best of our knowledge, this represents the first attempt to derive properties of two-part tariffs, when applied by health plans in paying for pharmaceuticals. Our aim is to study the effect on two-part pricing on the health plan's access decision and the corresponding distribution of surpluses across drug producers and the health plan. This is a highly relevant issue both for public and private health plans.

The remainder of the paper is structured as follows. In Section 2 we present the model. In Section 3 we compare uniform pricing and two-part tariffs in the case where the market is served by a monopolist. Therapeutic competition is then introduced in Section 4, where we make the same comparison of pharmaceutical pricing schemes when the market is (potentially) served by the producers of two therapeutically substitutable drugs. In Section 5 we extend the analysis in

[3] In a setting where different producers sell to a common retailer, potential pro-competitive effects of two-part tariffs have been identified by Gabrielsen and Sørgard (1998) and Cachon and Kök (2010). These effects are similar in nature to the ones identified in the present analysis, although the institutional setting is very different.

three different directions. First, in the case of therapeutic competition we include the possibility of exclusive contracting by the health plan. Second, we check whether our main results are robust to an alternative price determination process, in which drug prices result from Nash bargaining between the health plan and the drug producer(s). Finally, we analyse the implications of alternative payment schemes for drug innovation incentives. Section 6 provides a discussion and some concluding remarks.

2 Model

Consider a therapeutic market where either one or two on-patent prescription drugs are available for patients. If there is only one drug in the market, it is located at one of the endpoints of a unit line. If there are two drugs in the market, they are located at different endpoints of the same line; drug 1 at the left endpoint and drug 2 at the right endpoint. Patients are uniformly distributed on the line with total patient mass equal to one. We can interpret the unit line as a 'disease space' where a patient's therapeutic benefit of a particular drug is higher the closer the patient is located to the drug, all else equal. Thus, the distance between the location of a particular drug and the location of a particular patient reflects the degree of therapeutic mismatch between the two.

Each patient attends a physician who prescribes what is considered the most appropriate treatment for the patient, which is either one unit of a drug from the available choice set, or no drug treatment. When making this decision, the prescribing physician takes into account both the patient's health benefit and the price(s) of the drug(s). More specifically, let $v_i > 0$, $p_i > 0$ and $z_i \in \{0, 1\}$ denote the quality, price and therapeutic location, respectively, of drug i, where $i = 1, 2$. If one unit of this drug is prescribed to a patient located at $x \in [0, 1]$, the utility assigned to this choice by the prescribing physician is

$$u_i(x) = v_i - t|x - z_i| - \beta p_i. \tag{1}$$

For each patient, the physician will prescribe the drug that yields the highest utility, as specified by (1), but only if this utility is non-negative. Otherwise, no drug treatment is given.

The utility function given by (1) consists of two components. The first component is the patient's health benefit of being prescribed drug i, which is given by $v_i - t|x - z_i|$. The health benefit depends on the quality of the drug and the therapeutic match between the patient and the drug, where the

relative importance of the latter is reflected by the mismatch cost parameter $t > 0$.[4] Notice that, in case of two available drugs in the prescription choice set, t also measures (inversely) the degree of therapeutic substitutability, and thus the intensity of therapeutic competition, between the two drugs.[5]

The second component in the utility function reflects the cost of drug treatment. The parameter $\beta \in \left(\frac{1}{2}, 1\right]$ measures how sensitive the physician's prescription decision is to drug prices. The physician is an agent both for the patient and the third-party purchaser (i.e., the health plan). If $\beta = 1$, the physician takes drug prices fully into account and acts as a perfect agent for the health plan that maximises total health benefits net of purchasing costs. However, in the more general case of $\beta < 1$, the influence of patient agency makes health benefits more important relative to drug prices for the prescribing physician. Notice that our interpretation of β is sufficiently general to incorporate patient co-payments, where a higher co-payment rate implies a higher value of β.[6] Although physicians who prescribe drugs covered by health plans are aware of lower, or even zero, patient copayment, their choices also reflect health plans' prescription policies and cost control strategies (see, e.g., Danzon and Chao (2000)).[7] In order to ensure equilibrium existence throughout our analysis, we impose a lower bound on β equal to one half.[8]

Each drug is produced by a profit-maximising firm. The payment for drug i includes the per-unit price p_i and potentially also a fixed fee f_i, depending on the type of payment contracts used. Assuming a constant marginal cost c of drug production, equal for both drugs, the profit of producer i is given by

$$\pi_i = (p_i - c)\, y_i + f_i, \tag{2}$$

where y_i is the demand for drug i, which is derived from drug prescription decisions that maximise

[4] The mismatch cost can be interpreted as the subjective effectiveness of the drug therapy for a given patient, including also possible side-effects of the drug that are specific to the patient.

[5] There is a large literature on competition from generic and therapeutic alterantives. See Ellison et al. (1997) for an early contribution on this topic and the literature review on pharmaceutical markets by Scott Morton and Kyle (2011).

[6] Consider a patient located at x who is prescribed drug i and pays σp_i, where $\sigma \in (0, 1)$ is the copayment rate. The utility associated with this prescription choice is $v_i - \sigma p_i - t\,|x - z_i|$ from a patient perspective and $v_i - p_i - t\,|x - z_i|$ from a third-party purchaser perspective. If the prescribing physician maximises a weighted average of patient utility and purchaser utility, with a weight α given to the latter, the resulting physician payoff function is identical to (1) for $\beta := \alpha\,(1 - \sigma) + \sigma$, implying that β is increasing in the copayment rate (σ) and in the weight given to purchaser utility (α). If the prescribing physician places equal weights on patient and purchaser utility, i.e., $\alpha = \frac{1}{2}$, then $\beta = \frac{1}{2}\,(1 + \sigma) \in \left(\frac{1}{2}, 1\right)$.

[7] There is also evidence that physicians do take into account patients' out-of-pocket expenses when making prescription choices; see e.g., Carrera et al. (2018).

[8] More specifically, local stability of the Nash equilibrium under therapeutic competition with two-part tariffs requires $\beta > 1/2$.

(1) for each patient.

The available number of drugs in the market is determined by a monopoly purchaser (health plan) who decides whether to include one or both of the drugs in its health plan. The objective of the health plan is to maximise its surplus, defined as total health benefits to patients net of drug expenditures. If only drug i is included in the plan, the surplus is given by

$$S_i = H_i - p_i y_i - f_i, \tag{3}$$

where

$$H_i = \int_0^{y_i} (v_i - tx)\, dx = \left(v_i - \frac{t y_i}{2}\right) y_i \tag{4}$$

is the total health benefit of drug i, and where the market is partially (fully) covered if $y_i < (=) 1$. If both drugs are included, and there is viable competition between them, implying that the market is fully covered, the surplus is

$$S_{12} = H_{12} - p_1 y_1 - p_2 (1 - y_1) - f_1 - f_2, \tag{5}$$

where

$$H_{12} = \int_0^{y_1} (v_1 - tx)\, dx + \int_{y_1}^1 (v_2 - t(1 - x))\, dx = v_1 y_1 + v_2 (1 - y_1) - \frac{t}{2}(1 - 2(1 - y_1) y_1) \tag{6}$$

is the total health benefit of including both drugs in the health plan.

Finally, we impose two parameter restrictions:

A1 $\min\{v_1, v_2\} > c + t$.

A2 $|v_1 - v_2| < 3t$

The first assumption states that the net health benefit (health benefit minus production cost) of each drug is positive for all patients in the market. This implies that the socially efficient outcome is a fully covered market, where each patient is given drug treatment, regardless of whether the prescription choice set consists of one or two drugs. Another implication of this assumption is that there is a monotonic relationship between total health benefit and total welfare, defined as total health benefit net of production costs. The second assumption applies to the case of therapeutic

competition analysed in Section 4, ensuring that the quality difference between the two drugs is sufficiently low to ensure equilibrium existence under both payment schemes considered, when both drugs are included in the health plan. Given the above assumptions, we consider the following game:

1. The drug producers simultaneously and non-cooperatively submit bids (p_i, f_i).
2. The purchaser decides whether to include one or both of the drugs in the health plan (or none of the drugs if a positive surplus cannot be achieved).
3. Each patient is prescribed a drug from the available choice set (or no prescription if drug treatment does not yield a positive utility).

We will consider two different versions of this game, where the payment scheme is either based on uniform pricing ($f_i = 0$) or two-part tariffs ($f_i > 0$).[9] As usual, the game is solved by backwards induction to find the subgame perfect Nash equilibrium.

3 Monopoly

We start out by considering the case of a monopoly market, where only one drug exists. Alternatively, we can interpret this case as the quality difference between the two drugs being so large that therapeutic competition is infeasible, effectively turning the market into monopoly for the high-quality drug. Given that the existing (high-quality) drug is located at one of the endpoints of the unit line, and assuming that the drug is prescribed to all patients for which the utility given by (1) is non-negative, total demand for the drug is given by[10]

$$y = \min\left\{\frac{v - \beta p}{t}, 1\right\}. \tag{7}$$

The monopoly version of the game described towards the end of the last section is simply that the producer makes a take-it-or-leave-it offer which the purchaser can either accept or reject. Thus, the monopoly producer's problem is to maximise profits under the constraint that the purchaser's surplus, which in this case is given by (3), is non-negative. In the following we will solve this

[9] By assuming that drug producers make take-it-or-leave-it offers, we implicitly assume that the producers have all the bargaining power vis-à-vis the purchaser. This seems a reasonable approximation in on-patent pharmaceutical markets where a patent holder sells its drug to a large number of separate purchasers (markets). An alternative assumption would be to consider Nash bargaining between producers and purchasers. In Section 5 we check the robustness of our main results to such an alternative price determination process.

[10] To ease notation, we drop the subscript i in the monopoly case.

maximisation problem under two different assumptions about the price offer, namely that (i) the producer offers a uniform price (p), or (ii) the producer offers a two-part tariff (p, f).

3.1 Uniform pricing

Given that the drug is included in the health plan, the profit-maximising price is either an interior solution where $y < 1$, or a corner solution where the price is set such that the prescribing physician is indifferent between prescribing or not the drug to the patient with the lowest health benefit; i.e., $u(1) = 0$. The producer's choice of whether to implement an interior solution or not depends generally on the quality of the drug and on the price sensitivity of the prescription decision. However, inclusion in the health plan is not guaranteed at any price. The producer is therefore also constrained by the condition that the offered price must give the health plan a non-negative surplus. If this condition binds, the profit-maximising price is implicitly given by $S = 0$, where S is defined by (3) for $f = 0$.

When considering both types of potential corner solutions, stemming from the prescription decisions and from the participation constraint of the purchaser, it can be shown (see Appendix A for details) that the optimal price is given by[11]

$$p^M = \begin{cases} \frac{v+\beta c}{2\beta} & if \quad v \leq 2t + \beta c \text{ and } \beta > \beta^*, \\ \frac{v}{2-\beta} & if \quad v \leq 2t + \beta c \text{ and } \beta^{**} < \beta \leq \beta^*, \\ \frac{v-t}{\beta} & if \quad v > 2t + \beta c \text{ and } \beta > \beta^{**}, \\ v - \frac{t}{2} & if \quad \beta \leq \beta^{**}, \end{cases} \tag{8}$$

where $\beta^*, \beta^{**} \in \left(\frac{1}{2}, 1\right)$. The corresponding demand for the drug is given by

$$y^M = \begin{cases} \frac{v-\beta c}{2t} & if \quad v \leq 2t + \beta c \text{ and } \beta > \beta^*, \\ \frac{2(1-\beta)v}{(2-\beta)t} & if \quad v \leq 2t + \beta c \text{ and } \beta^{**} < \beta \leq \beta^*, \\ 1 & if \quad v > 2t + \beta c \text{ or } \beta \leq \beta^{**}. \end{cases} \tag{9}$$

Given that the drug is included in the health plan, the profit-maximising price is an interior solution, in which not all patients are prescribed the drug, if the drug quality is sufficiently low ($v < 2t + \beta c$). Intuitively, the scope for such an interior solution is larger if drug demand is more price sensitive or if mismatch costs are larger. However, unless demand is sufficiently price

[11] We use superscript M to denote equilibrium values under monopoly.

sensitive, the unconstrained price will be too high to yield a non-negative surplus for the health plan. In this case, which occurs if $\beta \leq \beta^*$, the optimal price must be adjusted downwards in order to satisfy the participation constraint of the purchaser. Such a corner solution can imply either partial or full market coverage, with full coverage occurring if β is sufficiently low ($\beta \leq \beta^{**}$). Thus, the monopoly solution under uniform pricing implies full market coverage if (i) the quality of the drug is sufficiently high ($v > 2t + \beta c$), *or* if (ii) the price sensitivity of demand is sufficiently low ($\beta \leq \beta^{**}$). If neither of these two conditions are met, the monopoly solution implies partial market coverage.

The producer profits and health plan surplus in the monopoly solution are given by, respectively,

$$\pi^M = \begin{cases} \frac{(v-\beta c)^2}{4t\beta} & if \quad v \leq 2t + \beta c \text{ and } \beta > \beta^* \\ \frac{2v(1-\beta)(v-(2-\beta)c)}{t(2-\beta)^2} & if \quad v \leq 2t + \beta c \text{ and } \beta^{**} < \beta \leq \beta^* \\ \frac{v-t}{\beta} - c & if \quad v > 2t + \beta c \text{ and } \beta > \beta^{**} \\ v - c - \frac{t}{2} & if \quad \beta \leq \beta^{**} \end{cases} \tag{10}$$

and

$$S^M = \begin{cases} \frac{(v(3\beta-2)-(2-\beta)\beta c)(v-\beta c)}{8t\beta} & if \quad v \leq 2t + \beta c \text{ and } \beta > \beta^* \\ 0 & if \quad v \leq 2t + \beta c \text{ and } \beta^{**} < \beta \leq \beta^* \\ \frac{(2-\beta)t-2(1-\beta)v}{2\beta} & if \quad v > 2t + \beta c \text{ and } \beta > \beta^{**} \\ 0 & if \quad \beta \leq \beta^{**} \end{cases} . \tag{11}$$

Consistent with the above discussion of the producer's optimal pricing incentives, we see that the monopoly solution leaves the health plan with zero surplus if the price sensitivity of drug demand is sufficiently low, such that the participation constraint of the purchaser binds (this case includes two regimes, where the market is either partially or fully covered). Thus, complete surplus extraction through uniform pricing might happen only if the health plan's cost of purchasing the drug is not fully internalised by the prescribing physician (i.e., if $\beta < 1$), making drug demand less price elastic. On the other hand, if the purchasing cost is fully internalised ($\beta = 1$), we obtain the standard result that the monopolist's profits are maximised at a price that leaves the purchaser with a positive surplus, regardless of whether this price yields an interior solution ($y^M < 1$) or not ($y^M = 1$).[12]

We summarise the above analysis as follows:

[12] Notice that full market coverage ($y^M = 1$) does not necessarily imply full surplus extraction ($S^M = 0$), nor does partial market coverage ($y^M < 1$) necessarily imply that the purchaser is left with a positive surplus ($S^M > 0$).

Proposition 1 *Under monopoly with uniform pricing, (i) the market is fully covered if the quality of the drug is sufficiently high or if the price sensitivity of demand is sufficiently low. Otherwise, the market is partially covered. (ii) If the price sensitivity of drug demand is sufficiently low, the health plan is left with zero surplus.*

3.2 Two-part tariffs

With two-part tariffs, the producer can extract the purchaser's entire surplus through the fixed fee f. From (3), this implies that, for any unit price p, the optimal fixed fee is given by

$$f = \left(v - p - \frac{ty}{2}\right) y. \tag{12}$$

The producer's profits can therefore be expressed as

$$\pi = (p - c)\, y + \left(v - p - \frac{ty}{2}\right) y = \left(v - c - \frac{ty}{2}\right) y. \tag{13}$$

For a given demand, a reduction in the unit price increases the health plan's surplus by the same amount as it reduces the firm's profits, implying that the profit loss of a unit price reduction can be fully recaptured by increasing the fixed fee. Thus, with two-part tariffs, the unit price affects profits only to the extent that it affects total demand, and it is optimally set to induce the level of demand that maximises the health plan's surplus. This demand level is clearly $y = 1$, given the assumption that the health benefits net of production costs are positive for all patients, i.e., $v > c + t$. Thus, the maximum profit under two-part tariffs is therefore

$$\widehat{\pi}^M = v - c - \frac{t}{2}, \tag{14}$$

which is independent of the unit price p as long as this price is low enough to induce a fully covered market. The characterisation of the optimal two-part tariff follows straightforwardly:

Proposition 2 *Under monopoly with two-part tariffs, there is a continuum of payoff-equivalent contracts characterised by*

$$\widehat{p}^M \leq \frac{v - t}{\beta} \tag{15}$$

and

$$\widehat{f}^M = v - \widehat{p}^M - \frac{t}{2}. \tag{16}$$

For any of these contracts, the producer's profits are given by (14) and the purchaser's surplus is zero.

Let us now compare the outcomes under the two different payment schemes. Whereas a socially efficient outcome (i.e., $y = 1$) is always guaranteed with a two-part tariff, uniform pricing produces an efficient outcome only if the drug quality is sufficiently high or if the price responsiveness of drug demand is sufficiently low. However, from the viewpoint of the purchaser, uniform pricing is strictly preferred to a two-part tariff if β is sufficiently high. Even if uniform pricing yields a worse health outcome, in the sense that some patients are not given the drug treatment because of an excessively high unit price, the additional health benefits obtained by the use of a two-part tariff are lower than the corresponding increase in total drug payments.

The producer, on the other hand, is obviously at least as well off with a two-part tariff, since this type of payment contract provides an additional instrument to extract surplus from the purchaser. However, if the price responsiveness of drug demand is sufficiently low, $\beta < \beta^{**}$, both the producer and the purchaser are indifferent between the two payment schemes.

The gains and losses from the two payment schemes under monopoly can be characterised as follows:

Proposition 3 *In case of monopoly, two-part tariffs yield weakly higher producer profits, health benefits and total welfare, but a weakly lower surplus for the health plan, than uniform pricing.*

4 Therapeutic competition

Consider now the case where two therapeutically substitutable drugs exist (or, alternatively, that the quality difference between the two drugs is sufficiently low to make therapeutic competition feasible). If both drugs are included in the health plan, and if prescription choices are made to maximise (1) for every patient, demand for drug i is given by

$$y_i = \frac{1}{2} + \frac{\Delta v - \beta \Delta p}{2t}, \tag{17}$$

$i, j = 1, 2$, $i \neq j$, where $\Delta v := v_i - v_j$ is the quality difference and $\Delta p := p_i - p_j$ is the price difference between drug i and drug j. An interior solution with positive demand for both drugs requires that $\Delta v < \beta \Delta p + t$.

In the following we will once more consider the two cases of uniform pricing and two-part tariffs, now solving the full game outlined in Section 2 with price competition between the two drug producers at the first stage of the game. Throughout the analysis we assume that the drug qualities are sufficiently high to make therapeutic competition viable, implying that the market is fully covered in equilibrium.

4.1 Uniform pricing

Suppose that the payment contracts are linear, such that $f_i = 0$, $i = 1, 2$. Consider first the decision made by the purchaser at the second stage of the game. Suppose that p_i and p_j are such that $S_i \geq S_j \geq 0$, implying that it is always optimal to include at least drug i in the health plan. In this case, by substituting (17) into (3)-(5) and setting $f_i = 0$ for $i = 1, 2$, we find that both drugs will be included if

$$S_{12} - S_i = \frac{\left((2 - \beta)\,\Delta p + t - \Delta v\right)\left(\beta \Delta p + t - \Delta v\right)}{4t} \geq 0, \tag{18}$$

which is true for $\Delta v < \beta \Delta p + t$. Thus, if p_i and p_j are such that both drugs have positive demand if they belong to the available prescription choice set, including the second drug will always increase the surplus of the health plan, since utility is increased for the patients who are prescribed the second drug.

We now turn to the first stage of the game, where each producer sets a uniform price under the anticipation that both drugs will be included in the health plan. By substituting (17) into (2) and maximising with respect to p_i, we derive the best-response function of producer i, which is given by[13]

$$p_i\left(p_j\right) = \frac{1}{2}\left(c + p_j + \frac{t + \Delta v}{\beta}\right). \tag{19}$$

As expected, the prices are strategic complements. A higher price set by producer j shifts demand

[13] The second-order conditions of the profit-maximisation problems are satisfied, since

$$\frac{\partial^2 \pi_i}{\partial p_i^2} = -\frac{\beta}{t} < 0.$$

towards drug i, which consequently makes the demand for drug i less price elastic, and producer i optimally responds by increasing the price.

The Nash equilibrium at the price bidding stage is given by[14,15]

$$p_i^D = c + \frac{t}{\beta} + \frac{\Delta v}{3\beta}, \tag{20}$$

which implies

$$\Delta p^D = \frac{2\Delta v}{3\beta} \tag{21}$$

and

$$y_i^D = \frac{1}{2} + \frac{\Delta v}{6t}. \tag{22}$$

We see that Assumption A2 ensures $y_i > 0$ for $i = 1, 2$. The remaining properties of this equilibrium are straightforward and intuitive. The price level of each drug is increasing in the marginal production cost and in the quality of the drug, while decreasing in the degree of therapeutic substitutability and in the price responsiveness of demand. Furthermore, the producer of the high-quality drug charges the higher price.

The above derived equilibrium outcome ensures that the condition in (18) is satisfied, which implies that either both or no drugs will be included in the health plan. If both drugs are included, the profits of the drug producers are

$$\pi_i^D = \frac{(3t + \Delta v)^2}{18t\beta}, \tag{23}$$

implying $\pi_i^D > \pi_j^D$ if $\Delta v > 0$. The total health benefit in this equilibrium is given by

$$H_{12}^D = \overline{v} - \frac{t}{4} + \frac{5\,(\Delta v)^2}{36t}, \tag{24}$$

and the total surplus of the health plan is

$$S_{12}^D = \overline{v} - c - \frac{(4+\beta)\,t}{4\beta} - \frac{(4-5\beta)\,(\Delta v)^2}{36t\beta}, \tag{25}$$

[14] The Nash equillibrium is (locally) stable, since

$$\frac{\partial^2 \pi_1}{\partial p_1^2}\frac{\partial^2 \pi_2}{\partial p_2^2} - \frac{\partial \pi_1}{\partial p_1 \partial p_2}\frac{\partial \pi_2}{\partial p_1 \partial p_2} = \frac{3\beta^2}{4t^2} > 0.$$

[15] We use superscript D (for duopoly) to denote equilibrium values under therapeutic competition.

where $\overline{v} := (v_1 + v_2)/2$ is the average drug quality. The final condition needed for both drugs to be included in the health plan is that $S_{12}^D \geq 0$, which requires that the quality of each drug is sufficiently high (see Appendix B for a derivation of a specific sufficient condition for $S_{12}^D \geq 0$). We assume that this condition is satisfied.

The equilibrium outcome under therapeutic competition with uniform pricing is summarised as follows:

Proposition 4 *Suppose that the producers of two therapeutically substitutable drugs compete in uniform prices. In this case, both drugs are included in the health plan if $\Delta v < 3t$, and the producer of the high-quality drug charges a higher price than the producer of the low-quality drug.*

4.2 Two-part tariffs

Suppose now that each of the drug producers bids a unit price p_i and a fixed fee f_i at the first stage of the game. Similarly to the monopoly case, this means that, for each producer, the additional surplus created by making drug i available at price p_i can be captured by the fixed fee f_i. However, this surplus extraction is conditional on the drug being included in the health plan. As we will show below, there are two different competition regimes arising under two-part tariffs, one in which both drugs are included in the plan, leading to *de facto therapeutic competition*, and one in which only one of the drugs is included, leading to a situation of *potential therapeutic competition*, where the non-included drug places a competitive pressure on the included one. The third possible outcome, where therapeutic competition is blockaded in the sense that it is not profitable for a rival with a therapeutic substitute to submit a bid (due to large quality differences), is captured by the monopoly case in the previous section.

4.2.1 *De facto* therapeutic competition

Suppose that drug j is included in the health plan. In this case, drug i will also be included if $S_{12} \geq S_j$. By substituting (17) into (3)-(5), the inclusion criterion for drug i is given by

$$\overline{v} - \overline{p} - \frac{t}{4} + \frac{(\Delta v - \beta \Delta p)(\Delta v - (2-\beta)\Delta p)}{4t} - f_i - f_j \geq v_j - p_j - \frac{t}{2} - f_j, \tag{26}$$

where $\overline{p} := (p_1 + p_2)/2$ is the average unit price. This condition can be re-formulated as

$$f_i \leq \widehat{f}_i := \frac{(t + \Delta v - (2-\beta)\Delta p)(t + \Delta v - \beta \Delta p)}{4t}. \tag{27}$$

Since a higher f_i implies a non-distortionary transfer from the buyer to the seller, the inequality in (27) will bind in equilibrium. Thus, given that drug j is included in the health plan, the producer of drug i can extract the additional surplus created by its inclusion through the fixed fee.

Setting $f_i = \widehat{f}_i$ in the profit function of producer i and maximising (2) with respect to the unit price p_i yields the following first-order condition for an interior solution:[16]

$$\frac{\partial \pi_i}{\partial p_i} = -\frac{\beta}{2t}(\beta p_i + (1-\beta) p_j - c) = 0. \tag{28}$$

The best-response functions are therefore given by

$$p_i(p_j) = \frac{c}{\beta} - \left(\frac{1-\beta}{\beta}\right) p_j, \quad i, j = 1, 2, \quad i \neq j. \tag{29}$$

There are two important observations that can be made here. First, we see that the nature of strategic interaction is fundamentally changed when the producers set two-part tariffs. Contrary to the case of uniform pricing, with two-part tariffs the unit prices are *strategic substitutes* at the price bidding stage (i.e., $\partial p_i / \partial p_j < 0$). In order to grasp the intuition behind this result, it is instructive to write the first-order condition (28) on a more general form, as

$$\frac{\partial \pi_i}{\partial p_i} = \frac{\partial ((p_i - c) y_i)}{\partial p_i} + \frac{\partial \widehat{f}_i}{\partial p_i} = 0. \tag{30}$$

Under two-part tariffs, a price increase affects variable profits ($(p_i - c) y_i$), as it does under uniform pricing, but it also affects the fixed fee that the firm can charge. More precisely, a higher price for drug i reduces the additional surplus generated by inclusion of the drug in the health plan, thus implying that the producer has to reduce the fixed fee in order to ensure inclusion. How do these marginal profit effects depend on the price level of the competing drug j? An increase in p_j shifts demand from drug j to drug i, which implies that the demand for drug i becomes less price elastic.

[16] The second-order conditions of the profit-maximisation problems are satisfied, since

$$\frac{\partial^2 \pi_i}{\partial p_i^2} = -\frac{\beta^2}{2t} < 0.$$

All else equal, this increases the first term in (30) and gives producer i an incentive to set a higher price. This is the only strategic effect under uniform pricing, leading to strategic complementarity between the prices. However, under two-part tariffs, the price of drug j also affects the magnitude of the second term in (30), which can be written as

$$\frac{\partial \widehat{f}_i}{\partial p_i} = \left(2ty_i - (1-\beta)\,\Delta p\right)\frac{\partial y_i}{\partial p_i} - (1-\beta)\,y_i < 0. \tag{31}$$

Consider again an increase in p_j , which leads to higher demand for drug i (i.e., $\partial y_i/\partial p_j > 0$). This demand increase implies that a marginal price increase for drug i will now affect a larger number of patients and therefore lead to a larger reduction in the additional surplus created by the inclusion of drug i in the health plan, leading in turn to a larger drop in the maximum fixed fee that producer i can charge. All else equal, this gives producer i an incentive to *reduce* the unit price. It turns out that the negative effect of p_j on the second term in (30) outweighs the positive effect on the first term, leading to strategic substitutability between the unit prices set by the two producers.

The second important observation we can make from the set of best-response functions, (29), is that they are completely symmetric and thus not dependent on the quality difference between the two drugs. This implies that the Nash equilibrium is also symmetric in terms of unit prices (but not in terms of fixed fees), and given by marginal cost pricing for both producers:[17]

$$\widehat{p}_i^{\,D} = c, \quad i = 1, 2. \tag{32}$$

The intuition follows from the explanation for strategic substitutability, as outlined above. For any pair of unit prices higher than marginal cost, an increase in variable profits due to a higher price is more than offset by a reduction in the fixed fee, regardless of the quality differences between the drugs, thus implying that neither of the drug producers have any incentives to set their unit prices above marginal cost. This means that profits are fully extracted through the fixed fees. Equilibrium profits (and fixed fees) are therefore given by

$$\widehat{\pi}_i^{\,D} = \widehat{f}_i^{\,D} = \frac{(t+\Delta v)^2}{4t}, \quad i, j = 1, 2, \quad i \neq j, \tag{33}$$

[17] The condition for (local) stability of the Nash equilibrium is

$$\frac{\partial^2 \pi_1}{\partial p_1^2}\frac{\partial^2 \pi_2}{\partial p_2^2} - \frac{\partial^2 \pi_1}{\partial p_1 \partial p_2}\frac{\partial^2 \pi_2}{\partial p_1 \partial p_2} = \frac{\beta^2\,(2\beta - 1)}{4t^2} > 0,$$

which holds for $\beta > 1/2$.

implying that the high-quality producer charges a higher fee and consequently earns a higher profit in equilibrium.

With marginal cost pricing, equilibrium demand is given by

$$\widehat{y}_i^D = \frac{1}{2} + \frac{\Delta v}{2t}, \quad i, j = 1, 2, \quad i \neq j. \tag{34}$$

An interior solution therefore requires that the drug quality difference is sufficiently low: $\Delta v < t$. Notice that this is *not* guaranteed for all parameter configurations covered by Assumption A2. Under the condition that $\Delta v < t$, both drugs will be included in the health plan, and the total health benefits are

$$\widehat{H}_{12}^D = \overline{v} - \frac{t}{4} + \frac{(\Delta v)^2}{4t}, \tag{35}$$

whereas the total surplus of the health plan is

$$\widehat{S}_{12}^D = \overline{v} - c - \frac{3t}{4} - \frac{(\Delta v)^2}{4t}. \tag{36}$$

It is easily confirmed that $\widehat{S}_{12}^D > 0$ for all parameter values satisfying Assumption A1.

4.2.2 Potential therapeutic competition

Consider now the case of $t < \Delta v < 3t$, for which *de facto* therapeutic competition is not feasible, since the quality difference is too large for the low-quality drug to obtain positive demand, given the contract offered by the high-quality producer in the candidate equilibrium. However, this does not allow the high-quality producer to extract all surplus with a two-part tariff, as in the monopoly case, because the purchaser can credibly threaten to replace the high-quality drug with the low-quality drug in the health plan. Thus, if drug i is the high-quality drug, the bid submitted by producer i is constrained by the condition that the health plan's surplus obtained by including only drug i must be at least as high as the surplus obtained if drug i is replaced by drug j in the health plan. Under the assumption of full market coverage, this condition is given by

$$v_i - p_i - \frac{t}{2} - f_i \geq v_j - p_j - \frac{t}{2} - f_j, \tag{37}$$

which, using our previously defined notation, can be re-written as

$$f_i \leq \Delta v - \Delta p + f_j. \tag{38}$$

All else equal, the most profitable contract offer for each producer is a contract that maximises total surplus and extracts profits only through the fixed fee, which implies that the unit price is set equal to marginal cost. This implies that the optimal contract offer by producer i is a contract given by $p_i = c$ and f_i, where f_i is just low enough to make the condition in (37) hold for all weakly profitable contract offers by producer j. In the latter set of contracts, the one that maximises the surplus of the health plan is a contract with $p_j = c$ and $f_j = 0$. Based on (38), this implies that the optimal contract offer by producer i has $f_i = \Delta v$.

Thus, for $t < \Delta v < 3t$, only the high-quality drug is included in the health plan, and the equilibrium two-part tariff is given by[18]

$$\widehat{p}_i^d = c \text{ and } \widehat{f}_i^d = \Delta v. \tag{39}$$

Equilibrium profits, total health benefit and total surplus for the health plan are given by, respectively,

$$\widehat{\pi}_i^d = \Delta v, \tag{40}$$

$$\widehat{H}_i^d = v_i - \frac{t}{2} \tag{41}$$

and

$$\widehat{S}_i^d = v_i - c - \frac{t}{2} - \Delta v = v_j - c - \frac{t}{2} > 0. \tag{42}$$

Thus, the equilibrium profits of the high-quality producer are given by the quality difference between the two drugs, whereas the equilibrium surplus of the health plan is given by the maximum value of the purchaser's outside option, i.e., the maximum surplus that can be obtained by replacing drug i by drug j in the health plan.

The next proposition summarises the main insights from the analysis of therapeutic competition with two-part tariffs.

Proposition 5 *Suppose that the producers of two therapeutically substitutable drugs compete in*

[18] We use superscript d to denote equilibrium values under potential therapeutic competition.

two-part tariffs. In this case, both drugs are included in the health plan if the drug quality difference is sufficiently low, $\Delta v < t$, whereas only the high-quality drug is included if $t \leq \Delta v < 3t$. Regardless of whether one or both drugs are included in the health plan, the equilibrium two-part tariffs have unit prices equal to marginal cost.

4.3 Comparison of payment contracts under therapeutic competition

We conclude this section by comparing the equilibrium outcomes under the two different payment contracts. Who gains and who loses from the use of two-part tariffs instead of uniform pricing? The answer to this question is summarised in the following proposition (see Appendix C for a proof):

Proposition 6 *(i) Under de facto therapeutic competition ($\Delta v < t$), the surplus of the health plan is always higher with two-part tariffs than with uniform pricing. The same is true under potential therapeutic competition ($t \leq \Delta v < 3t$), unless the drug quality difference and the price sensitivity of demand are both sufficiently high.*

(ii) The profit of the low-quality producer is always higher with uniform pricing than with two-part tariffs, regardless of whether there is de facto or potential competition. The same is true for the profit of the high-quality producer, if the price sensitivity of demand is sufficiently low. On the other hand, if demand is sufficiently price sensitive, the profit of the high-quality producer is higher with two-part tariffs under de facto competition if the quality difference is sufficiently large, and it is always higher with two-part tariffs under potential competition.

(iii) The total health benefit (and thus total welfare) is always higher with two-part tariffs, regardless of whether there is de facto or potential competition.

The impact of therapeutic competition on the purchaser's preference ranking of the two payment schemes is quite striking. In complete contrast to the monopoly case, in which the monopoly supplier is able to extract all surplus by using a two-part tariff, the health plan obtains a higher surplus under therapeutic competition if the drug payments are based on two-part tariffs instead of uniform pricing. The only exception is the case in which the quality difference is high and drug demand is sensitive to price changes.

There are two different reasons why two-part tariffs are generally preferable to uniform pricing from the viewpoint of the purchaser. First, with therapeutic competition between drug suppliers, each supplier is only able to extract the additional surplus created by the inclusion of its drug in

the health plan. The purchaser's credible threat of excluding one of the drugs from the health plan implies that the producers' surplus extraction is lower than in the case of a monopoly producer. Thus, *competition for access* between the two producers ensures that a larger share of the surplus is captured by the purchaser. The second reason is that the use of two-part tariffs eliminates an allocational inefficiency that is present under uniform pricing, thereby creating a larger surplus. If the two producers compete by setting uniform prices, the high-quality producer exploits its demand advantage by setting a higher price than the low-quality producer. This creates an allocative inefficiency where some patients would have been better off being treated with the high-quality drug, but are instead being prescribed the low-quality drug because it is less expensive. In other words, too few patients are being prescribed the high-quality drug under uniform pricing. This inefficiency is eliminated under two-part tariffs, because of the producers' incentives to set unit prices equal to marginal cost, regardless of drug quality differences.

The allocative inefficiency caused by uniform pricing explains why the total health benefits are always higher under two-part tariffs. This is also true under *potential* therapeutic competition, even if the use of two-part tariffs instead of uniform pricing in this case implies a reduction in the prescription choice set from two to one drugs. However, remember that this case arises for $\Delta v > t$, which implies that all patients in the market obtains a higher health benefit if they are treated by the high-quality drug. This illustrates again the allocative inefficiency related to uniform pricing. For $\Delta v \in (t, 3t)$, some patients are treated with the therapeutically less appropriate drug under uniform pricing because of the price difference caused by this payment scheme.

Since the purchaser generally prefers two-part tariffs under therapeutic competition, the producers tend to prefer uniform pricing, again in contrast to the monopoly case. However, while this is always the case for the low-quality producer, the high-quality producer might also prefer two-part tariffs if demand is sufficiently price sensitive and if the quality difference is sufficiently large. The partially aligned interests of the purchaser and the high-quality producer can be explained by the above mentioned efficiency gain from using two-part tariffs, which increases in the quality difference between the drugs.

Overall, our analysis shows how the presence (or not) of therapeutic competition is of crucial importance for identifying the winners and losers from the two different payment schemes under consideration. Although the use of two-part tariffs always maximises the total surplus, and therefore also maximises total health benefits, the distribution of this surplus depends crucially on the

competition regime. Under monopoly, a switch from uniform pricing to two-part tariffs generally benefits drug producers at the expense of drug purchasers. In the presence of therapeutic competition, the opposite tends to be the case.

Our main results regarding payment scheme preferences are summarised in Table 1, where we indicate the preferred payment scheme – uniform pricing (UP) or two-part tariffs (2PT) – for each of the interested parties under different market structures.

Table 1: Overview of payment scheme preferences

	Monopoly	*Therapeutic competition*
Health plan	UP	2PT (unless β and Δv are high)
Producers	2PT	UP*
Patients	2PT	2PT
Total surplus	2PT	2PT

*If β and Δv are sufficiently high, the high-quality producer prefers 2PT

5 Extensions

In this section we extend our main analysis in three different directions. First, we consider the possibility of exclusive contracting. Second, we check whether our main results are robust to an alternative price determination process, in which drug prices result from Nash bargaining between the health plan and the drug producer(s). Finally, we analyse and discuss the implications of alternative payment schemes for drug innovation incentives. We summarise our results for each extension in Propositions.

5.1 Exclusive contracting

We now extend our analysis of therapeutic competition by considering an alternative version of the game, where the two producers compete for an *exclusive contract* with the health plan. This implies a modification of the rules at stage two of the game outlined in Section 2, where we now assume that the purchaser only includes the drug that yields the highest surplus for the health plan. Thus, drug i will obtain exclusivity if $S_i \geq S_j$. Clearly, this condition holds in equilibrium only if drug i is the high-quality drug, i.e., if $v_i \geq v_j$. In this case, the producer of drug i can

obtain exclusivity by making a bid that outperforms all profitable bids by the producer of drug j. The unique Nash equilibrium is then that producer j makes a bid that yields zero profits if the bid is accepted, while producer i makes a bid that is marginally more favourable for the health plan. Assuming a tie-breaking rule where producer i wins the contest if both bids yield equal surplus for the health plan, the equilibrium bid by producer i is implicitly given by

$$S_i = S_j|_{\pi_j=0}\,. \tag{43}$$

If the two producers compete for exclusivity by using *two-part tariffs*, the game is identical to the one described under *potential therapeutic competition* in Section 4.2.2, and the winning bid is therefore given by (39), which gives the health plan a surplus given by (42). On the other hand, if the producers compete in *uniform prices*, the bid that gives zero profits for producer j is $p_j = c$. Using (3)-(4), and assuming that v_i is large enough to ensure full market coverage in equilibrium, the equilibrium condition (43) under uniform pricing then becomes

$$v_i - p_i - \frac{t}{2} = v_j - c - \frac{t}{2}, \tag{44}$$

which implies that the winning bid is given by

$$p_i = c + \Delta v. \tag{45}$$

Since total demand is one, this gives producer i a profit of Δv, which is exactly the same as under competition for exclusivity in two-part tariffs. The surplus of the health plan is therefore also the same, and given by (42). Thus:

Proposition 7 *If the health plan offers an exclusive contract, the equilibrium outcome is the same regardless of whether the producers compete for the contract by uniform or two-part pricing. In both cases, the high-quality producer wins the contract and obtains a profit equivalent to the quality difference between the drugs, while the health plan obtains a surplus equal to the maximum value of its outside option.*

Would the health plan benefit from offering an exclusive contract? And would such a contract lead to a better or worse health outcome for patients? The answers to these questions are not *a priori* obvious, because exclusive contracting will generally affect both the purchasing costs,

through changes in the degree of competition between the producers, and the total health gain, through changes in mismatch costs. By making the relevant comparisons of health gains and purchaser surplus with and without contract exclusivity, we arrive at the following conclusions (proof in Appendix C):

Proposition 8 *(i) If drug payment is based on two-part tariffs, and if $\Delta v < t$, offering an exclusive contract always yields a lower total health benefit but a higher surplus for the health plan. For $\Delta v \in [t, 3t)$, exclusive contracting has no implications for the equilibrium outcome under two-part pricing.*

(ii) If drug payment is instead based on uniform pricing, non-exclusivity yields a higher surplus for the health plan if demand is sufficiently price sensitive and the drug quality difference is sufficiently large. Otherwise, the health plan would be better off by offering an exclusive contract. The total health benefit is higher (lower) with exclusive contracting if the drug quality difference is sufficiently large (small).

With two-part pricing, there are two counteracting effects. On the one hand, since competition on the market yields marginal cost pricing and therefore no allocative inefficiencies, exclusive contracting always leads to higher mismatch costs and therefore lower total health gains, because patients lose access to a valuable therapeutic substitute. On the other hand, letting the producers compete for an exclusive contract always intensifies the competition between them (as long as *de facto* therapeutic competition would ensue if both drugs are included in the plan) and therefore allows the health plan to reduce its purchasing costs. Perhaps surprisingly, the pro-competitive effect of exclusive contracting dominates for the entire (relevant) range of parameters. Although exclusivity leads to a worse health outcome, this is more than outweighed by a reduction in purchasing costs. In order to understand this result, notice that what increases the health loss of exclusive contracting also increases the amount of purchasing costs that can be saved by letting the producers compete for exclusivity. Consider an increase in the mismatch cost parameter t. This increases the health loss of excluding one of the drugs from the market, all else equal. However, a higher t also reduces competition between the producers on the market if both drugs are included in the health plan. Since the exclusive contract equilibrium does not depend on t, higher mismatch costs therefore increase the pro-competitive effect of contract exclusivity.

With *uniform pricing*, the effects are more complicated. In this case, exclusive contracting can

lead to either higher or lower health gains, and also higher or lower purchasing costs. The direction of both effects depends crucially on the quality difference between the two drugs. Consider first the effect on health gains. Since a larger quality difference increases the share of patients who are better off with the high-quality drug, the additional health benefit of having access to a second (lower-quality) drug is reduced. Furthermore, and in contrast to the case of two-part pricing, the allocative inefficiency created by uniform pricing (as discussed in Section 4.3) implies that, if the quality difference is sufficiently large, the total health gain is higher if only the high-quality drug is included in the health plan. Thus, exclusive contracting leads to an overall *increase* in health benefits for a sufficiently large quality difference between the drugs.[19] Consider next the effect of exclusive contracting on total purchasing costs. Once more, this effect depends crucially on the drug quality difference. Although a higher quality difference dampens the degree of competition on the market if both drugs are included, it leads to an even stronger reduction in the competition for an exclusive contract and, importantly, this difference is larger the more price sensitive drug demand is.[20] In fact, for sufficiently price sensitive demand *and* a sufficiently large quality difference, the purchasing costs are *higher* with exclusive contracting than under non-exclusivity. Overall, it turns out that, for sufficiently low quality differences, the negative health effect of exclusive contracting is more than outweighed by the reduction in purchasing costs, leading to an increase in the health plan's surplus. However, for sufficiently high quality differences, and if in addition demand is sufficiently price sensitive, the *positive* health effect of exclusive contracting is more than outweighed by the *increase* in purchasing costs, leading to a *reduction* in the surplus of the health plan.[21]

The possibility of offering exclusive contracts has interesting implications for the relative merits of uniform versus two-part pricing from the health plan's point of view. From Proposition 8 we know that the health plan always (weakly) prefers contract exclusivity under two-part pricing, but from Proposition 7 we know that, under contract exclusivity, the health plan is indifferent between

[19] An increase in patients' health benefits as a result of exclusive contracting occurs if $\Delta v > \frac{3}{5}t$. This implies that there exists a parameter set, given by $\Delta v \in \left(\frac{3}{5}t, t\right)$, in which having access to only the high-quality drug yields higher health benefits than having access to both drugs at distorted prices, even if the total health benefits would be maximised with a strictly positive share of patients being prescribed the low-quality drug. In other words, since too many patients are prescribed the low-quality drug under non-exclusivity with uniform pricing, total health benefits would increase by removing access to the low-quality drug.

[20] Notice that while more price sensitive demand (i.e., a higher β) intensifies price competition in the market duopoly, the exclusive contract equilibrium is unaffected by the price sensitivity of demand.

[21] The parameter space for which contract exclusivity leads to a lower surplus for the health plan under uniform pricing coincides with the parameter space in which the purchaser, under non-exclusivity, prefers uniform over two-part pricing. From Proposition 6 we know that this parameter space is a subset of $\Delta v \in (t, 3t)$.

uniform pricing and two-part tariffs. In other words, any benefit from exclusive contracting for the health plan can just as well be obtained using uniform pricing, which has the following immediate implication:

Corollary 1 *If offering exclusive contracts is a possibility, the health plan will always weakly prefer uniform pricing over two-part tariffs.*

Finally, it should be noted that the analysis of exclusive contracting in this section relies importantly on the assumption of a single, centralised drug procurer (i.e., health insurer). As pointed out in the concluding section, competition between health plans may alter the equilibrium properties of two-part tariffs and uniform prices, including the role played by exclusive contracts.

5.2 Bargaining

In the main analysis we have assumed that drug prices are determined by take-it-or-leave it offers from the drug producers. We now test the robustness of this assumption by considering an alternative price determination process, where drug prices result from Nash bargaining between the producers and the health plan. As in the main analysis, we will distinguish between the cases of monopoly and therapeutic competition. In order to facilitate the analysis, we make the simplifying assumption that, in case of therapeutic substitution, the drugs have equal quality; i.e., $\Delta v = 0$.

5.2.1 Monopoly

If there is only one available drug in the market, and if the health plan and the drug producer bargain over a *uniform price*, the solution can be either an interior or a corner solution where the market is partially or fully covered, respectively. Suppose first that the monopoly producer and the health plan bargain over a set of prices that will all result in a fully covered market ($y = 1$), which requires that v is sufficiently large. In this case, and assuming that the disagreement point is zero for both parties, the uniform price resulting from Nash bargaining is the price that maximises the Nash product

$$\Omega^M := S^{\alpha}\pi^{1-\alpha}, \tag{46}$$

where $\alpha \in (0,1)$ is the relative bargaining power of the health plan, and $S = v - (t/2) - p$ and $\pi = p - c$. The price that maximises Ω^M is

$$p^M(\alpha) = \alpha c + (1-\alpha)\left(v - \frac{t}{2}\right). \tag{47}$$

Thus, the uniform drug price under Nash bargaining is monotonically decreasing in the relative bargaining power of the health plan and ranges from marginal cost (if $\alpha = 1$) to the price that captures the entire surplus of the health plan (if $\alpha = 0$).[22] The corresponding profits are

$$\pi^M(\alpha) = (1-\alpha)\left(v - c - \frac{t}{2}\right), \tag{48}$$

while the health plan surplus is given by

$$S^M(\alpha) = \alpha\left(v - c - \frac{t}{2}\right). \tag{49}$$

Notice that, since we consider a parameter set for which total demand is completely inelastic, the total surplus is unaffected by the bargaining outcome (i.e., bargaining is efficient) and the two bargaining parties share this surplus according to their relative bargaining strength.

Suppose instead that the producer and the health plan bargain over a set of prices that all result in a partially covered market; i.e., $y < 1$. From our analysis in Section 3, we know that such an outcome requires a sufficiently price sensitive demand. For analytical convenience, let us therefore consider the special case of $\beta = 1$ (where the prescribing physician is a perfect agent for the health plan). In this case, the outcome of Nash bargaining is the uniform price that maximises Ω^M when $S = (v - ty/2 - p)\, y$ and $\pi = (p - c)\, y$, where $y = (v - p)/t$. This price is given by

$$\widetilde{p}^M(\alpha) = \frac{1+\alpha}{2}c + \frac{1-\alpha}{2}v \tag{50}$$

and yields an interior solution if drug quality is sufficiently low.[23] The corresponding profits and

[22] This price yields a corner solution with a fully covered market if

$$v > \frac{2c\alpha\beta + (2 - (1-\alpha)\beta)\, t}{2\alpha\beta + 2(1-\beta)}.$$

[23] More specifically,

$$v < c + \frac{2t}{1+\alpha}.$$

health plan surplus are given by

$$\widetilde{\pi}^M(\alpha) = \frac{\left(1-\alpha^2\right)(v-c)^2}{4t} \tag{51}$$

and

$$\widetilde{S}^M(\alpha) = \frac{(1+\alpha)^2(v-c)^2}{8t}, \tag{52}$$

respectively. Notice that bargaining is not efficient in this case, since the total surplus is monotonically increasing in the relative bargaining power of the health plan. When total demand is elastic, a lower price will increase demand and therefore increase the total surplus (under our basic assumption A1).

If the health plan and the monopoly producer instead bargain over a *two-part tariff*, the two parties will optimally agree on setting the unit price at a level which maximises the total surplus and bargain over the division of this surplus through the fixed fee. Thus, the outcome is always a fully covered market and bargaining is efficient. The bargained outcome is a unit price p and a fixed fee f that maximise Ω^M when $S = v - (t/2) - p + f$ and $\pi = p - c + f$. The solution is given by

$$\widehat{p}^M = c \quad \text{and} \quad \widehat{f}^M(\alpha) = (1-\alpha)\left(v - c - \frac{t}{2}\right). \tag{53}$$

The resulting profits and health plan surplus are identical to (48) and (49); i.e., $\widehat{\pi}^M(\alpha) = \pi^M(\alpha)$ and $\widehat{S}^M(\alpha) = S^M(\alpha)$.

This means that both the health plan and the producer are indifferent between bargaining over a uniform price or a two-part tariff as long as the bargaining outcome yields a fully covered market. In both cases, bargaining is efficient and each party receives a share of the surplus equivalent to its relative bargaining power. This is not the case if bargaining over a uniform price yields a partially covered market. The relevant profit comparison yields

$$\widehat{\pi}^M(\alpha) - \widetilde{\pi}^M(\alpha) = (1-\alpha)\left(v - c - \frac{t}{2} - \frac{(1+\alpha)(v-c)^2}{4t}\right) > 0 \tag{54}$$

and

$$\widehat{S}^M(\alpha) - \widetilde{S}^M(\alpha) = \alpha\left(v - c - \frac{t}{2}\right) - \frac{(1+\alpha)^2(v-c)^2}{8t} < 0, \tag{55}$$

where the signs of these two expressions are determined by invoking the condition $v - c < 2t/(1+\alpha)$, which is required to ensure an interior solution under uniform price bargaining. Thus, whereas the

producer prefers to bargain over a two-part tariff, the health plan prefers to bargain over a uniform price only. These results are in line with the ones we derive in Section 3.

5.2.2 Therapeutic competition

In the case of therapeutic competition, we assume simultaneous pair-wise Nash bargaining between the health plan and each of the two producers. If there are two available therapeutic substitutes, this positively affects the disagreement point of the health plan compared with the monopoly case. Using the standard assumptions for Nash-in-Nash bargaining (see Horn and Wolinsky, 1988, for details), the bargained outcome is the solution to the following pair of Nash products:

$$\Omega_1^D := (S_{12} - S_2)^{\alpha} \pi_1^{1-\alpha}, \tag{56}$$

$$\Omega_2^D := (S_{12} - S_1)^{\alpha} \pi_2^{1-\alpha}, \tag{57}$$

where the disagreement point for the health plan in the bilateral Nash bargaining problem between the health plan and Producer i is the anticipated surplus from reaching an agreement with Producer j, $i, j = 1, 2$, $i \neq j$.

If the participants bargain over a *uniform price* only, the bargaining outcome is a pair of prices, p_1 and p_2, that maximise Ω_1^D and Ω_2^D, respectively, when π_i, S_i, S_{12} and y_i are given by (2), (3), (5) and (17), respectively, with $f_i = 0$, $i = 1, 2$. The symmetric solution is given by

$$p_i^D(\alpha) = \frac{2\alpha c + (1-\alpha)(\beta c + t)}{2\alpha + (1-\alpha)\beta}, \tag{58}$$

and the corresponding profits and health plan surplus are given by, respectively,

$$\pi_i^D(\alpha) = \frac{(1-\alpha)t}{2(2\alpha + (1-\alpha)\beta)}, \tag{59}$$

and

$$S_{12}^D(\alpha) = v - c - \frac{t(4 - 2\alpha + (1-\alpha)\beta)}{4(2\alpha + (1-\alpha)\beta)}. \tag{60}$$

If instead the participants bargain over a two-part tariff, the bargaining outcome is a pair of unit prices and fixed fees, where (p_1, f_1) maximises Ω_1^D and (p_2, f_2) maximises Ω_2^D. The symmetric

solution is given by

$$\widehat{p}_i^D = c \text{ and } \widehat{f}_i^D(\alpha) = \frac{(1-\alpha)t}{4}, \tag{61}$$

and the corresponding profits and health plan surplus are

$$\widehat{\pi}_i^D(\alpha) = \widehat{f}_i(\alpha) = \frac{(1-\alpha)t}{4}, \tag{62}$$

and

$$\widehat{S}_{12}^D(\alpha) = v - c - \left(\frac{3-2\alpha}{4}\right)t. \tag{63}$$

With therapeutic competition, due to our assumption of a fully covered market, notice that bargaining is always efficient for $\Delta v = 0$. Thus, regardless of whether the health plan and the producers bargain over a two-part tariff or over a uniform price only, the relative bargaining strength of the participants only affects the division but not the magnitude of the total surplus. However, for a given value of α, and in contrast to the monopoly case under inelastic demand, the division of the surplus is also affected by the bargaining scope. A comparison of (59)-(60) and (62)-(63) yields

$$\widehat{\pi}_i^D(\alpha) - \pi_i^D(\alpha) = -\frac{(1-\alpha)^2(2-\beta)t}{4(2\alpha+(1-\alpha)\beta)} < 0 \tag{64}$$

and

$$\widehat{S}_{12}^D(\alpha) - S_{12}^D(\alpha) = \frac{(1-\alpha)^2(2-\beta)t}{2(2\alpha+(1-\alpha)\beta)} > 0. \tag{65}$$

Thus, under therapeutic competition and Nash bargaining, the results are once more similar to the main analysis (in Section 4). Regardless of the relative bargaining strength of the participants, the health plan prefers to bargain over a two-part tariff, whereas each of the drug producers prefers to bargain over a uniform price only.

Overall, although we have made some simplifying assumptions in order to facilitate the analysis, our results in this subsection, which are summarised in the next proposition, suggest that our main results are robust to the assumption of Nash bargaining instead of price bidding.

Proposition 9 *Suppose that drug prices are determined by Nash bargaining between the health plan and the drug producer(s). (i) In case of monopoly, both the health plan and the drug producer are indifferent between bargaining over a uniform price or over a two-part tariff if the bargaining outcome yields a fully covered market. However, in case of a partially covered market, the health*

plan prefers to bargain over a uniform price, whereas the producer prefers to bargain over a two-part tariff. (ii) In case of therapeutic competition between two drugs of equal quality ($\Delta v = 0$), the health plan always prefers to bargain over a two-part tariff, whereas the drug producers prefer to bargain over a uniform price.

5.3 Drug innovation

Our main analysis rests on the implicit assumption that drug innovation has already taken place. It is therefore an analysis of how different payment schemes affect static efficiency and the division of a surplus created by an exogenously given set of drugs. In this subsection we extend our main analysis in order to derive some implications of different drug payment schemes also for dynamic efficiency. In particular, we are interested in how different payment schemes might affect incentives for different types of innovation strategies. Obviously, drug producers' incentives for spending resources on drug innovation depend on the profits they can expect to earn in case of successful innovation. This might in turn depend on the type of payment schemes that will be applied to new drugs. One of the main results of our analysis is that drug producers tend to prefer two-part tariffs under monopoly but uniform pricing under therapeutic competition. This suggests that a (permanent) switch from uniform pricing to two-part tariffs as the main drug payment scheme makes it relatively more profitable for new drugs to obtain a monopoly position and might therefore stimulate drug producers to spend more resources on 'drastic' innovation relative to so-called 'me-too' innovation.

In order to illustrate and explore the above suggested mechanism within the context of our modelling framework, consider the following simple modification of our main model. Suppose that there are two separate therapeutic markets, denoted by A and B, that are identical in terms of the locations and characteristics of patients, as laid out in Section 2. In Market A there is currently a monopoly supplier of a drug, located at one of the endpoints. Market B, on the other hand, is only a 'potential market'; i.e., it consists of patients with a type of disease for which no drug treatment currently exists.

Consider now the innovation incentives of a producer (different from the current monopolist in Market A) with a given R&D budget that can be used in one of two different ways: (i) developing a therapeutic substitute to the existing drug in Market A (located at the other endpoint of the Hotelling line), or (ii) trying to develop a pioneer drug for Market B (located at one of the endpoints

of the Hotelling line). We will refer to (i) as *me-too innovation* and (ii) as *drastic innovation*. Suppose for simplicity that the costs of both alternatives are the same, and lower than expected ex post profits, but that the probability of success is larger for me-too innovation than for drastic innovation. More specifically, suppose that a therapeutic substitute can be developed with certainty while the probability of successful drastic innovation is $\phi < 1$. Suppose also that the health benefit parameter v and the cost parameter c are equal in both cases, and equal to the respective parameters of the existing drug in Market A. These simplifications imply that the only difference between the two alternatives (in case of successful innovation) is the market structure that the producer of the new drug faces. Suppose also that the parameters of the model are such that the market is always fully covered in the monopoly case, regardless of payment scheme. Finally, we assume that drug prices result from Nash bargaining along the lines of the analysis in the previous subsection, with the relative bargaining power of the health plan given by $\alpha \in (0, 1)$.

Under the above set of assumptions, if the drug producer expects that the payment scheme for the new drug will be *uniform pricing*, the expected profit gain of pursuing drastic innovation instead of me-too innovation is found by comparing (48) and (59), and is given by

$$\pi_i^M(\alpha) - \pi_i^D(\alpha) = \phi(1-\alpha)\left(v - c - \frac{t}{2}\right) - \frac{(1-\alpha)t}{2(2\alpha + (1-\alpha)\beta)}. \tag{66}$$

This expected profit gain is positive, implying that the producer will spend its resources on drastic innovation, if $\phi > \phi^u$, where

$$\phi^u := \frac{t}{(2\alpha + (1-\alpha)\beta)(2(v-c) - t)}. \tag{67}$$

It is straightforward to verify that $\phi^u < 1$ when imposing the parameter restrictions consistent with the above set of assumptions. Thus, drastic innovation will be chosen if the probability of success is sufficiently high. Otherwise, me-too innovation will be chosen.

We can contrast this decision rule with the socially optimal rule. If a new drug is developed in Market B, this creates a net health benefit of $v - t/2$ while total production costs are c. On the other hand, if a therapeutic substitute is developed in Market A, this will result in an overall reduction of therapeutic mismatch costs by $t/4$ at no extra production costs. Thus, in terms of total expected surplus, spending a given amount of resources on drastic innovation instead of me-too

innovation is socially efficient if

$$\phi\left(v-\frac{t}{2}-c\right)>\frac{t}{4}. \tag{68}$$

This condition holds if $\phi>\phi^*$, where

$$\phi^*:=\frac{t}{4\left(v-c-\frac{t}{2}\right)}. \tag{69}$$

However, notice that

$$\phi^u-\phi^*=\frac{(1-\alpha)(2-\beta)t}{4\left(2\alpha+(1-\alpha)\beta\right)\left(v-c-\frac{t}{2}\right)}>0, \tag{70}$$

which implies that, under uniform pricing, drug producers have socially inefficient innovation incentives. If the probability of successful drastic innovation is $\phi\in(\phi^*,\phi^u)$, the drug producer will choose me-too innovation even if the expected total surplus would be higher by spending the same resources on drastic innovation. Furthermore, it is easily verified that $\phi^*-\phi^u$ is monotonically decreasing in α, implying that the distortion of innovation incentives is larger when drug producers have more bargaining power in price determination.

On the other hand, if the drug producer expects that the payment scheme for the new drug will be *two-part tariffs*, the expected profit of drastic innovation is the same as under uniform pricing, but the profit of me-too innovation is given by (62). The expected profit gain of pursuing drastic innovation instead of me-too innovation is therefore given by

$$\widehat{\pi}_i^M(\alpha)-\widehat{\pi}_i^D(\alpha)=\phi(1-\alpha)\left(v-c-\frac{t}{2}\right)-\frac{(1-\alpha)t}{4}. \tag{71}$$

It is easily verified that this profit gain is positive (negative) if $\phi>(<)\,\phi^*$. Thus, under two-part tariffs, the producer has socially optimal innovation incentives. The following proposition summarises our results on payment schemes and innovation incentives.

Proposition 10 *For any given degree of relative bargaining power, (i) bargaining over two-part tariffs yields socially optimal incentives for spending resources on drastic innovation relative to me-too innovation, whereas (ii) bargaining over a uniform price yields too strong incentives for me-too innovations relative to drastic innovations, compared to what is socially optimal.*

Although based on a special case of the main model, with full symmetry in case of therapeutic competition, the analysis in this section identifies a potential second efficiency gain from using two-

part tariffs instead of uniform pricing. Whereas our main analysis shows that a switch from uniform pricing to two-part tariffs is likely to improve static efficiency, the extended analysis suggests that such a switch of payment scheme might also improve dynamic efficiency by redirecting resources from me-too innovations to drastic innovations in a socially optimal way.

The main analysis also showed that, under therapeutic competition, it is generally in the health plan's interest to generate the static efficiency gains brought about by two-part tariffs, since this payment scheme distributes the surplus in a way that is generally more favourable for the health plan. However, this is less likely to be the case for the dynamic efficiency gains identified in this extension. If the health plan has zero bargaining power, as in the main analysis, stimulating drastic innovation is clearly not in the health plan's interest, since all the surplus generated by such innovation will be captured by the innovating producer. More generally, if the health plan faces the possibility of stimulating different types of innovation incentives through the choice of payment scheme, the following trade-off arises. On the one hand, drastic innovation yields a larger health gain than me-too innovation. On the other hand, me-too innovation increases competition and yields lower drug expenditures, all else equal. The relative importance of these two considerations depends on the health plan's relative bargaining strength. More bargaining power allows the health plan to capture a larger share of the surplus created by drastic innovation, while simultaneously reducing the benefit of increased competition in markets with existing drugs. Thus, stronger bargaining power increases the importance of the former consideration while reducing the importance of the latter. This can be illustrated by a comparison of (49) and (63). Under two-part tariffs, having access to two innovative drugs (monopoly in two different markets) is better for the health plan than having therapeutic competition in Market A and no drugs in Market B if

$$2\widehat{S}^{M}(\alpha) - \widehat{S}_{12}^{D}(\alpha) = (2\alpha - 1)\left(v - c - \frac{3}{4}t\right) > 0, \tag{72}$$

which is true if $\alpha > 1/2$. In other words, the health plan prefers drastic innovation over me-too innovation only if it has more bargaining power than drug producers in price negotiations.

Finally, it is also not immediately clear whether the change innovation incentives brought about by a change in the payment scheme is in the interest of the drug producers. Notice that the introduction of two-part tariffs makes drastic innovation relatively more attractive not because the profitability of such innovation increases, but because the profitability of me-too innovation is

reduced. Within the time horizon of the current model, this is not in the interest of the innovating firm, all else equal. On the other hand, if we take a longer time perspective, weaker incentives for me-too innovations also means that, once the innovator of a pioneer drug has obtained a monopoly position, the probability of facing therapeutic competition from other producers in the future is lower.

6 Concluding remarks

In this paper we have studied the effects of adopting a new payment scheme (the Netflix model) for pharmaceuticals in a setup where drug producers submit bids to a purchaser for being included in the health plan. In particular, we have analysed distributional effects (who gains and who loses) of switching from a traditional system with uniform pricing to a new payment scheme with two-part pricing, where the health plan pays an access (subscription) fee in addition to a low unit price (equal to marginal costs). A key insight from our analysis is that the effects of two-part pricing crucially depend on whether or not there is competition in the relevant therapeutic market. If a drug producer is a monopolist (i.e., there exists no viable therapeutic alternatives), two-part pricing is beneficial for the drug producer but not for the health plan, because it enables the monopolist to extract all of the surplus from the health plan. However, this result is reversed if there is competition among drug producers in the therapeutic market. In this case, the drug producers compete more aggressively under two-part pricing than uniform pricing for being included in the health plan. Two-part pricing benefits the health plan in two ways; it improves allocative efficiency (unit prices equal to marginal costs) and lowers drug expenditures (intensified competition). This conclusion relies, however, on the assumption of non-exclusivity, meaning that the health plan will include all drugs that contributes positively to the health plan's surplus. If we also open up for the possibility of exclusive contracting, where only one drug is included in the plan, then uniform pricing is always weakly preferred by drug purchaser.

Although the distributional effects of different payment schemes depend crucially on the market structure, two-part tariffs unambiguously promote efficient allocation of drugs due to marginal cost pricing, regardless of whether the drug producers are engaged in therapeutic competition or not. Thus, a switch from uniform pricing to two-part tariffs improves static efficiency. Interestingly, we have also identified an effect whereby a similar switch might also improve dynamic efficiency. If drug innovators expect that new drugs will be paid for by two-part tariffs instead of uniform

prices, this might spur a reallocation of R&D resources from 'me-too' innovation to more drastic innovation, leading to larger health gains for patients. However, since less me-too innovation means less therapeutic competition and therefore more expensive drugs, all else equal, the health plan will only benefit from such dynamic efficiency gains if it has sufficient bargaining power vis-à-vis drug producers.

While the paper employs a duopoly Hotelling model to study the effects of two-part pricing, the results hold more generally. Allowing for more (than two) drug companies will make competition more intense and thus reduce the distortions in consumption under uniform prices due to prices above marginal costs. However, this generalisation will not change our results in qualitative terms; neither the comparison of monopoly versus competition nor the comparison of uniform versus two-part pricing under competition. The same is true for a more general demand structure under standard assumptions. The key results depend on whether the health plan has an outside option when deciding on the inclusion, and to what extent two-part pricing results in marginal cost pricing.

By way of conclusion, we would like to point at some limitations of our study. First, our analysis has not accounted for possible externalities across markets of adopting two-part pricing. In the presence of parallel trade, drug producers may be reluctant to offer unit prices at marginal costs, as this may result in parallel export to countries with higher, uniform prices. A similar type of spillover may occur across health plans. If prices are public information, then drug producers may be reluctant to offer unit prices equal to marginal costs under two-part pricing, as this may have a negative impact on the price-setting to other health plans (using uniform pricing). Such cross-market externalities can reduce the scope for two-part pricing.

Second, throughout the analysis we have assumed a monopoly health insurer that can include one or both drugs in its health plan. Although this applies to many countries and public health insurers, we are not able to capture the interplay between competition and market power in the health insurance market and the pharmaceutical market (see, e.g., Gaynor an Vogt (2000)). Imperfect competition between health insurers can enrich the analysis by dealing with vertical restraints that may appear between drug suppliers and health insurers. Gal-Or (1997) or Bardey and Bourgeon (2011) study such issues in a context of bilateral oligopoly between health plans and hospitals, and this could be adapted to the pharmaceutical sector and payment schemes. Another interesting topic for future research is the role of third-party pharmaceutical benefit managers (PBM) that negotiate prices and administer claims for insurance providers.

Third, we model physicians' prescription choices in a highly reduced form. In practice, physicians' prescription choices are likely to be affected by many factors, including their contract with the health plan, remuneration schemes, the competitive environment, marketing by drug companies, etc. We have reduced these circumstances to a single parameter (β), measuring the weight that the physician puts on patients' health benefits relative to the price of the drug. While we believe this is a key element of physicians' prescription choices, a more careful analysis of physician behaviour is needed to fully understand the effects of different payment schemes. This includes the impact of two-part tariff schemes on health insurers' design of copayment schemes (see, e.g., Lakdawalla and Sood (2013, 2019) or Bardey et al. (2016)). However, this is beyond the scope of the current paper, and we leave this issue, as well as competition between health insurers and cross-market spillovers, to future research.

Appendix

A. Optimal uniform pricing under monopoly

Suppose that the drug is included in the health plan (i.e., that the participation constraint of the purchaser holds). Substituting the demand function, (7), into the profit function, (2), and maximising with respect to the unit price p, the optimal (unconstrained) solution is

$$p = \begin{cases} \frac{v+\beta c}{2\beta} & if \quad v \leq 2t + \beta c \\ \frac{v-t}{\beta} & if \quad v > 2t + \beta c \end{cases}, \tag{A1}$$

where $y < (=) \, 1$ if $v < (\geq) \, 2t + \beta c$.

The next step is to check what is required for the purchaser's participation constraint to hold. From (3), under uniform pricing, this constraint is given by

$$S = \left(v - p - \frac{ty}{2} \right) y \geq 0. \tag{A2}$$

Consider first the case of $v < 2t + \beta c$, for which unconstrained monopoly pricing yields $p = (v + \beta c)/2\beta$ and $y = (v - \beta c)/2t$, so (A2) becomes

$$S = \left(v \left(3\beta - 2 \right) - \left(2 - \beta \right) \beta c \right) \frac{(v - \beta c)}{8t\beta} \geq 0. \tag{A3}$$

It follows that $S \geq 0$ if $\beta \geq \beta^*$, where

$$\beta^* := 1 - \left(\frac{3v - \sqrt{9v^2 - 4c\,(v-c)}}{2c} \right). \tag{A4}$$

If β is below this threshold level, the drug producer's price setting is constrained by the condition that the health plan's surplus must be non-negative. In this case, the optimal (constrained) price solves

$$\left(v - p - \frac{ty}{2} \right) y = 0, \tag{A5}$$

where y is either an interior solution given by $y = (v - \beta p)/t$, or a corner solution given by $y = 1$. The optimal (constrained) price that implements an interior solution is therefore given by

$$p = \frac{v}{2 - \beta}, \tag{A6}$$

whereas the price that implements a corner solution is

$$p = v - \frac{t}{2}. \tag{A7}$$

aBy a simple comparison of profits, we find that the optimal price is given by (A6) if $\beta > \beta^{**}$ and by (A7) if $\beta < \beta^{**}$, where

$$\beta^{**} := \frac{2\,(v-t)}{2v - t}. \tag{A8}$$

Consider now the case of $v > 2t + \beta c$, which implies a fully covered market in the profit-maximising solution. Setting $p = (v-t)/\beta$ and $y = 1$ in (A2), the purchaser's participation constraint is given by

$$S = \frac{t\,(2-\beta) - 2\,(1-\beta)\,v}{2\beta} \geq 0. \tag{A9}$$

This constraint holds if $\beta \geq \beta^{**}$. On the contrary, if $\beta < \beta^{**}$, the producer must offer the purchaser a lower price. In this case, we have already found that the highest price that the purchaser is willing to accept is the price given by (A7).

Finally, a comparison of the two threshold values β^* and β^{**} shows that

$$\beta^* - \beta^{**} = \frac{(2v-t)\left(\sqrt{9v^2 - 4c\,(v-c)} - 3v \right) + 2ct}{2c\,(2v-t)}. \tag{A10}$$

Thus, $\beta^* > \beta^{**}$ iff

$$(2v-t)\sqrt{9v^2-4c(v-c)} > 3v(2v-t)-2ct, \tag{A11}$$

which is equivalent to

$$(2v-t)^2\left(9v^2-4c(v-c)\right) > \left(3v(2v-t)-2ct\right)^2, \tag{A12}$$

or, after re-arranging terms,

$$t(5v-2(c+t))-2v(v-c) > 0. \tag{A13}$$

This condition holds if v is sufficiently low. Recall that the threshold value β^* is only relevant in the parameter set defined by $v < 2t+\beta c$. Although it is not feasible to show analytically, numerical simulations confirm that $\beta^* \geq \beta^{**}$ for all $v \leq 2t+\beta^* c$, implying that $\frac{1}{2} < \beta^{**} \leq \beta^* < 1$ in the relevant range of parameters. This completes the characterisation of all the relevant regimes and the monopoly solution under uniform pricing is fully characterised by (8).

B. Equilibrium existence under therapeutic competition with uniform pricing

The Nash equilibrium given by (20) exists if the surplus of the health plan, S_{12}^D, is non-negative. To confirm this, assume without loss of generality that $v_i \geq v_j$, implying $\Delta v \geq 0$. From (25) we derive

$$\frac{\partial S_{12}^D}{\partial v_i} = \frac{1}{2} + \frac{(5\beta-4)\Delta v}{18t\beta} > 0 \text{ for } \Delta v < 3t. \tag{B1}$$

Thus, for a given value of v_j, S_{12}^D is monotonically increasing in v_i, which implies that the scope for $S_{12}^D \geq 0$ is minimised for $v_i = v_j$. Evaluating S_{12}^D at $v_i = v_j$ yields

$$S_{12}^D\big|_{v_i=v_j} = v_i - c - \frac{4+\beta}{4\beta}t. \tag{B2}$$

It follows that a sufficient (but not necessary) condition for $S_{12}^D \geq 0$ is

$$\min\{v_i, v_j\} \geq c + \frac{4+\beta}{4\beta}t. \tag{B3}$$

C. Proofs

Proof of Proposition 6

Let drug i be the high-quality drug and drug j the low-quality drug, implying $\Delta v > 0$.

(i) Consider first the case of *de facto* therapeutic competition, which requires $\Delta v < t$. From (25) and (36), a comparison of the *health plan surplus* under uniform pricing and two-part tariffs yields

$$S_{12}^D - \widehat{S}_{12}^D = \frac{(7\beta - 2)(\Delta v)^2 - (2-\beta)\,9t^2}{18t\beta}. \tag{C1}$$

The sign of (C1) depends on the sign of the numerator, which is monotonically increasing in Δv. Setting Δv at the upper bound, $\Delta v = t$, the numerator reduces to $-4t^2(5-4\beta) < 0$. Thus, $S_{12}^D < \widehat{S}_{12}^D$ for all $\Delta v < t$.

Consider next the case of *potential* therapeutic competition, which requires $\Delta v \in (t, 3t)$. From (25) and (42), a similar *health plan surplus* comparison yields

$$S_{12}^D - \widehat{S}_i^d = \frac{9t\left((2\Delta v + t)\,\beta - 4t\right) + (5\beta - 4)(\Delta v)^2}{36t\beta}. \tag{C2}$$

The sign of (C2) depends on the sign of the numerator, which we define as A, where

$$\frac{\partial A}{\partial \beta} = (3t + \Delta v)(3t + 5\Delta v) > 0, \tag{C3}$$

implying that the numerator is monotonically increasing in β. Evaluating A at the lower bound of β yields

$$A|_{\beta=\frac{1}{2}} = -\frac{3}{2}\left(3t(7t - 2\Delta v) + (\Delta v)^2\right) < 0 \text{ for } \Delta v \in (t, 3t). \tag{C4}$$

Evaluating A at the upper bound of β yields

$$A|_{\beta=1} = 9t(2\Delta v - 3t) + (\Delta v)^2, \tag{C5}$$

which is monotonically increasing in Δv. It is easily verified that $A|_{\beta=1} < 0$ for $\Delta v = t$ and $A|_{\beta=1} > 0$ for $\Delta v = 3t$. Thus, $A > 0$, implying $S_{12}^D > \widehat{S}_i^d$ if both β and Δv are sufficiently high. Otherwise, $S_{12}^D < \widehat{S}_i^d$.

(ii) Consider first the case of *de facto* therapeutic competition, which requires $\Delta v < t$. From (23) and (33), a comparison of profits under uniform pricing and two-part tariffs for the *high-quality*

producer yields

$$\pi_i^D - \widehat{\pi}_i^D = \frac{9t^2(2-\beta) - \Delta v\left((9\beta-2)\Delta v - 6t(2-3\beta)\right)}{36t\beta}. \tag{C6}$$

The sign of (C6) is given by the sign of the numerator, which we define as B, where

$$\frac{\partial B}{\partial \beta} = -9(t+\Delta v)^2 < 0, \tag{C7}$$

implying that the numerator is monotonically decreasing in β. Evaluating B at the lower bound of β yields

$$B|_{\beta=\frac{1}{2}} = \frac{1}{2}(3t-\Delta v)(9t+5\Delta v) > 0 \text{ for } \Delta v < t. \tag{C8}$$

Evaluating B at the upper bound of β yields

$$B|_{\beta=1} = 3t(3t-2\Delta v) - 7(\Delta v)^2. \tag{C9}$$

It is easily verified that $B|_{\beta=1} > 0$ if Δv is sufficiently close to zero, and $B|_{\beta=1} < 0$ if Δv is sufficiently close to t. Thus, $\pi_i^D < \widehat{\pi}_i^D$ if β and Δv are both sufficiently high. Otherwise, $\pi_i^D > \widehat{\pi}_i^D$. A similar profit comparison for the *low-quality producer* yields

$$\pi_j^D - \widehat{\pi}_j^D = \frac{9t^2(2-\beta) - \Delta v\left((9\beta-2)\Delta v + 6t(2-3\beta)\right)}{36t\beta}. \tag{C10}$$

The sign of (C10) depends on the sign of the numerator, which we define as C, where

$$\frac{\partial C}{\partial \beta} = -9(t-\Delta v)^2 < 0, \tag{C11}$$

implying that the numerator is monotonically decreasing in β. Evaluating C at the upper bound of β yields

$$C|_{\beta=1} = \Delta v(6t-7\Delta v) + 9t^2 > 0 \text{ for } \Delta v < t. \tag{C12}$$

Thus, $C > 0$, implying $\pi_j^D > \widehat{\pi}_j^D$ for all $\beta \in \left(\frac{1}{2}, 1\right]$ and $\Delta v < t$.

Consider next the case of *potential* therapeutic competition, which requires $\Delta v \in (t, 3t)$. In this regime, the *low-quality producer* is included in the health plan under uniform pricing and excluded from the health plan under two-part tariffs, which obviously means that the profits are higher in

the former case. From (23) and (40), a profit comparison for the *high-quality producer* yields

$$\pi_i^D - \widehat{\pi}_i^d = \frac{3t\left(3t - 2\Delta v\left(3\beta - 1\right)\right) + \left(\Delta v\right)^2}{18t\beta}. \tag{C13}$$

The sign of (C13) depends on the sign of the numerator, which we define as E, where

$$\frac{\partial E}{\partial \beta} = -\frac{\left(3t + \Delta v\right)^2}{18t\beta^2} < 0, \tag{C14}$$

implying that the numerator is monotonically decreasing in β. Evaluating E at the lower bound of β yields

$$E|_{\beta=\frac{1}{2}} = \frac{3t\left(3t - \Delta v\right) + \left(\Delta v\right)^2}{9t} > 0 \text{ for } \Delta v \in \left(t, 3t\right), \tag{C15}$$

while evaluating E at the upper bound of β yields

$$E|_{\beta=1} = \frac{\left(\Delta v\right)^2 - 3t\left(4\Delta v - 3t\right)}{18t} < 0 \text{ for } \Delta v \in \left(t, 3t\right). \tag{C16}$$

Thus, $\pi_i^D > (<)\, \widehat{\pi}_i^d$ if β is sufficiently low (high).

(iii) Consider first the case of *de facto* therapeutic competition, which requires $\Delta v < t$. From (24) and (35), a comparison of the *total health benefit* under uniform pricing and two-part tariffs yields

$$H_i^D - \widehat{H}_i^D = -\frac{\left(\Delta v\right)^2}{9t} < 0. \tag{C17}$$

For the case of *potential* competition, which requires $\Delta v \in (t, 3t)$, a similar comparison yields

$$H_i^D - \widehat{H}_i^d = \frac{\left(3t - 5\Delta v\right)\left(3t - \Delta v\right)}{36t} < 0 \text{ for } \Delta v \in \left(t, 3t\right). \tag{C18}$$

Proof of Proposition 8

(i) For $\Delta v < t$, the health benefit of exclusive contracting when the producers use two-part tariffs, denoted by ΔH, is given by a comparison of (41) and (35), yielding

$$\Delta H = v_i - \frac{t}{2} - \left(\overline{v} - \frac{t}{4} + \frac{\left(\Delta v\right)^2}{4t}\right) = -\frac{\left(t - \Delta v\right)^2}{4t} < 0. \tag{C19}$$

The effect of exclusive contracting on the surplus of the health plan, denoted by ΔS, is given by a comparison of (42) and (36), yielding

$$\Delta S = v_j - c - \frac{t}{2} - \left(\overline{v} - c - \frac{3t}{4} - \frac{(\Delta v)^2}{4t} \right) = \frac{(t - \Delta v)^2}{4t} > 0. \tag{C20}$$

For $\Delta v \in [t, 3t)$, the Nash equilibrium under exclusive contracting is identical to the Nash equilibrium under non-exclusivity (which is characterised by *potential therapeutic competition*). Contract exclusivity is therefore irrelevant in this parameter range.

(ii) Under uniform pricing, the health benefit of non-exclusivity (instead of exclusive contracting) is given by (C18) in the proof of Proposition 6, where in this case the relevant parameter range is $\Delta v < 3t$. It is straightforward to verify that this expression is positive (negative) if $\Delta v < (>) \frac{3}{5}t$. Similarly, the effect of non-exclusivity on the health plan's surplus is given by (C2) in the proof of Proposition 6, once more with the relevant parameter range being $\Delta v < 3t$. It is easily verified that the condition for a positive (resp. negative) sign of this expression, which in the proof of Proposition 6 is derived for $\Delta v \in [t, 3t)$, also extends to $\Delta v < t$. Thus, (C2) is positive if both β and Δv are sufficiently high, and negative otherwise.

References

[1] Armstrong, M. and J. Vickers, 2001. Competitive price discrimination. *The RAND Journal of Economics*, 32(4), pp. 579–605.

[2] Bardey, D., A. Bommier and B. Jullien, 2010. Retail price regulation and innovation: reference pricing in the pharmaceutical industry. *Journal of Health Economics*, 29(2), pp. 303–316.

[3] Bardey, D. and J-M. Bourgeon, 2011. Health Care Network Formation and Policyholders' Welfare. *The B.E. Journal of Economic Analysis & Policy*, 11 (2).

[4] Bardey, D., B. Jullien and J.-M. Lozachmeur, 2016. Health insurance and diversity of treatment. *Journal of Health Economics*, 47, pp. 50–63.

[5] Brekke, K.R., I. Königbauer and O.R. Straume, 2007. Reference pricing of pharmaceuticals. *Journal of Health Economics*, 26(3), pp. 613–642.

[6] Brekke, K.R., C. Canta and O.R. Straume, 2016. Reference pricing with endogenous generic entry. *Journal of Health Economics,* 50, pp. 312–329.

[7] Cachon, G.P. and A.G. Kök, 2010. Competing manufacturers in a retail supply chain: on contractual form and coordination. *Management Science*, 56(3), pp. 571–589.

[8] Carrera, M., Goldman, D.P., Joyce, G., and N. Sood, 2018. Do Physicians Respond to the Costs and Cost-Sensitivity of Their Patients? *American Economic Journal: Economic Policy*, 10(1), pp. 113-152.

[9] Danzon, P. M. and L.-W. Chao, 2000. Does Regulation Drive out Competition in Pharmaceutical Markets? *The Journal of Law and Economics*, 43(2), pp. 311-358.

[10] Danzon, P.M. and A.J. Epstein, 2012. Effects of Regulation on Drug Launch and Pricing in Interdependent Markets. In: Bolin, K. and Kaestner, R. (Ed.) The Economics of Medical Technology (Advances in Health Economics and Health Services Research, Vol. 23), Emerald Group Publishing Limited, Bingley, pp. 35-71.

[11] Ellison, S.F., Cockburn, I., Griliches, A. and J. Hausman, 1997. Characteristics of demand for pharmaceutical products: An examination of four cephalosporins. *RAND Journal of Economics*, 28, pp. 426-446

[12] Gabrielsen, T.S. and L. Sørgard, 1998. The pro-competitive effect of two-part tariffs. *International Journal of the Economics of Business*, 5(1), pp. 47–55.

[13] Gal-Or, E., 1997. Exclusionary Equilibria in Health-Care Markets. *Journal of Economics & Management Strategy*, 6(1), pp. 5–43.

[14] Gaynor, M. and W. B. Vogt, 2000. Antitrust and competition in health care markets. *Handbook of Health Economics*, Elsevier, Volume 1, Part B, pp. 1405-1487,

[15] González, P., I., Macho-Stadler, and D. Pérez-Castrillo, 2016. Private versus social incentives for pharmaceutical innovation. *Journal of Health Economics*, 50, pp. 286–297.

[16] Horn, H. and A. Wolinsky, 1988. Bilateral monopolies and incentives for merger. *RAND Journal of Economics*, 19(3), pp. 408–419.

[17] Kanavos, P., J. Costa-Font and A. McGuire, 2007. Product differentiation, competition and regulation of new drugs: the case of statins in four European countries. *Managerial and Decision Economics*, 28(4-5), pp. 455–465.

[18] Lakdawalla, D. and N. Sood, 2009. Innovation and the welfare effects of public drug insurance. *Journal of Public Economics*, 93, pp. 541–548.

[19] Lakdawalla, D. and N. Sood, 2013. Health insurance as a two-part pricing contract. *Journal of Public Economics*, 102, pp. 1–12.

[20] Lu, Z. J. and W. S. Comanor, 1998. Strategic Pricing of New Pharmaceuticals. *The Review of Economics and Statistics*, 80(1), pp. 108–118.

[21] Mailankody, S. and V. Prasad, 2016. Implications of proposed medicare reforms to counteract high cancer drug prices. *JAMA*, 316(3), pp. 271–272.

[22] McCann, NC, T.H. Horn, E.P. Hyle and R.P. Walensky, 2020. HIV antiretroviral therapy costs in the United States, 2012-2018. *JAMA Internal Medicine*,180(4), pp. 601–603.

[23] Miraldo, M., 2009. Reference pricing and firms' pricing strategies. *Journal of Health Economics*, 28(1), pp. 176–197.

[24] Moon, S. and E. Erickson, 2019. Universal medicine access through lump-sum remuneration - Australia's approach to hepatitis C. *New England Journal of Medicine*, 380(7), pp. 607–610.

[25] Oi, W., 1971. A Disneyland dilemma: two-part tariffs for a Mickey Mouse monopoly. *Quarterly Journal of Economics*, 85(1), pp. 77–90.

[26] Scott Morton, F. and M. Kyle, 2011. Markets for Pharmaceutical Products. Chapter 12 in Handbook of Health Economics (eds. M.V. Pauly, T.G. Mcguire, P.P. Barros), 2, pp. 763-823.

[27] Yin, X., 2004. Two-part tariff competition in duopoly. *International Journal of Industrial Organization,* 22(6), pp. 799–820.