Pricing above value: selling *to* an adverse selection market

Jan Boone*

September 21, 2021

Abstract

This paper shows that selection incentives in downstream markets affect upstream prices. In fact, it is possible for inputs to be priced above the value that the good has for final consumers. This happens in adverse selection markets where low types mimic high types. High input prices then help to separate consumer types and reduce information rents. We use the example of pharmaceutical companies selling drugs to a health insurance market at prices exceeding treatment value. Another feature of the model is an excessive private incentive to reduce market size, e.g. in the form of personalized medicine.

Keywords: adverse selection, pricing above value, vertical relations, pharmaceutical prices

JEL codes: I13, I11

*Tilburg University, Department of Economics, Tilec and CEPR, E-mail: *j.boone@uvt.nl*.

1 Introduction

Consider a value chain where firms in an upstream market U sell inputs to firms in a downstream market D and the latter sell to final consumers. When considering the input prices set by firms in U, there are a number of common sense results: 1. do not set the input price above the final consumers' valuation of your input; 2. try to innovate to make your product attractive to a bigger group of final consumers and 3. reduce your price as the downstream market becomes more competitive.

However, when U sells to a market D with adverse selection problems, these three results do not necessarily hold. If low types in D mimic the high types, it is optimal for U to set prices above final consumers' valuation of the input. Limiting the types of final consumers that value your product raises profits. And input prices increase with the competition intensity in downstream market segments.[1]

An interesting market to apply this framework to is the health insurance market (D) where pharmaceutical companies in market U sell drugs to insurers to be included in their insurance contracts. Health insurance markets are known for their adverse selection problems and almost daily there are stories in the news about pharmaceutical companies charging "outrageous" prices for their treatments.[2] Howard et al. (2015) document price developments in the market for anticancer drugs. Although it is hard to put a value on an additional year of life to see whether prices are above treatment value, they argue (pp. 149) that "in 1995 patients and their insurers paid $54,100 for a year of life. A decade later, 2005, they paid $139,100 for the same benefit. By 2013, they were paying $207,000." Most regulators in the world use less than $200k as the monetary value of a life year and indeed Cavalli (2013) states that at the World Oncology Forum the "prevailing opinion was that ... the cost of the new generation of drugs is getting out of all proportion to the added benefit." Note that many of these expensive drugs are not the first to treat these forms of cancer. Their added benefit is compared to existing (sometimes generic) drugs.

Finally, Howard et al. (2015) note that "launch prices of new anticancer drugs and

[1] Note that this can happen with linear pricing and double marginalization. As the downstream market becomes more competitive, the "second marginalization" decreases, allowing for higher input prices. Double marginalization disappears if upstream firms use two-part tariffs. In our paper also with two-part tariffs it is the case that input prices can increase with downstream competition.

[2] Examples include `https://nyti.ms/2NvLoZx`, `https://nyti.ms/2bHXkFj` and `https://nyti.ms/2nKMJ2a`.

other drugs in the so-called 'specialty' pharmaceutical market have been increasing over time" (pp. 140). The promise of specialty or precision medicine was "to give 'the right drug to the right patient' to maximize the effectiveness and safety of the treatment" (Garattini et al., 2015). However, up till now this targeting of treatments has not lived up to this promise and one reason is that these treatments turn out to be extremely expensive. It is not clear that we can afford precision medicine; see Doble (2016) for a discussion in the case of oncology with examples of treatments costing $300,000 while they "only result in minimal benefit".

We propose a set-up with upstream pharmaceutical companies selling drugs to downstream insurers for inclusion in their health insurance contracts. The health insurance market suffers from adverse selection, which we model with two consumer types. The condition we need for our results is that low risk types try to mimic high risk types which we model following the countervailing incentives literature (Lewis and Sappington, 1989). In this case, insurance coverage of upstream drugs helps to separate downstream types and reduce the low type's information rent. This allows pharmaceutical firms to charge prices in excess of their treatments' value for patients and still have their treatments covered by insurance. In other words, to reduce rents the high type's allocation is distorted upwards by covering a treatment that is (too) expensive. Any change that makes the low type's incentive compatibility constraint "more binding" (i.e. makes the high type contract more attractive to the low type) makes covering this treatment more attractive and increases the excess profits that upstream firms can earn. We denote these excess profits –i.e. profits in excess of treatment value– supra profits.

We show that supra profits are increasing in the competitiveness of the high type market: as the high type market becomes more competitive, the insurance premium tends to fall making the high contract more attractive to the low type. Further, targeting of pharmaceutical R&D investments on subgroups tends to increase supra profits: as long as the targeting reduces the efficacy among low types faster than among high types it helps to separate types and raises supra profits. We derive conditions under which such targeting of R&D is privately profitable and socially wasteful. We interpret the last result in the context of specialty drugs or personalized/precision medicine which leads to higher prices more than better treatment results. Finally , the introduction of generic drugs helps to increase prices of patented drugs. To the extent that people can decide to go without insurance, the premium should not exceed the total value of the insurance contract. Generic drug competition leads to prices below the value of treatment thereby creating the

"space" for patented drugs to charge prices above value.

The main assumption we need is that low types want to mimic high types. We provide two rationales for this assumption. First, in the main model the high type market segment is more competitive than the low type segment in the insurance market. In our health insurance context this implies that high risk individuals are more sensitive to value differences between insurance contracts than low risk types. There are a number of reasons why this is the case. Parente et al. (2004) find that the insurance premium elasticity is substantially higher for employees with a chronic condition than without. One explanation for this is that high risk types are more likely to have experience with healthcare and therefore know more about quality differences between treatments. As they are better informed, they are more sensitive to quality differences between plans (Gaynor et al., 2015).

A second intuition is that high risk types tend to have lower incomes and therefore pay more attention to value differences between plans.[3] Ho et al. (2017) find that consumers' insurer switching probability falls with income. Also Atherly et al. (2004), Auerbach and Ohri (2006) and Saltzman (2019) find that people on low income are more price sensitive when choosing health insurance. This is in line with Royalty and Solomon (1999) where the higher educated and the wealthier are less price sensitive.[4]

Next to differences in competitiveness, a second rationale why low risk types mimic high risk types is a lack of information/rationality. For example, Handel and Kolstad (2015) document that low risk types buy the contract with low out-of-pocket expenditures (targeted at high risk types) because they (incorrectly) believe

[3]The correlation between health status and income is well documented in the empirical health literature, see for example Frijters et al. (2005), Finkelstein and McGarry (2006), Gravelle and Sutton (2009) or Munkin and Trivedi (2010). Potential explanations for this correlation include the following. High income people are better educated and hence know the importance of healthy food, exercise etc. Healthy food options tend to be more expensive and therefore better affordable to high income people. Or (with causality running in the other direction) healthy people are more productive and therefore earn higher incomes.

[4]There are also papers suggesting that people with lower health status tend to be less price elastic when choosing insurance. A number of these papers are based on age as an indicator of health status: older people tend to have lower health status (Strombom et al., 2002; Royalty and Solomon, 1999). But Costa and Garcia (2003) find no difference in elasticity between age groups. As pointed out by Beaulieu (2002), another reason for the lower price elasticity is that older people (people with low health status in general) tend to be better informed about the quality of the different health plans and the treatments they cover. This can explain why they react less to price changes and to plan quality information published by the government or an employer.

this contract is more generous than it actually is, e.g. because they believe it covers more treatments and a wider network of providers. Below we show how this lack of rationality/information strengthens our results.

Our analysis is related to the following strands of literature. Howard et al. (2015) discuss a number of explanations for high drug prices that are found in the literature. However, these cannot fully explain why prices would exceed treatment value. To illustrate, an explanation that is often mentioned is that with health insurance people want a treatment, no matter what the cost since the insurer pays (most of) the price. This is true ex post: once I have insurance, the effects of the cost of treatment are reduced for me. Economists tend to refer to this as moral hazard. However, why would I buy insurance coverage for a treatment that costs more than the benefit it provides? Dropping such a treatment from the contract leads to a bigger reduction in the premium than the loss in expected utility. Hence, an insurer –whether or not it has market power– benefits from removing such treatments from its insurance contract. This threat of not being covered by an insurance contract limits the price a pharmaceutical company can ask for patented drugs.

Similarly, Garrison and Towse (2017) mention the high sunk costs of R&D to explain high treatment prices. Although high fixed/sunk costs can explain high prices in competitive markets (by limiting entry), this mechanism is not obvious for a monopoly market where a firm is protected by a patent. Since a monopolist tries to appropriate most or all of the surplus from its customers, its sunk fixed costs are not directly relevant for setting prices. And also here, if the treatment is too expensive, it should be dropped by the insurer from the contract.[5]

When comparing drug prices between countries, a number of institutional features are often mentioned. One is whether the health insurance market is run by the government or via a market with private insurers. To illustrate, in public healthcare systems politicians tend to find it hard to refuse reimbursement of a treatment on the basis of its price; this creates upward pressure on prices.[6] A classic example is the coverage of proton beam therapy in the NHS before any cost-benefit analysis was done: "proton beam therapy has not been the subject of a technology appraisal by the National Institute for Health and Clinical Excellence" and at the time there

[5]In free market systems, the insurer can decide what to cover or not in its contracts. In regulated market systems, like the Netherlands, the government prescribes which conditions need to be covered by basic insurance but does not define which treatments need to be covered.

[6]Note that citizens may actually agree with politicians' point of view here (McCabe et al., 2005).

was no "reliable, objective evidence that proton beam treatment improved clinical outcomes" (Hawkes, 2012). Economists have compared proton beam therapy to the death star (Langreth, 2012).

Other explanations for cross country price differences include GDP per head (richer countries tend to pay more), whether the drug prices are bargained over centrally or by insurers individually, country regulations regarding reference pricing etc. (Danzon and Taylor, 2010). What we add to this literature is that the extent of selection problems in the health insurance market affect drug prices. To the best of our knowledge there is no direct evidence on this relation. However, we do know that the US tends to pay far more for patented drugs than other countries and that price differences between European countries are relatively small (Mulcahy et al., 2021; Young et al., 2017). Adverse selection problems are bigger in the US (with many people going without insurance) than in Europe where all countries have relatively high insurance coverage (often mandatory or automatic insurance) and in this sense reduced selection problems.

Kamphorst and Karamychev (2021) focus on the relation between prices and disease rarity: the lower the prevalence of a disease, the higher the prices for drugs treating it. In the context of orphan drugs, this is sometimes referred to as payers valuing rarity (Medic et al., 2017; Messori et al., 2010). They show that if the insurance premium is determined by the budget constraint of the marginal insured, an inverse relation exists between disease prevalence and drug price. We offer a complementary explanation for this inverse relation via the selection incentives on the health insurance market.

In terms of the mechanism design literature, our focus on a binding incentive compatibility constraint for low types is in line with the countervailing incentives literature (Lewis and Sappington, 1989). This literature considers type dependent outside options, which we generate through differing demand elasticities for different types. Although we use health insurance and adverse selection to illustrate our model, the mechanism applies in any screening model where the incentive compatibility constraint is binding for the low type. It is then optimal to distort the high type's allocation upward (Laffont and Martimort, 2002). One way to get the upward distortion is to add a feature to the allocation at a price that exceeds the feature's value. Boone and Schottmüller (2017) and references therein consider adverse selection markets with a violation of single crossing. There it also happens that the incentive compatibility constraint of the low risk type is binding. These papers then analyze the equilibrium outcomes and the welfare properties of the equilibria, but

not the effect on price and innovation incentives in an upstream market.

Finally, pricing above value can happen in an industrial organization context where there are negative externalities. In the context of R&D, this externality can be the business stealing effect. An early analysis of how an R&D lab can extract more than the social value of an innovation is Katz and Shapiro (1986). They focus on the licensing mechanism that can be used to extract the maximum profit out of an innovation. This literature on auction mechanisms to maximize revenue in the context of a negative externality was further developed by Jehiel et al. (1996). Papers in this literature do not consider innovation incentives for firms selling to an adverse selection market.

Our paper is organized as follows. First, we illustrate our main effect in a simple insurer monopoly model. Then we present our general framework with competing insurers. We argue that supra profits appear when the elasticity/competition difference between the segments of the two risk types exceeds their expected cost difference. We use a Hotelling competition model to illustrate that parameter values exist under which this happens. Then we present our results in a general framework. We discuss the effects of bounded rationality and conclude with a discussion of policy implications.

2 Simple example

To see how it is possible at all that an insurer pays more for a treatment than the treatment's value to the insured, consider the following simple example with two treatments and a monopolist insurer. The example introduces the notation and illustrates the mechanics of the result: the high type's allocation is distorted upwards which makes it less attractive to the low type and hence reduces the low type's information rent.

Denote the two treatments 1 and 2; both treatments are produced at marginal costs normalized to zero: $c_1 = c_2 = 0$. Treatment 1 is under patent while 2 is off patent and sold by competing firms at a price equal to marginal costs. The values of these treatments are given by v_1, v_2 resp. and are the same for each patient. Value v_i captures things like life years gained, improvement in quality of life, increased productivity etc. (Garrison and Towse, 2017). Although it is not straightforward to measure this in practice, conceptually the value of a treatment is well defined. We aim to show that pharmaceutical companies can profitably charge a price in excess of this value and still be covered by insurance plans.

The monopolist insurer sells insurance at a premium σ and faces a customer who can be either of type l (probability ϕ) or type h (probability $1-\phi$). Type $k = l, h$ needs treatments 1,2 with probability ψ_{1k}, ψ_{2k}. We assume single crossing: $\psi_{ih} \geq \psi_{il}$ for $i = 1, 2$ and to simplify notation in this example assume that $\psi_{2h} = \psi_{2l} = \psi_2$. Hence, the high risk consumer has a strictly higher probability of needing treatment 1: $\psi_{1h} > \psi_{1l}$.

We assume that the consumer buys insurance to get access to the treatment(s). That is, without insurance, the consumer goes without treatment.[7] It is well documented that people without health insurance tend to forgo treatment as they have difficulty financing it. These access issues have been stressed both in the popular press (Cohn, 2007) and in academic journals (Nyman, 1999; Schoen et al., 2008, 2010). Many governments are concerned about health consumption inequality caused by income differences and design policies to make healthcare accessible to low income families (Schokkaert and van de Voorde, 2011). In terms of modeling, risk aversion would complicate the expressions for utility (by introducing a non-linearity) without adding insight.

The insurer offers two contracts (which can be identical in case of a pooling outcome), each contract aimed at a consumer type. We write the value/utility of the contract for type $k = 1, 2$ as follows:

$$u_k = \alpha\psi_{1k}x_{1k}v_1 + \psi_2 x_{2k}v_2 - \sigma_k \tag{1}$$

where $x_{ik} \in [0, 1]$ denotes the probability that treatment i is covered by contract k and σ_k denotes the price/premium of contract k. Value v_i denotes the utility of receiving the treatment in case the consumer needs it (with probability ψ_{ik}) compared to not receiving this treatment. Note that falling ill in itself can cause a disutility for the individual. Taking this into account would add a constant to the expression in (1) which we leave out to ease notation.

The use of $\alpha \in [0, 1]$ is a "technical trick". Strictly speaking, it denotes the probability that treatment 1 is available to the insurer.[8] We think of α as being equal to one throughout the paper; but can deduce the value of treatment 1 for the insurer by considering the effect on the insurer's profits of a small decrease in α. This gives us the convenience of differentiating instead of taking the discrete difference

[7] To simplify notation, we normalized $c_1 = c_2 = 0$. For the access to care interpretation, think of c_1, c_2 being high enough that a patient without insurance cannot afford them even at cost price.

[8] In this sense, x_{ik} denotes the probability that i is covered conditional on it being available to the insurer.

between profits with ($\alpha = 1$) and without ($\alpha = 0$) treatment 1 being covered. In particular, to understand how $p_1 > v_1$ is possible, we will derive conditions under which the insurer's profits are strictly increasing in α *even if* $p_1 = v_1$.

The incentive compatibility (IC) constraints for these contracts can be written as follows.

$$u_h \geq \alpha\psi_{1h}x_{1l}v_1 + \psi_2 x_{2l}v_2 - \sigma_l \tag{2}$$

$$u_l \geq \alpha\psi_{1l}x_{1h}v_1 + \psi_2 x_{2h}v_2 - \sigma_h \tag{3}$$

That is, IC_h implies that the high type is better off choosing the high contract (yielding utility u_h) than to buy the low type's contract (which yields her utility equal to the right hand side of (2)). And, similarly, IC_l implies that the low type is better off buying the l contract –yielding u_l – than buying the h contract which yields her utility equal to the right hand side of (3).

The individual rationality (IR) constraints make sure that each type is better off buying a contract than not buying a contract at all; the constraints IR_h, IR_l can be written as:

$$u_h \geq \bar{u}_h$$

$$u_l \geq \bar{u}_l$$

where $\bar{u}_k$ denotes the utility of type k's outside option of not buying insurance at all. In the main model below, $\bar{u}_k$ is endogenized by the outside option of buying from a competing insurer. Here we simply assume that the values $\bar{u}_h, \bar{u}_l$ are exogenously given.

As we just want to show that $p_1 > v_1$ is possible, we simply assume that IC_l and IR_h are binding as is done in the literature on countervailing incentives (Lewis and Sappington, 1989). Below we derive this as an equilibrium outcome in a model where insurers compete.[9]

Then by combining (1) and (3), we write the binding IC_l as

$$u_l = \bar{u}_h - \alpha v_1 x_{1h}(\psi_{1h} - \psi_{1l})$$

We write the premium as

$$\sigma_k = \alpha\psi_{1k}x_{1k}v_1 + \psi_2 x_{2k}v_2 - u_k$$

[9]Note that our focus on a binding IC_l implies that we do not need to introduce co-payments which help separate types when IC_h is binding (Rothschild and Stiglitz, 1976).

Further, treatment 1 is bought from its producer at price p_1 and for treatment 2 we assumed that competition leads to marginal cost pricing, $p_2 = 0$. Hence the monopoly insurer's profit can be written as

$$\begin{aligned}\Pi = &\phi(\alpha\psi_{1l}x_{1l}(v_1 - p_1) + \psi_2 x_{2l} v_2 - u_l) \\ &+ (1-\phi)(\alpha\psi_{1h}x_{1h}(v_1 - p_1) + \psi_2 x_{2h} v_2 - \bar{u}_h) \\ &+ \lambda_l(u_l - \bar{u}_h + \alpha v_1 x_{1h}(\psi_{1h} - \psi_{1l}))\end{aligned}$$

where λ_l denotes the Lagrange multiplier on the IC_l constraint. Profits consist of share of low [high] types $\phi[1-\phi]$ multiplied by the difference between the premium σ and the expected cost of the contract $\alpha\psi_{1l}x_{1l}p_1[\alpha\psi_{1h}x_{1h}p_1]$. From this it is straightforward to derive that

$$\left.\frac{d\Pi}{dx_{1h}}\right|_{p_1=v_1} = \alpha\lambda_l v_1(\psi_{1h} - \psi_{1l}) > 0$$

as $\psi_{1h} > \psi_{1l}$. It follows that the corner solution with $x_{1h} = 1$ maximizes profits and we find that

$$\left.\frac{d\Pi}{d\alpha}\right|_{p_1=v_1} = \lambda_l v_1(\psi_{1h} - \psi_{1l}) > 0$$

In words, even if $p_1 = v_1$, the insurer's profits are strictly increasing in α. This implies that the producer of treatment 1 can ask more than $p_1 = v_1$ –the final consumers' valuation of the treatment– and the insurer will still cover this treatment in its health insurance contract. For our purposes here, there is no need to characterize the solution further.

The reason why the insurer is willing to cover a treatment which is sold at a price in excess of its value to consumers is that the treatment helps to reduce the low type's information rent. In other words, the treatment has a value for the insurer in addition to the utility created by the treatment for the insured. Covering the treatment relaxes the IC_l constraint. The value of relaxing IC_l is given by its shadow price $\lambda_l > 0$. Since the h type is more likely to need the treatment than the l type, covering the treatment makes the h contract less attractive to the l type. This allows the insurer to increase σ_l and profits.

Note the role of generic drugs, here captured by treatment 2, being sold at a price below their value ($p_2 < v_2$). Generic drugs are needed for our argument to "create space" for patented firms to charge prices above their treatments' values. To see this, consider the case where $p_2 = v_2$ and the outside option is normalized to $\bar{u}_k = 0$. Then the IR constraint is of the form $\alpha\psi_{1k}x_{1k}v_1 + \psi_{2k}x_{2k}v_2 - \sigma_k \geq 0$. Insurers charge a price that (at least) covers their costs: $\sigma_k \geq \alpha\psi_{1k}x_{1k}p_1 + \psi_{2k}x_{2k}v_2$.

Thus we find $p_1 \leq v_1$ in case $p_2 \geq v_2$. Without (generic) drugs being sold at a price below value we cannot have $p_1 > v_1$.[10]

Next we introduce our general framework to analyze the effects of insurer competition and to endogenize that IC_l is binding. We will see that price exceeding value can happen in both pooling and separating equilibria.

3 Framework

Let P denote the set of treatments that are currently under patent and O the set of treatments where the patent has run out ("open" as in open source). To simplify the exposition we assume that $\psi_{il} < \psi_{ih}$ for each $i \in P$.[11] For $j \in O$ we assume $\psi_{jl} \leq \psi_{jh}$. This ensures that single crossing is satisfied in our set up.[12]

Insurers $\iota \in \{a, b, c, ..., n\}$ offer contracts $((x^\iota_{il}, x^\iota_{jl}, \sigma^\iota_l)_{i\in P, j\in O}, (x^\iota_{ih}, x^\iota_{jh}, \sigma^\iota_h)_{i\in P, j\in O})$, where the first contract is intended for type l and the second for h. A contract specifies the probability $x \in [0, 1]$ (below we have that x is either 0 or 1, except when explicitly mentioned) that a treatment is covered and a premium σ.

Then the utility for type $k = l, h$ of buying the contract meant for k is given by

$$u_k = \sum_{i\in P} \alpha_i \psi_{ik} x_{ik} v_i + \sum_{j\in O} \psi_{jk} x_{jk} v_j - \sigma_k \tag{4}$$

where we drop the ι superscript to ease notation. Utility consists of the probability that consumer k needs the treatment $\psi_{ik}[\psi_{jk}]$ multiplied by the probability that the treatment is covered by the contract $\alpha_i x_{ik}[x_{jk}]$ times the value of the treatment $v_i[v_j]$ minus the premium σ_k. As above, $\alpha_i \in [0, 1]$ is a technical convenience to determine the value of treatment i for the insurer using differentiation. The assumption is that the agent cannot afford the treatments without insurance and hence v denotes the value of the treatment compared to the best affordable (i.e. without insurance) alternative treatment.

[10] Note that this is different with risk averse agents. Insurance then generates value beyond the value of treatment by reducing risks. Some patented treatments may try to capture part of this insurance rent by charging high prices. This strengthens the argument that it is profitable to charge $p_1 > v_1$.

[11] Otherwise we need to distinguish treatments i where the inequality in $\psi_{il} \leq \psi_{ih}$ is strict or not below. This is tedious while not adding to our understanding of the problem.

[12] See Boone and Schottmüller (2017) for an analysis of the health insurance market where single crossing is violated.

The incentive compatibility constraints can be written as

$$u_l \geq u_h - \sum_{i \in P} \alpha_i x_{ih} v_i (\psi_{ih} - \psi_{il}) - \sum_{j \in O} x_{jh} v_j (\psi_{jh} - \psi_{jl}) \qquad (IC_l)$$

$$u_h \geq u_l + \sum_{i \in P} \alpha_i x_{il} v_i (\psi_{ih} - \psi_{il}) + \sum_{j \in O} x_{jl} v_j (\psi_{jh} - \psi_{jl}) \qquad (IC_h)$$

As we have insurer competition here, we assume that competition is intense enough that each type's IR constraint is satisfied (i.e. σ_k is low enough).

To simplify notation, we assume that treatments in the set O are sold under Bertrand competition with price equal to marginal costs, which we normalize to zero: $p_j = c_j = 0$ for $j \in O$. We use equation (4) to write

$$\sigma_k = \sum_{i \in P} \alpha_i \psi_{ik} x_{ik} v_i + \sum_{j \in O} \psi_{jk} x_{jk} v_j - u_k \qquad (5)$$

The marginal cost of contract k for the insurer is given by the expected costs of the contract where costs are determined by treatment prices.

$$c_k = \sum_{i \in P} \alpha_i \psi_{ik} x_{ik} p_i \qquad (6)$$

As the price of $j \in O$ treatments is normalized to 0, the expected cost of type k is given by the sum over all patented treatments of the probability that k will receive treatment times the price p_i of treatment. Insurer ι's profits can then be written as follows:

$$\begin{aligned}
\Pi^\iota = {} & \phi q^\iota(u_l^\iota, u_l^{-\iota}, \theta_l)\Big(\sum_{i \in P} \alpha_i \psi_{il} x_{il}^\iota (v_i - p_i) + \sum_{j \in O} \psi_{jl} x_{jl}^\iota v_j - u_l^\iota\Big) \\
& + (1-\phi) q^\iota(u_h^\iota, u_h^{-\iota}, \theta_h)\Big(\sum_{i \in P} \alpha_i \psi_{ih} x_{ih}^\iota (v_i - p_i) + \sum_{j \in O} \psi_{jh} x_{jh}^\iota v_j - u_h^\iota\Big) \\
& + \lambda_l\Big(u_l^\iota - u_h^\iota + \sum_{i \in P} \alpha_i x_{ih}^\iota v_i (\psi_{ih} - \psi_{il}) + \sum_{j \in O} x_{jh}^\iota v_j (\psi_{jh} - \psi_{jl})\Big) \\
& + \lambda_h\Big(u_h^\iota - u_l^\iota - \sum_{i \in P} \alpha_i x_{il}^\iota v_i (\psi_{ih} - \psi_{il}) - \sum_{j \in O} x_{jl}^\iota v_j (\psi_{jh} - \psi_{jl})\Big)
\end{aligned} \qquad (7)$$

where $q^\iota(u_k^\iota, u_k^{-\iota}, \theta_k)$ denotes the market share of the insurer's contract on the $k = l, h$ market as a function of ι's own utility offered u_k^ι, the utilities offered by ι's competitors $u_k^{-\iota}$ and of factors θ_k affecting the demand elasticity like the competition intensity on this market. We denote the Lagrange multiplier on constraint (IC_k) by λ_k. We assume that q^ι is increasing and concave in ι's offered utility and decreasing in the utility levels offered by ι's competitors.

An insurer maximizes profits over coverage and utility (i.e. the premium): x_{ik}, x_{jk}, u_k for $i \in P, j \in O, k \in \{l, h\}$. We can already see the following properties of the solution.

Lemma 1 *Assume $p_i = v_i$ and $\alpha_i \in \langle 0, 1]$ for $i \in P$, then we have:*

- $\lambda_h > 0$ *implies* $x_{il} = 0$*;*
- *it cannot be the case that both IC constraints bind;*
- $\lambda_l > 0$ *implies that* $\lambda_h = 0$*,* $x_{ih} = x_{jh} = x_{jl} = 1$ *and* $x_{il} \in [0, 1]$*.*

Finally, $p_i > v_i$ *implies* $x_{il} = 0$*.*

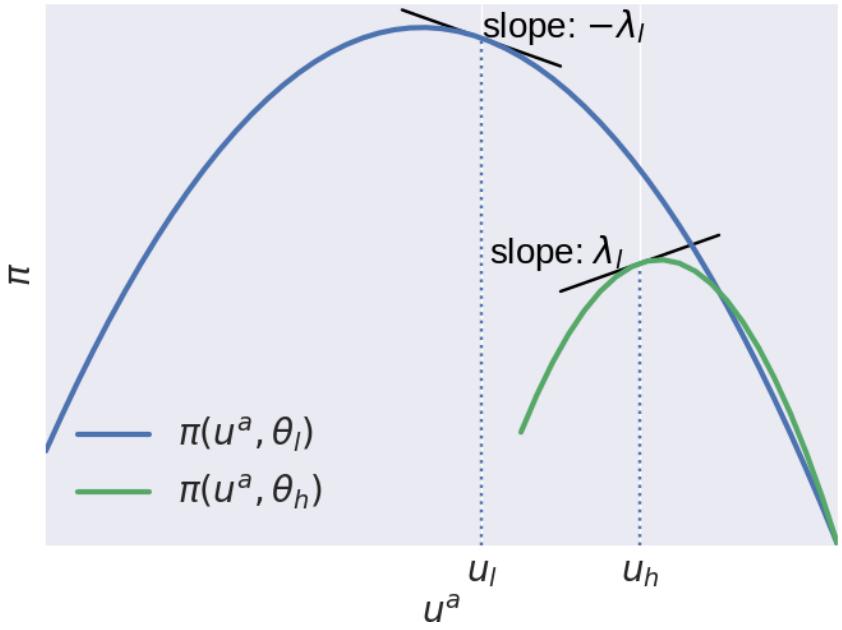

Figure 1: Insurer a's profits $\pi(u^a, \theta_l), \pi(u^a, \theta_h)$ as a function of u^a.

If (IC_h) is binding, coverage for the low type is distorted downwards. This is the distortion in a standard insurance model where types only differ in costs (Rothschild and Stiglitz, 1976). However, we focus on the opposite case with $\lambda_l > \lambda_h = 0$. Because single crossing is satisfied in our model, it cannot be the case that each type wants to mimic the other type; hence only one IC constraint is binding. With $p_i \leq v_i, p_j \leq v_j$, both contracts can cover all treatments. But $p_i > v_i$ implies that the low type's contract does not cover treatment i. For the low type the treatment price cannot exceed value. This can be seen as follows: $\lambda_h = 0$ implies

that $\partial\Pi/\partial x_{il} = \phi q \alpha_i \psi_{il}(v_i - p_i) < 0$ for $p_i > v_i$. Hence, optimal $x_{il} = 0$ (corner solution) in this case. Intuitively, l consumers do not want to pay more for treatment i than it is worth and x_{il} does not help the insurer to separate consumer types. With $\lambda_h = 0$ no one tries to mimic l. Thus the insurer implements the first best outcome: do not cover a treatment with a price above its value. But $x_{ih} = 1$ is possible in case $p_i > v_i$, as we will see below.

We focus on the case where (IC_l) is binding, while (IC_h) is not. We show in Section 4 that such a case indeed exists in a model with Hotelling competition. Figure 1 illustrates how this can happen.[13] As illustrated in the figure, the l market is more profitable than the h market in that profits on the l market exceed h profits over the relevant range of u. This is intuitive since l customers are cheaper in expectation than h types. In addition, for the insurer to attract customers on the h market, a larger utility level needs to be left to its customers than on the l market. This can be due to the fact that h customers have more experience with treatments and are better able to compare the values offered by different insurance plans. Or h types (high risk/low health status) tend to have low income and hence pay a relatively low premium σ_h leading to high u_h. Hence, the profit maximizing level of u_h exceeds that of u_l. But the utility left to h customers cannot be too high, because this would induce l customers to buy the h contract.[14] The distance $u_h - u_l$ in the figure is determined by IC_l holding with equality.

In the figure, u_l denotes the symmetric equilibrium outcome on the l market and u_h on the h market. The former is chosen higher than its profit maximizing level and the latter lower, if one would consider each market in isolation. This is caused by the constraint that the difference between u_h and u_l cannot be too big. If this difference would be bigger, the contracts would no longer be incentive compatible and everyone would buy the h contract. The first order conditions for u_h, u_l show that the marginal profits $\partial\pi/\partial u_k^a$ on each market should be equal (in absolute value) and equal to λ_l. At the margin, the loss in profits of not being able to lower u_l equals the loss of not being able to increase u_h.

[13] This figure is based on the example in Section 4 with parameter values: $\phi = 0.5, v_1(\psi_{1h} - \psi_{1l}) = 1, \psi_2 v_2 = 5, t_h = 1, t_l = 3, \eta = 1, y_l = y_h = 0$.

[14] If the h market would become too small, an insurer can decide to stop serving this market segment. We assume throughout the paper that the h market is big enough that the insurer keeps on selling to h types.

4 Competition example

Above we work in a framework with elastic demand on two market segments and IC constraints on both segments. We focus on the case where the IC constraint of the low type is binding ($\lambda_l > 0$). This section presents a health insurance model where demand elasticities differ on the market segments due to differences in travel costs and income. We present example parameter values such that indeed IC_l is binding and IC_h is slack. As this section aims to illustrate results, it can be skipped on first reading.

Since we want to show that something is possible, we simplify by considering (only) two treatments. Treatment 1 is under patent, treatment 2 is not. As in the simple example above, we focus on $p_1 = v_1$, $p_2 = c_2 = 0$ and $\psi_{2l} = \psi_{2h} = \psi_2$. To interpret demand elasticities below, we introduce a simple way to capture income effects which we explain and motivate below. If insurance contract k from insurer a gives utility u_k^a, then an insured's utility is given by

$$(y_k + u_k^a)^\eta \tag{8}$$

with $y_k > 0, \eta \in \langle 0, 1]$. There are two insurers competing on a Hotelling beach of length 1 with consumers distributed uniformly (with density 1) along the beach. Insurer a is on the left hand side of the beach and insurer b on the right hand side. A fraction ϕ of consumers (on each location) is type l and a fraction $1-\phi$ is h. The travel cost over the Hotelling beach for type k is denoted by $t_k, k = l, h$. If insurer a offers utility u_k^a and insurer b offers u_k^b, then a's market share q_k is given by the indifferent consumer at position $q_k \in [0, 1]$ on the k market:

$$(y_k + u_k^a)^\eta - t_k q_k = (y_k + u_k^b)^\eta - t_k(1 - q_k)$$

For the indifferent consumer the utility from buying from a minus the travel cost to a equals the utility from b minus the travel cost to b. Hence, a's market share can be written as

$$q_k = \frac{1}{2} + \frac{(y_k + u_k^a)^\eta - (y_k + u_k^b)^\eta}{2t_k}$$

Then we can define type k's demand elasticity as

$$\varepsilon_k = \left| \frac{\partial q_k}{\partial \sigma_k^a} \frac{\sigma_k^a}{q_k} \right| = \frac{\eta}{t_k} \frac{\sigma_k}{(y_k + u_k)^{1-\eta}} \tag{9}$$

in symmetric equilibrium with $u_k^a = u_k^b = u_k, \sigma_k^a = \sigma_k^b = \sigma_k$.

Before interpreting these results, let us motivate the functional forms used. The Hotelling model of competition is fairly standard (Tirole, 1988). It allows us to

model competition with inelastic market demand but elastic demand for the firm. The inelastic demand is useful here for two reasons. First, it is easy to combine it with IC constraints for the two customer types. Second, it is straightforward to compare the value of the treatment v_i with the price of the treatment p_i charged by the manufacturer. If consumers have differing values for v_i (elastic demand), it is not clear to which value v_i the price p_i should be compared to claim that pricing is excessive (the average v_i, the median or maximum v_i?). With this set-up we can transparently make the claim that pricing is excessive in the sense that the price exceeds the treatment's value for each customer.

If both insurers offer the same utility level, a's market share equals $1/2$. If $u_k^a > u_k^b$, a is relatively more attractive and its market share exceeds $1/2$. The pace at which a's market share increases with the difference between u_k^a and u_k^b is determined by travel cost t_k and parameter η. The lower t_k, the more market share responds to utility differences offered by insurers and the more competitive the market is.

A similar logic is used when choosing the parameterization in equation (8). Because u_k^a enters linearly in this expression, we can use the standard IC constraints and compare v_i to the price p_i that the manufacturer charges. We interpret y_k as type k's income that is spent on other goods and services. It acts to make overall utility less elastic as income increases in case $\eta < 1$. Someone with a low income is more sensitive to changes in, say, the insurance premium than someone with a high income (Ho et al., 2017).

As equation (9) shows, the demand elasticity of type k decreases with t_k: the more competitive the Hotelling market (lower t_k), the more elastic is type k's demand for insurer a's contract. Lower y_k also leads to more elastic demand in case $\eta < 1$.

We finish the section with an example where indeed IC_l is binding. As the goal is to show that something is *possible* we simplify further by assuming $\eta = 1, y_k = 0$ for $k = l, h$. We first assume that IC_h is slack (and check later that this assumption is correct; see the proof of lemma 2). With $p_1 = v_1$ and only one treatment in each of the sets P, O, we write equation (7) as follows:

$$\begin{aligned}\max_{u_l^a, u_h^a, x_{1h}^a, x_{2l}^a, x_{2h}^a} \; & \phi \left(\frac{1}{2} + \frac{u_l^a - u_l^b}{2t_l} \right) (\psi_2 x_{2l}^a v_2 - u_l^a) \\ & + (1-\phi) \left(\frac{1}{2} + \frac{u_h^a - u_h^b}{2t_h} \right) (\psi_2 x_{2h}^a v_2 - u_h^a) \\ & + \lambda_l (u_l^a - u_h^a + \alpha v_1 x_{1h}^a (\psi_{1h} - \psi_{1l}))\end{aligned} \tag{10}$$

We derive the following equilibrium outcome.

Lemma 2 *In the Hotelling model, assume that* $p_1 = v_1$ *and* $t_l - t_h > v_1(\psi_{1h} - \psi_{1l})$. *Then in equilibrium, it is the case that* $x_{1h} = x_{2h} = x_{2l} = 1$, $x_{1l} \in [0, 1]$ *and*

$$
\begin{aligned}
u_h &= \psi_2 v_2 + \frac{\frac{\phi}{t_l}}{\frac{\phi}{t_l} + \frac{1-\phi}{t_h}} \alpha v_1 (\psi_{1h} - \psi_{1l}) - \frac{1}{\frac{\phi}{t_l} + \frac{1-\phi}{t_h}} \\
u_l &= \psi_2 v_2 - \frac{\frac{1-\phi}{t_h}}{\frac{\phi}{t_l} + \frac{1-\phi}{t_h}} \alpha v_1 (\psi_{1h} - \psi_{1l}) - \frac{1}{\frac{\phi}{t_l} + \frac{1-\phi}{t_h}} \\
\lambda_l &= \frac{1}{2} \frac{t_l - t_h - \alpha v_1 (\psi_{1h} - \psi_{1l})}{\frac{t_h}{1-\phi} + \frac{t_l}{\phi}} > 0 \\
\left. \frac{d\Pi}{d\alpha} \right|_{p_1 = v_1} &= (\psi_{1h} - \psi_{1l}) v_1 \lambda_l > 0 \qquad (11)
\end{aligned}
$$

The assumption in the lemma is that the h market is (sufficiently) more competitive than the l market ($t_h < t_l$). Under this assumption an equilibrium exists where IC_l is binding and we derive the utility levels for both types (offered by both firms). These utility levels can be understood as follows. With $p_1 = v_1$, the price of treatment 1 equals the utility it generates. Hence, its net utility for the insured equals zero. The utility generated by the insurance contract is due to the cheap ($p_2 = c_2 \leq v_2$) treatment 2. A well known property of the Hotelling model is the following. If there would be only one agent type k, utility would equal $\psi_2 v_2 - t_k$. The higher t_k (the less competitive the market), the higher the premium and the lower utility would be. However, here we have two types and their markets are linked. Hence, we do not simply subtract t but the (weighted) harmonic mean of t_h, t_l: $1/(\phi/t_l + (1-\phi)/t_h)$. We do not need to require $u_h = u_l$; we are allowed to differentiate $u_h - u_l \geq 0$ as long as $u_h - u_l \leq \alpha v_1 x(\psi_{1h} - \psi_{1l})$ to satisfy (IC_l).

Hence, we increase u_h and we reduce u_l in a way that causes (IC_l) to hold with equality. The relative weights of this increase, decrease resp. are determined by the market size and competition differences between the market segments. Finally, the shadow price λ_l of the (IC_l) constraint is positive under the assumptions that we made: the l type wants to mimic the h type who gets a relatively better deal as the h market is more competitive than the l market.

Even when treatment 1 is priced at its value ($p_1 = v_1$), we find that a reduction in α strictly reduces profits. With $p_1 = v_1$, the consumers are indifferent whether treatment 1 is covered or not. But a fall in α strictly reduces insurers' profits. Hence, insurers are willing to pay more than $p_1 = v_1$ to have treatment 1 covered in their contracts. Treatment 1 generates value in excess of the utility of its treatment, v_1. The intuition is that treatment 1 helps the insurers to reduce information rents.

This is why it is so profitable to sell to an adverse selection market. In the proof of the lemma, we also solve the asymmetric case where $\alpha^a < \alpha^b = 1$: if firm 1 tries to increase the rents from selling its treatment, it can threaten insurer a to reduce its access to treatment 1 ($\alpha^a < 1$) while competing with an insurer b that has full access to this treatment ($\alpha^b = 1$). It turns out that the expressions are similar to the ones given above and they have the same properties. So here we focus on the simpler symmetric equations.

We also show in the appendix that λ_l is increasing in t_l and decreasing in t_h. In words, an increase in competition in the h market makes IC_l "more binding".

For reference below, the condition in the lemma can be written as $t_l - t_h > c_h - c_l$ where the costs for the insurer of type k equals $c_k = \psi_{1k} p_1 = \psi_{1k} v_1$. That is, the difference in competition intensities between the two markets exceeds the difference in expected costs. What happens if the condition is not satisfied; that is, $c_h - c_l \geq t_l - t_h$? In this case, we can solve the insurer's problem (10) without the IC_l constraint. It is routine to verify that this yields $u_k = \psi_2 v_2 - t_k$. Substituting this solution in the IC_l constraint shows that it is indeed satisfied. Hence, if the difference in competition intensities t_l, t_h is not big enough compared to the cost difference, the insurer sets a premium on the l contract that is so low (due to the low cost) that the l type does not want to mimic the h type and $\lambda_l = 0$. If the cost difference increases further, the h contract becomes so expensive that the high type wants to mimic the low type. This is the more traditional adverse selection problem (Rothschild and Stiglitz, 1976). Only when sufficient market power on the l market compensates for the lower c_l, the h contract becomes attractive to the l type.

5 Supra profits

We consider two contractual arrangements that lead to a profit for the patent holder that exceeds the value of its innovation; we call these extra rents "supra profits". First, we consider the innovator using a two-part tariff which captures most non-linear pricing schemes. Then we consider the case where the innovator can only use a linear fee (a price per unit).

5.1 two-part tariff

The easiest way to see the main effects of this paper is to assume that innovators sell treatments to insurers using two-part tariffs. It turns out that the intuitions we

find here, carry over to the case of linear pricing.

To characterize the optimal prices set by innovator i, we first derive the insurers' equilibrium response to the linear part p_i of the tariff. The insurers set a premium σ which can be different for the h and l markets. Or equivalently (see equation (5)) insurers set utility levels u_l, u_h.

The first order condition for maximizing (7) with respect to u_h can be written as:

$$\begin{aligned}\lambda_l &= (1-\phi)\left(-q_h + (\sigma_h - c_h)\frac{\partial q_h}{\partial u_h}\right) \\ &= -(1-\phi)q_h\left(1 + \frac{\sigma_h - c_h}{\sigma_h}\frac{\partial q_h}{\partial \sigma_h}\frac{\sigma_h}{q_h}\right) \\ &= (1-\phi)q_h(\mu_h\varepsilon_h - 1) \end{aligned} \tag{12}$$

where we use that $\partial q_h/\partial u_h = -\partial q_h/\partial \sigma_h$, c_h is given by (6) and the insurer's mark-up equals $\mu_h = (\sigma_h - c_h)/\sigma_h$. As above, we define the elasticity of the h type with respect to the premium σ_h as $\varepsilon_h = |\frac{\partial q_h}{\partial \sigma_h}\frac{\sigma_h}{q_h}|$. The reason why we write the first order condition like this is that we assume that ε_h and ε_l are (locally) constant when we do comparative statics.

We assume that $\varepsilon_h > \varepsilon_l$; one can think of two reasons why this is the case. First, high risk agents are likely to have experienced more care in the past and hence understand better what the different insurance contracts offer. This makes it easier for them to compare the insurance contracts (captured by $t_l \leq t_h$ in equation (9)). Second, if h agents with low health status tend to earn a lower income, they may have a higher incentive than l agents to find the better deal (captured by $y_l > y_h$ in equation (9)).

The first order condition for maximizing Π^ι with respect to u_l can be written as:

$$-\lambda_l = \phi q_l(\mu_l\varepsilon_l - 1) \tag{13}$$

where $\mu_l = (\sigma_l - c_l)/\sigma_l$.

In symmetric equilibrium ($q_l = q_h = 1/n$ where n denotes the number of insurers) adding the first order conditions for u_h and u_l, we find that

$$\phi(\mu_l\varepsilon_l - 1) + (1-\phi)(\mu_h\varepsilon_h - 1) = 0 \tag{14}$$

If both markets would be served independently (i.e. without being linked via an IC constraint), the insurer would set its price cost margin equal to $\mu = 1/\varepsilon$; the well known Lerner expression for a profit maximizing mark-up. Because the markets

are linked via (IC_l), the weighted sum of these two terms equals zero. That is the mark-up is too high in one market (compared to $\mu = 1/\varepsilon$) and this is compensated by a mark-up which is too low in the other market.

Here we consider the case where R&D firm i sets $p_i = v_i$ and uses the fixed part of the tariff to appropriate the supra profits. Lemma 1 implies that $x_{ih} = x_{il} = x_{jh} = x_{jl} = 1$. Hence, we have a pooling contract.[15] Therefore, μ_l and μ_h in (14) are based on the same premium σ. Thus, we can solve for σ. Since we have $\sigma_l = \sigma_h = \sigma$, we have $\mu_l > \mu_h$ because $c_h > c_l$.

Lemma 3 *With $p_i \leq v_i$, we find the following expressions for the premium and the mark-ups on the two markets:*

$$\sigma = \frac{\phi\varepsilon_l c_l + (1-\phi)\varepsilon_h c_h}{\phi\varepsilon_l + (1-\phi)\varepsilon_h - 1} \tag{15}$$

$$\mu_h = \frac{c_h - \phi\varepsilon_l(c_h - c_l)}{\phi\varepsilon_l c_l + (1-\phi)\varepsilon_h c_h} \tag{16}$$

$$\mu_l = \frac{c_l + (1-\phi)\varepsilon_h(c_h - c_l)}{\phi\varepsilon_l c_l + (1-\phi)\varepsilon_h c_h}$$

$$\lambda_l = \frac{1}{n}\phi(1-\phi)\left(\frac{\varepsilon_h c_h(1-\varepsilon_l) - \varepsilon_l c_l(1-\varepsilon_h)}{\phi c_l\varepsilon_l + (1-\phi)c_h\varepsilon_h}\right) \tag{17}$$

As a proxy of the supra profits that i can appropriate, consider the following derivative of equation (7):

$$\tau_i = \left.\frac{d\Pi^\iota}{d\alpha_i}\right|_{p_i=v_i} = \lambda_l v_i \Delta\psi_i \tag{18}$$

Even if treatment i is priced at consumer value $p_i = v_i$, reducing α_i still strictly reduces insurers' profits with $\lambda_l > 0$. This part of the supra profits is captured using the fixed tariff t_i. The level of t_i is directly related to τ_i.

The following assumption ensurers that we are in the relevant parameter space that allows for supra profits. First, we need that (IC_l) is binding: $\lambda_l > 0$. Further, for the insurers' optimization problem to be well defined, we need that the "average elasticity" on the two markets exceeds 1.[16] Throughout this paper, we make the following assumptions such that both conditions are satisfied.

[15] Given that we have a pooling contract, we could have solved the model by assuming the pooling contract from the start and not bother with IC constraints. Our approach has the following advantages. First, we can prove under which conditions pooling is optimal. Second, we use the same approach to derive the pooling and the separating outcomes. Third, we get the Lagrange multiplier λ_l as a measure of the cost of the information rent to the insurer.

[16] If the average elasticity is below 1, it is optimal for the insurer to set the premium at $+\infty$. Discarding this possibility is without loss of generality.

Assumption

$$c_h\varepsilon_h(1-\varepsilon_l) - c_l\varepsilon_l(1-\varepsilon_h) > 0 \tag{19}$$

$$\phi\varepsilon_l + (1-\phi)\varepsilon_h > 1 \tag{20}$$

As the denominator of (17) is positive, $\lambda_l > 0$ if and only if inequality (19) holds. Equation (20) makes sure that σ in (15) has a finite value. The following lemma derives two conditions under which (19) is satisfied.

Lemma 4 *We find that $\lambda_l > 0$ if either*

- $\varepsilon_l \leq 1$
- *or*

$$\frac{c_h - c_l}{c_l} < \frac{\varepsilon_h - \varepsilon_l}{\varepsilon_h(\varepsilon_l - 1)} \tag{21}$$

If $\varepsilon_l \leq 1$ then (20) implies that $\varepsilon_h > 1$. Insurers then want to set such a high price on the l market that IC_l binds. If $\varepsilon_l > 1$, we have that $\varepsilon_h > \varepsilon_l > 1$. This implies that the fraction on the right hand side of (21) is bigger than 0. Hence, there is a range of values for $c_h > c_l$ for which (21) is satisfied. Put differently, for given $c_h - c_l > 0$, there exist ε_h big enough and $\varepsilon_l > 1$ small enough such that (21) holds.

Note that the inequality on $c_h - c_l$ is reminiscent of the condition in Lemma 2: the difference in costs between the types is bounded by the difference in elasticities. Intuitively, if the cost difference is too big (compared to the difference in elasticities), the h contract will be too expensive to be attractive for the l type. IC_l is not binding in this case. If the cost difference is small, the higher elasticity on the h market reduces the mark-up on the h contract compared to the l contract to such an extent that the h contract is attractive to the l type: IC_l is binding.

The following lemma derives properties of insurance markets that tend to lead to high supra profits.

Lemma 5 *The shadow price λ_l on the low type's IC constraint is:*

- *increasing in ε_h and decreasing in ε_l,*
- *decreasing in c_h and increasing in c_l.*

Recall that supra profits τ_i are increasing in λ_l. Hence, we find that supra profits are increasing in the difference between ε_h and ε_l. The more competitive the h market becomes, compared to the l market, the lower the mark-up on the h contract compared to the l contract. This makes the h market more attractive to the l type and hence the latter's IC constraint "more binding". Similarly, as the difference $c_h - c_l$ falls, the h market becomes more attractive for insurers to compete on which, in turn, makes the h contract more attractive to l types.

When explaining differences in pharmaceutical prices between countries, attention is focused on institutional features like income per head, who bargains with pharmaceutical companies (e.g. government vs. individual insurers), use of reference pricing and cost effectiveness analysis etc.; see Morton and Kyle (2011) and references therein. Lemma 5 suggests that selection incentives on the health insurance market can play a role as well. As a preliminary illustration of this point, selection problems in US health insurance tend to be seen as worse than in European countries with competitive health insurance markets (like the Czech Republic, Germany, Israel, The Netherlands). The latter countries have automatic/mandatory coverage levels above 90% of the population in contrast to the US where a large share of the population goes without health insurance. Also branded drug prices in the US are far higher than in Europe (Mulcahy et al., 2021). The analysis here suggests that these two phenomena could be connected.

Another factor determining treatment prices is the (individual rationality) constraint $u_k \geq 0$ where the outside option is normalized to 0. If drug prices increase to such an extent that $u_k < 0$, consumers stop buying health insurance because the expected benefit is lower than the price (premium). The introduction of generic drugs in the past decades which –due to competition– are priced below value ($p_j < v_j$) creates the space for patented drugs to price above value and still the value of insurance is bigger than or equal to σ. Mulcahy et al. (2021) show that generic drug prices are lower in the US than in Europe which allows branded drug producers to charge higher prices before the individual rationality constraint is binding.

The previous lemma compares the effects of different health insurance systems, e.g. differences between countries. Next, we compare –within a system– which treatments claim the highest supra profits.

Consider two treatments, 1 and 2; which treatment captures a higher supra profit compared to the value it offers? Equation (18) implies

$$\frac{\tau_1/v_1}{\tau_2/v_2} = \frac{\Delta\psi_1}{\Delta\psi_2}$$

From this it follows immediately that:

Corollary 1 *Relative to other patented treatments in P, the supra profit of treatment i increases in $\Delta\psi_i$.*

Hence drugs with a clear distinction between heavy users and a low probability for low risk people tend to earn high supra profits. One can think of two reasons why $\Delta\psi_i$ is high for a treatment i: the first is exogenous to the R&D lab and the second endogenous. First, it can be a matter of biology: some people suffer from diabetes and others do not.[17] The difference between the prevalence of diabetes among high and low risk types determines $\Delta\psi_i$ and firm i cannot change this.

Second, R&D lab i can invest in projects that are targeted at sub-populations of patients with a disease. Dugger et al. (2018) describe this as "a transition away from the production of 'one-size-fits-all' treatments towards targeted treatments". With precision or personalized medicine, the treatment takes the patient's underlying mechanism of the disease into account. Such targeted therapies require the co-development of diagnostic tools to identify the optimal treatment for individual patients. Biomarkers are used to define the subset of patients who benefit from the treatment. The use of biomarkers in clinical trials has increased substantially over time (Chandra et al., 2018).

Advantages of precision medicine include faster development and smaller/cheaper trials because the drugs are targeted at smaller groups (Chandra et al., 2018). Better clinical results for the sub-population of patients targeted by the treatment. Ideally, lower healthcare expenditure because of the cheaper development and the fact that the drugs are used for smaller groups. However, the last effect has not materialized as personalized medicine turns out to be very expensive.

Pharmaceutical companies need to select the most promising among the drug targets identified in early stages of research to pursue further (Emmerich et al., 2021; Knowles and Gromo, 2003). There are two margins along which they can decide to focus on targeted drugs. The extensive margin where they prioritize targeted projects above general projects. The intensive margin where they decide to narrow down a given project by investing in the discovery of (more) biomarkers. This decision is partly informed by science but there is an important role for marketing and financial directors (Knowles and Gromo, 2003).

[17]Note the following distinction: type 2 diabetes tends to be endogenous to an individual's lifestyle but exogenous to an R&D lab. Here we focus on the exogeneity with respect to the lab.

Although biomarkers "divide the market of treatable patients into groups and clusters, thus reducing market share" (Jakka and Rossbach, 2013), this targeting can be profitable in its own right beyond the (socially) beneficial effects mentioned above. In particular, we show that this partitioning of the market is profitable even if no extra value is created by targeting the treatment. In this sense, there is an excessive incentive for pharmaceutical companies to target treatments with precision medicine.

To capture this idea of excessive targeting, we introduce a parameter ζ with the properties that $d\psi_{ih}/d\zeta < 0$ and $d\Delta\psi_i/d\zeta > 0$. This we call "high type targeting". In words, innovator i focuses on a high type targeting strategy (lab i increases ζ_i) if its treatment will be effective for only a subset of high types ($d\psi_{ih}/d\zeta_i < 0$) but for an even smaller set of low types ($d\Delta\psi_i/d\zeta_i > 0$). To illustrate, consider a disease with different strains. Focusing treatment i on a particular strain that is more prevalent under high than low types, leads to a fall in i's market share under h types –as not all h types have this strain– and to an even bigger fall in market share under low types.

The corollary shows that a treatment which is more high type targeted leads to higher supra profits. This does not imply that targeting is necessarily a profitable strategy as it shrinks the market for the treatment. We come back to this below.

To the extent that the specialty pharmaceutical market and personalized medicine are examples of high type targeting, the corollary implies that they have contributed to the rise in treatment prices documented in the Introduction. This is also our explanation for the observation that "payers value rarity": prices tend to be high for (orphan) drugs with very small patient populations. If such drugs are hardly used by high types (low ψ_{ih}) and even fewer low types (high $\Delta\psi_i$) prices and supra profits will be high.

5.2 linear pricing and separation

Consider the case where the innovating firm cannot use a two-part tariff but only linear pricing. To capture the supra profits, firm i cannot use the fixed fee but has to set $p_i > v_i$. We know from lemma 1 that $p_i > v_i$ implies $x_{il} = 0$: the insurer drops treatment i in the l contract. Hence, firm i weighs a higher price p_i against a smaller group of customers using the treatment.

The first order conditions for u_h, u_l are unchanged, hence equations (12), (13) and (14) remain valid here. However, in this case, $x_{il} = 0$ while $x_{ih} = 1$; thus, there

is no pooling contract.

The following result characterizes the insurance premia charged on both market segments. The set E_h denotes the patented treatments that are exclusively available to h types: $x_{ih} = 1$ and $x_{il} = 0$ for all $i \in E_h$.

Lemma 6 *Let $\sigma = \sigma_l$. Then $\sigma_h = \sigma + \Delta\sigma$ where*

$$\Delta\sigma = \sum_{i \in E_h} \psi_{il} v_i \tag{22}$$

and σ is (implicitly) defined by

$$\sigma = \frac{\phi\varepsilon_l c_l + (1-\phi)\varepsilon_h c_h \frac{\sigma}{\sigma+\Delta\sigma}}{\phi\varepsilon_l + (1-\phi)\varepsilon_h - 1} \tag{23}$$

Interestingly, the additional premium $\Delta\sigma$ paid by the h type is not determined by the additional utility she gets from buying her own contract instead of the l contract. Because IC_l is binding, $\Delta\sigma$ is driven by the extra utility the l type would get if she bought the h contract which is the only contract covering treatments E_h. In particular, the coverage of E_h is the only difference in coverage between the two contracts. Hence, for l to be indifferent between the two contracts (i.e. binding IC_l constraint), the premium difference must equal the utility difference for l. If l would buy the h contract, she would gain additional treatment value equal to $\sum_{i \in E_h} \psi_{il} v_i$. Since ψ_{il} tends to be small, the effect of one treatment on the premium is small. But the set E_h can be quite big and people can over-estimate ψ_{il} as we discuss in the next section.

Note that with $\Delta\sigma = 0$, the expression in (23) is identical to (15): if the set E_h is empty, we are back to a pooling contract. In the pooling setting above, we know that $\mu_l > \mu_h$; the same is true here.

Lemma 7 *With $\Delta\sigma > 0$, we find that $\mu_l > \mu_h$.*

There are two ways to understand this result. First, with $\varepsilon_l < \varepsilon_h$, the insurer sets a higher margin on the l than on the h market. Because the IC_l constraint is binding, the margins are modified (from simple $\mu = 1/\varepsilon$), but not to the extent that the inequality is reversed. Second, because $\Delta\sigma$ is based on both v_i and ψ_{il}, while c_h is based on $p_i \geq v_i$ and $\psi_{ih} > \psi_{il}$, the increase in premium $\Delta\sigma$ is smaller than the increase in costs $\Delta c = c_h - c_l$ with $c_k = \sum_{i \in P} \psi_{ik} p_i$. Hence, the insurers make a lower margin on the h than on the l market.

Using the first order condition of profits (7) with respect to x_{ih}, we derive the expression for $p_i > v_i$:

$$\frac{1}{n}(1-\phi)\psi_{ih}(v_i - p_i) + \lambda_l v_i \Delta\psi_i = 0 \tag{24}$$

This equation needs to hold with equality. If it would be negative, p_i is so high that the insurer drops i also from the h contract ($x_{ih} = 0$). If it is positive, p_i can be increased further by firm i without losing sales. It follows that $\lambda_l > 0$ implies $p_i > v_i$. Hence, firm i faces a trade off: set $p_i = v_i$ and sell to both types or set $p_i > v_i$ and i's treatment is only covered by the h contract. The proposition presents the condition under which setting $p_i > v_i$ maximizes i's profits.

Proposition 1 *If* $\psi_{il} \leq \Delta\psi_i(1 - \varepsilon_l(1 - \frac{c_l}{\sigma}))$ *then*

$$\frac{p_i - v_i}{v_i} = \frac{\phi}{(1-\phi)\psi_{ih}}\Delta\psi_i(1 - \varepsilon_l(1 - \frac{c_l}{\sigma})) > 0 \tag{25}$$

where σ *is defined by equation* (23).

The intuition for the condition in the proposition is as follows. By charging $p_i > v_i$, i loses the l market with profit $\phi\psi_{il}v_i$. And it gains on the h market by charging $p_i > v_i$ instead of v_i; see equation (25):

$$(1-\phi)\psi_{ih}(p_i - v_i) = \phi\Delta\psi_i v_i(1 - \varepsilon_l(1 - \frac{c_l}{\sigma}))$$

Comparing the gain $\phi\psi_{il}v_i$ and loss $\phi\Delta\psi_i v_i(1 - \varepsilon_l(1 - c_l/\sigma))$ yields the condition in the proposition.

Intuitively, if the l market is big (high ψ_{il}), it is profitable to keep selling on this market segment as well: $p_i = v_i$. But if ψ_{il} is small compared to $\Delta\psi_i$, it is profitable to reap supra profits on the h segment.

We derive the following comparative static results with respect to the supra profits. When considering the effects of high type targeting, we make the following simplifying assumption. An insurance contract covers many treatments and a change in ψ_{ih}, ψ_{il} has a negligible effect on c_h, c_l and σ.[18]

Lemma 8 *The supra profit* $\frac{p_i - v_i}{v_i}$ *is*

- *increasing in* ε_h *and decreasing in* ε_l*;*

[18]Working with a continuum of treatments and writing $c_k = \int_P \psi_{ik}x_{ik}p_i di$ would make this argument mathematically consistent.

- *decreasing in c_h and increasing in c_l;*
- *increasing in high type targeting ζ_i.*

The results on the elasticities $\varepsilon_l, \varepsilon_h$ and cost levels c_l, c_h on supra profits are the same as in the pooling case.

As firm i targets more on high types, ψ_{il} falls and $\Delta\psi_i$ increases. Hence, it is easier to satisfy the condition in the proposition and –given that the condition is satisfied– the price increases as well with ζ_i because ψ_{ih} falls. As $\Delta\psi_i$ increases, treatment i becomes more important in separating the l from the h types. For this improved role in separating types, firm i can charge a higher price.

We finish the section with the question whether targeting on high types can be profitable for an R&D lab. Note that we use a rather restrictive definition of targeting: it reduces the probability that a treatment is effective for both h and l agents, without any benefits. This way we make the point that targeting can be excessive. Therefore, we derive conditions under which an increase in ζ_i (without any social benefits) is, in fact, profitable for the R&D lab. The profitability arises because of the supra profits earned by a "narrow" treatment.

Let π_i denote the profits earned on the insurance market by research lab i. Assuming that $p_i > v_i$, lab i's profits can be written as follows.[19]

$$\pi_i = (1-\phi)\psi_{ih}p_i = (1-\phi)\psi_{ih}v_i + \phi\Delta\psi_i v_i(1-\varepsilon_h(1-\frac{c_l}{\sigma}))$$

We can then derive the following result which does not hold in "normal" markets.

Corollary 2 *For each $d\psi_{ih}/d\zeta_i < 0$ and $d\Delta\psi_i/d\zeta_i > 0$, there exists $\phi \in [0,1]$ close enough to 1 such that $d\pi_i/d\zeta_i > 0$.*

Hence, when the share of high types in the population is small, R&D labs have an incentive to reduce their market by specializing in disease strains that are particularly prevalent among the high types. This only happens because the labs are selling to a downstream market plagued by adverse selection. In a "normal" market, reducing the appeal of your product to a subset of customers is not an optimal strategy (if this specialization has no other benefits). In fact, this strategy can be more profitable than described here if indication-based pricing is possible. This allows pharmaceutical companies to charge $p_i > v_i$ for the targeted group and $p_i \leq v_i$ for the low types (Chandra et al., 2018).

[19] Recall that we normalized lab i's marginal production costs to 0.

The argument is not that the specialty pharmaceutical market and personalized medicine are socially wasteful. By focusing the treatment on certain subgroups, side effects can be reduced and the development time of the treatment can be reduced. These effects are valuable from a social point of view. However, the analysis above does show two things. First, these developments to specialize medicines for subgroups contribute to rising drug prices even when the specialization does not increase production costs. Second, the private incentives for such specialization are excessive. Even if there are no social benefits from specialization, it is still profitable from a private point of view.

6 Bounded rationality

Above we assume that consumers are perfectly rational: they understand the value of the treatments covered by the insurance plan and they know their probability of needing a particular treatment. The main result is that part of the reason why drug prices are so high is that they are sold to a health insurance market with adverse selection problems. We formalize this by showing that even with rational consumers it can be optimal for insurers to over-pay for a treatment, that is pay more than the treatment's value for patients.

We know that consumers find healthcare markets difficult to (fully) understand. For instance, Handel and Kolstad (2015) show that people over-estimate the value of more generous insurance plans. In particular, low risk types tend to buy the more generous plan aimed at high risk consumers.

As an example of the effect of lack of rationality, we model this over-valuing by low types by assuming that they over-estimate the probability that they need a patented treatment: $\psi_{il}^e > \psi_{il}$ for $i \in P$ where ψ_{il}^e denotes the low type's expectation of ψ_{il}. We show that the effect of the over-estimation $\delta_{il} = \psi_{il}^e - \psi_{il} > 0$ on the price p_i for treatment i is non-monotone. First, it allows for $p_i > v_i$ even in a pooling contract. But for high δ_{il}, the price falls as treatment i is no longer effective in separating the high and low risk types.

It is routine to verify that profit function Π^ι in equation (7) can now be written

as:

$$\Pi^{\iota} = \phi q^{\iota}(u_l^{\iota}, u_l^{-\iota}, \theta_l)(\sum_{i \in P} \alpha_i x_{il}^{\iota}[\psi_{il}(v_i - p_i) + v_i \delta_{il}] + \sum_{j \in O} \psi_{jl} x_{jl}^{\iota} v_j - u_l^{\iota})$$
$$+ (1-\phi) q^{\iota}(u_h^{\iota}, u_h^{-\iota}, \theta_h)(\sum_{i \in P} \alpha_i \psi_{ih} x_{ih}^{\iota}(v_i - p_i) + \sum_{j \in O} \psi_{jh} x_{jh}^{\iota} v_j - u_h^{\iota})$$
$$+ \lambda_l(u_l^{\iota} - u_h^{\iota} + \sum_{i \in P} \alpha_i x_{ih}^{\iota} v_i(\Delta\psi_i - \delta_{il}) + \sum_{j \in O} x_{jh}^{\iota} v_j \Delta\psi_j)$$

where expressions with δ_{il} are added to the first and last line and we focus on the case with $\lambda_l > \lambda_h = 0$.

The first order conditions for most variables are the same as above, with two exceptions. First, the first order condition for x_{il}^{ι} shows that $x_{il} = 1$ if and only if

$$(v_i - p_i)\psi_{il} + v_i \delta_{il} \geq 0$$

or equivalently

$$p_i \leq v_i \left(1 + \frac{\delta_{il}}{\psi_{il}}\right) \tag{26}$$

Second, the first order condition for x_{ih}^{ι} implies that $x_{ih} = 1$ if and only if

$$p_i \leq v_i \left(1 + \lambda_l \frac{\Delta\psi_i - \delta_{il}}{(1-\phi)\frac{1}{n}\psi_{ih}}\right) \tag{27}$$

in symmetric equilibrium with $q^{\iota}(.) = 1/n$.

Hence, we can have a pooling contract with $x_{il} = x_{ih} = 1$ for $p_i > v_i$ as long as both (26) and (27) hold. Therefore, for $\delta_{il} > 0$ but small, we extend the range of parameters for which it is optimal for research lab i to induce a pooling contract instead of charging p_i so high that treatment i is dropped from l type's contract. As δ_i increases, so does p_i. However, for high δ_i, l types over-estimate the probability ψ_{il} to such an extent that treatment i is no longer useful in separating the types. If l types believe that their probability of needing i is the same as for h types, we have $\delta_{il} = \Delta\psi_i$ and equation (27) implies $p_i \leq v_i$.

Assuming that low risk types over-estimate (to some extent) the probability that they need treatment, the patent holder on this treatment can charge a (linear) price that exceeds the value of its treatment and the treatment is covered by health insurance for both risk types. In this sense, taking bounded rationality into account strengthens our result that a pharmaceutical company can profitably charge a price above the value of its treatment.

7 Policy implications

This paper introduces a framework where upstream firms sell inputs to a downstream market where the downstream market suffers from adverse selection problems. We find that if the low type's incentive compatibility constraint is binding in the downstream market, we find a number of results that are counter-intuitive at first sight. First, upstream firms can charge prices in excess of consumers' valuation of their product. Second, upstream firms have an incentive to narrow their market; that is, make their product attractive to a subset of final consumers. Third, making downstream market segments more competitive, tends to increase upstream prices.

An indication that the low type's incentive constraint is binding is that low types buy the (generous) contracts that are aimed at high types. Handel and Kolstad (2015) show that this happens in health insurance markets.

We have applied our framework to the pharmaceutical market which has been in the news in past years for charging exorbitant prices for drugs. Indeed, the model explains that drug prices tend to be high because pharmaceutical companies sell to a health insurance market with adverse selection problems. Theoretically, these prices can be in excess of treatment value. A couple of developments have contributed to the price increases in the pharmaceutical markets. First, the increased adoption of generic drugs has created the "space" for patented drugs to charge prices in excess of treatment value. Although the net value of coverage for some treatments is negative from the insured's point of view, the overall value of insurance is still positive. Second, the development to target treatments to subgroups of patients suffering from a disease also leads to upward pressure on drug prices. Moreover, we have shown that the private incentives for targeting treatments –say, through personalized medicine– are excessive.

We assume that pharmaceutical companies make take-it-or-leave-it offers to insurers. We show that these offers can lead to prices above value. If, instead, pharmaceutical companies and insurers bargain over prices and insurers have some market power, prices tend to be lower. The outcome can then still be a high price close to value because without insurer bargaining power prices exceed treatment value.

The implications of our analysis for policy can be summarized as follows. First, there have been numerous recent examples of drugs being sold at very high prices. The narrative usually is that it is "unfair" or not "ethical" for pharmaceutical companies to benefit from people's health problems. We show that it is not only unfair, it may well be inefficient. By charging a price in excess of a treatment's value,

R&D incentives are distorted: (i) incentives to do R&D can be excessive as firms earn supra profits: the private value of the innovation exceeds the social value; (ii) firms have an excessive incentive to target their treatment to subgroups of patients: even if there is no social value to targeting, it is still privately profitable. To reduce the excessive R&D incentives, a government can reduce tax breaks for R&D in the pharmaceutical sector and increase the industry's financial contribution to research by (public) universities both for fundamental research and for running trials to test new treatments. Further, the government can consider introducing price caps; for instance, in the form of not approving treatments for insurance coverage if the price per qaly (quality adjusted life year) is too high. This helps to keep the health-care system affordable and reduces excessive R&D incentives. As shown, relying on market forces to keep prices low does not work for an upstream sector selling to a downstream market with adverse selection problems.

Second, assuming that consumers stop buying insurance in case the expected value of the insurance plan is lower than the premium, treatment prices can be reduced by creating a separate insurance market for patented treatments. The separation would be similar to having basic and supplementary insurance markets as some countries have; but here there would be insurance for patented drugs and separate insurance for treatments where the patent has run out (which we called "open" above). "Open insurance" would cover all generic drugs which yield high patient utility compared to their cost. This leaves less rents for patented treatments on their insurance market segment to appropriate by charging a price above treatment value. Such a segmentation of the health insurance market helps to reduce treatment prices.

References

Adam Atherly, Bryan E. Dowd, and Roger Feldman. The effect of benefits, premiums, and health risk on health plan choice in the medicare program. *Health Services Research*, 39(4p1):847–864, 2004.

David Auerbach and Sabina Ohri. Price and the demand for nongroup health insurance. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 43(2):122–134, 2006.

Nancy Dean Beaulieu. Quality information and consumer health plan choices. *Journal of Health Economics*, 21(1):43 – 63, 2002. Health Plan Choice.

Jan Boone and Christoph Schottmüller. Health insurance without single crossing: why healthy people have high coverage. *Economic Journal*, 127(599):84–105, 2017.

Franco Cavalli. An appeal to world leaders: Stop cancer now. *The Lancet*, 381 (9865):425–426, 2013.

Amitabh Chandra, Craig Garthwaite, and Ariel Dora Stern. *Characterizing the Drug Development Pipeline for Precision Medicines*, pages 115–157. University of Chicago Press, May 2018.

J. Cohn. *Sick: the untold story of America's health care crisis–and the people who pay the price*. Harper Perennial, 2007.

Joan Costa and Jaume Garcia. Demand for private health insurance: how important is the quality gap? *Health Economics*, 12(7):587–599, 2003. ISSN 1099-1050. doi: 10.1002/hec.756.

Patricia M. Danzon and Erin Taylor. Drug pricing and value in oncology. *The Oncologist*, 15(S1):24–31, 2010.

Brett Doble. Budget impact and cost-effectiveness: can we afford precision medicine in oncology? *Scandinavian Journal of Clinical and Laboratory Investigation*, 76 (sup245):S6–S11, 2016.

S.A. Dugger, A. Platt, and D.B. Goldstein. Drug development in the era of precision medicine. *Nature Reviews Drug Discovery*, 17:183–196, March 2018. doi: 10.1038/nrd.2017.226.

C.H. Emmerich, L.M. Gamboa, M.C.J. Hofmann, M. Bonin-Andresen, O. Arbach, P. Schendel, B. Gerlach, K. Hempel, A. Bespalov, U. Dirnagl, and M.J. Parnham. Improving target assessment in biomedical research: the got-it recommendations. *Nature Reviews Drug Discovery*, 20:64–81, January 2021. doi: 10.1038/s41573-020-0087-3.

A. Finkelstein and K. McGarry. Multiple dimensions of private information: evidence from the long-term care insurance market. *The American Economic Review*, 96(4):938–958, 2006.

P. Frijters, J.P. Haisken-DeNew, and M.A. Shields. The causal effect of income on health: Evidence from german reunification. *Journal of Health Economics*, 24(5): 997–1017, 2005.

Livio Garattini, Alessandro Curto, and Nick Freemantle. Personalized medicine and economic evaluation in oncology: all theory and no practice? *Expert Review of Pharmacoeconomics & Outcomes Research*, 15(5):733–738, 2015.

Jr. Garrison, L.P. and A. Towse. Value-based pricing and reimbursement in personalised healthcare: Introduction to the basic health economics. *Journal of Personalized Medicine*, 7(3), 2017.

Martin Gaynor, Kate Ho, and Robert J. Town. The industrial organization of health-care markets. *Journal of Economic Literature*, 53(2):235–84, June 2015.

H. Gravelle and M. Sutton. Income, relative income, and self-reported health in britain 1979-2000. *Health Economics*, 18(2):125–145, 2009.

Benjamin R. Handel and Jonathan T. Kolstad. Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105(8):2449–2500, 2015.

Nigel Hawkes. Uk to spend £250m on proton beam treatment despite no appraisal by nice. *BMJ*, 344, 2012. doi: 10.1136/bmj.e2627.

Kate Ho, Joseph Hogan, and Fiona Scott Morton. The impact of consumer inattention on insurer pricing in the medicare part d program. *The RAND Journal of Economics*, 48(4):877–905, 2017.

David H. Howard, Peter B. Bach, Ernst R. Berndt, and Rena M. Conti. Pricing in the market for anticancer drugs. *Journal of Economic Perspectives*, 29(1):139–62, February 2015. doi: 10.1257/jep.29.1.139.

Sairamesh Jakka and Michael Rossbach. An economic perspective on personalized medicine. *The HUGO Journal*, 7(1):1, 2013.

Philippe Jehiel, Benny Moldovanu, and Ennio Stacchetti. How (not) to sell nuclear weapons. *The American Economic Review*, 86(4):814–829, 1996.

Jurjen Kamphorst and Vladimir A. Karamychev. Going through the roof: On prices for drugs sold through insurance, 2021.

Michael L. Katz and Carl Shapiro. How to license intangible property. *Quarterly Journal of Economics*, 101(3):567 – 589, 1986.

J. Knowles and G. Gromo. Target selection in drug discovery. *Nature Reviews Drug Discovery*, 2:63–69, January 2003.

J.J. Laffont and D. Martimort. *The theory of incentives: the principal-agent model*. Princeton University Press, 2002.

Robert Langreth. Prostate cancer therapy too good to be true explodes health cost. *Bloomberg Technology*, March 26 2012.

Tracy R Lewis and David E.M Sappington. Countervailing incentives in agency problems. *Journal of Economic Theory*, 49(2):294 – 313, 1989.

Christopher McCabe, Karl Claxton, and Aki Tsuchiya. Orphan drugs and the nhs: should we value rarity? *BMJ*, 331:1016–1019, October 2005.

G. Medic, D. Korchagina, K.E. Young, M. Toumi, M.J. Postma, M. Wille, and M. Hemels. Do payers value rarity? an analysis of the relationship between disease rarity and orphan drug prices in europe. *Journal of Market Access and Health Policy*, 5, 2017.

Andrea Messori, Americo Cicchetti, and Luigi Patregani. Relating price determination to disease prevalence. *BMJ*, 341, 2010. ISSN 0959-8138. doi: 10.1136/bmj.c4615.

Fiona Scott Morton and Margaret Kyle. Chapter twelve - markets for pharmaceutical products. In Thomas G. Mcguire Mark V. Pauly and Pedro P. Barros, editors, *Handbook of Health Economics*, volume 2, pages 763 – 823. Elsevier, 2011.

Andrew W. Mulcahy, Christopher M. Whaley, Mahlet Gizaw, Daniel Schwam, Nathaniel Edenfield, and Alejandro U. Becerra-Ornelas. *International Prescription Drug Price Comparisons: Current Empirical Estimates and Comparisons with Previous Studies*. RAND Corporation, Santa Monica, CA, 2021. doi: 10.7249/RR2956.

M.K. Munkin and P.K. Trivedi. Disentangling incentives effects of insurance coverage from adverse selection in the case of drug expenditure: a finite mixture approach. *Health Economics*, 19(9):1093–1108, 2010. ISSN 1099-1050.

John A. Nyman. The value of health insurance: the access motive. *Journal of Health Economics*, 18(2):141 – 152, 1999.

Stephen T. Parente, Roger Feldman, and Jon B. Christianson. Employee choice of consumer-driven health insurance in a multiplan, multiproduct setting. *Health Services Research*, 39(4p2):1091–1112, 2004.

M. Rothschild and J. Stiglitz. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4):629–649, 1976.

Anne Beeson Royalty and Neil Solomon. Health plan choice: Price elasticities in a managed competition setting. *The Journal of Human Resources*, 34(1):1–41, 1999. ISSN 0022166X.

Evan Saltzman. Demand for health insurance: Evidence from the california and washington aca exchanges. *Journal of Health Economics*, 63:197 – 222, 2019. doi: https://doi.org/10.1016/j.jhealeco.2018.11.004.

C Schoen, R Osborn, D Squires, M M Doty, R Pierson, and S Applebaum. How health insurance design affects access to care and costs, by income, in eleven countries. *Health Affairs*, 29(12):1–12, 2010.

Cathy Schoen, Sara R Collins, Jennifer L Kriss, and Michelle M Doty. How many are underinsured? trends among u.s. adults, 2003 and 2007. *Health affairs (Project Hope)*, 27(4):298–309, 2008.

Erik Schokkaert and Carine van de Voorde. Chapter 15 - user charges. In S. Glied and P. Smith, editors, *Oxford Handbook of Health Economics*, pages 329 – 353. Oxford University Press, 2011.

Bruce A Strombom, Thomas C Buchmueller, and Paul J Feldstein. Switching costs, price sensitivity and health plan choice. *Journal of Health Economics*, 21(1):89 – 116, 2002.

Jean Tirole. *The theory of industrial organization*. MIT Press, 1988.

Katherine Eve Young, Imen Soussi, Michiel Hemels, and Mondher Toumi. A comparative study of orphan drug prices in europe. *Journal of Market Access and Health Policy*, 5, 2017.

A Proof of results

Proof of Lemma 1 Because

$$\left.\frac{d\Pi}{dx_{il}}\right|_{p_i=v_i} = -\lambda_h \alpha_i v_i(\psi_{ih} - \psi_{il}) < 0 \tag{28}$$

for $\lambda_h > 0$, binding IC_h implies that $x_{il} = 0$ for all $i \in P$. Further, note that with $p_i \leq v_i$ it is the case that $d\Pi/dx_{ih}, d\Pi/dx_{jh} \geq 0$; hence we can set $x_{ih} = x_{jh} = 1$ for all $i \in P, j \in O$.

Suppose, by contradiction, that both constraints are binding; then we have

$$\sum_{i\in P} \alpha_i v_i(\psi_{ih} - \psi_{il}) + \sum_{j\in O} v_j(\psi_{jh} - \psi_{jl}) = u_h - u_l = \sum_{j\in O} x_{jl} v_j(\psi_{jh} - \psi_{jl})$$

with $x_{jl} \leq 1$. Since the left hand side is strictly bigger than the right hand side ($\alpha_i > 0$), this is a contradiction.

If $\lambda_l > 0$, from the previous point we know that $\lambda_h = 0$. Hence, equation (28) implies that any $x_{il} \in [0, 1]$ is optimal. It is routine to verify that the derivative of Π with respect to each of the variables x_{ih}, x_{jh}, x_{jl} is positive; hence the insurer sets them equal to 1.

Finally, if $p_i > v_i$ then $d\Pi/dx_{il} < 0$ (even if $\lambda_h = 0$) and hence the insurer sets $x_{il} = 0$. *Q.E.D.*

Proof of Lemma 2 First observe that the derivative of equation (10) with respect to x_{1h}, x_{2h}, x_{2l} is strictly positive (with $\lambda_l > 0$) and hence each of these variables equals 1, the derivative with respect to x_{1l} equals 0 and this variable can take on any value between 0 and 1. The first order conditions with respect to u_l^a and u_h^a can be written as

$$\frac{\phi}{2t_l}(\psi_2 v_2 - u_l^a) - \phi(\frac{1}{2} + \frac{u_l^a - u_l^b}{2t_l}) + \lambda_l = 0 \tag{29}$$

$$\frac{1-\phi}{2t_h}(\psi_2 v_2 - u_h^a) - (1-\phi)(\frac{1}{2} + \frac{u_h^a - u_h^b}{2t_h}) - \lambda_l = 0 \tag{30}$$

Adding both equations, looking at a symmetric equilibrium with $u_k^a = u_k^b$ on both market segments $k = l, h$ and using that IC_l holds with equality

$$u_l^a = u_h^a - \alpha v_1 x_{1h}^a(\psi_{1h} - \psi_{1l})$$

leads to the expressions for u_h and u_l in the lemma.

The expression for λ_l follows from equation (30) with the values for u_l, u_h substituted in.

It is straightforward to see that $d\Pi/d\alpha = \lambda_l v_1(\psi_{1h} - \psi_{1l})$ which leads to (11).

Finally, we need to check that IC_h is indeed satisfied. Using the values for u_l, u_h, we find that

$$\frac{\frac{\phi}{t_l}}{\frac{\phi}{t_l}+\frac{1-\phi}{t_h}}\alpha v_1(\psi_{1h}-\psi_{1l}) \geq -\frac{\frac{1-\phi}{t_h}}{\frac{\phi}{t_l}+\frac{1-\phi}{t_h}}\alpha v_1(\psi_{1h}-\psi_{1l}) + \alpha v_1 x_{1l}(\psi_{1h}-\psi_{1l})$$

which holds because $x_{1l} \leq 1$.

In fact, it is also straightforward to derive the asymmetric case where $\alpha^b = 1$ while $\alpha^a = \alpha < 1$. In words, treatment 1 is always sold to insurer b but firm 1 threatens not to sell to insurer a. How much is a willing to pay to be able to cover treatment 1?

To derive this asymmetric equilibrium, we write the constraint as

$$u_l^a = u_h^a - \zeta_\alpha \tag{31}$$

where, to simplify notation, we use $\zeta_\alpha = \alpha(\psi_{1h} - \psi_{1l})v_1$. Adding the equations (29) and (30) and substituting (31) we get

$$\begin{aligned}&\frac{\phi}{2t_l}(\psi_2 v_2 + \zeta_\alpha - u_h^a) - \phi(\frac{1}{2} + \frac{u_h^a - \zeta_\alpha - u_h^b + \zeta_1}{2t_l}) \\ &+ \frac{1-\phi}{2t_h}(\psi_2 v_2 - u_h^a) - (1-\phi)(\frac{1}{2} + \frac{u_h^a - u_h^b}{2t_h}) = 0\end{aligned}$$

and a similar first order condition for u_h^b. We can write these in matrix notation as

$$(\frac{\phi}{t_l} + \frac{1-\phi}{t_h}) \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} u_h^a \\ u_h^b \end{bmatrix} = \begin{bmatrix} \frac{\phi}{t_l}(\psi_2 v_2 + 2\zeta_\alpha - \zeta_1) + \frac{1-\phi}{t_h}\psi_2 v_2 - 1 \\ \frac{\phi}{t_l}(\psi_2 v_2 + 2\zeta_1 - \zeta_\alpha) + \frac{1-\phi}{t_h}\psi_2 v_2 - 1 \end{bmatrix}$$

Inverting the matrix on the left hand side and solving for u_h^a, u_h^b yields

$$\begin{aligned}u_h^a &= \psi_2 v_2 + \frac{\frac{\phi}{t_l}\zeta_\alpha - 1}{\frac{\phi}{t_l} + \frac{1-\phi}{t_h}} \\ u_h^b &= \psi_2 v_2 + \frac{\frac{\phi}{t_l}\zeta_1 - 1}{\frac{\phi}{t_l} + \frac{1-\phi}{t_h}}\end{aligned}$$

Hence u_h^a does not depend on the level of $\alpha^b = 1$ and u_h^b does not depend on the level of $\alpha^a = \alpha$. Via equation (31) the same is true for u_l^a, u_l^b resp.

We can also derive λ_l^a, λ_l^b.

$$\begin{aligned}\lambda_l^a &= \frac{1}{2}\frac{t_l - t_h + (\zeta_\alpha - 2\zeta_1)}{\frac{t_l}{\phi} + \frac{t_h}{1-\phi}} \\ \lambda_l^b &= \frac{1}{2}\frac{t_l - t_h + (\zeta_1 - 2\zeta_\alpha)}{\frac{t_l}{\phi} + \frac{t_h}{1-\phi}}\end{aligned}$$

Q.E.D.

One of the questions in this paper is: how does this extra profit $d\Pi/d\alpha|_{p_1=v_1}$ vary with competition? In our model, competition is inversely related to travel cost t. Hence, in the notation of the previous section, we would define our variable capturing competition as $\theta_k = 1/t_k$. As (11) illustrates, $d\Pi/d\alpha$ increases in λ_l.

Lemma In the Hotelling model, the multiplier λ_l is increasing in t_l and decreasing in t_h.

Proof of Lemma Taking the derivatives of λ_l with respect to t_h and t_l, it is routine to verify that $d\lambda_l/dt_h < 0$ and $d\lambda_l/dt_l > 0$. *Q.E.D.*

Proof of Lemma 3 We write equation (14) as

$$\frac{\sigma - c_l}{\sigma}\phi\varepsilon_l + \frac{\sigma - c_h}{\sigma}(1-\phi)\varepsilon_h = 1$$

and solve for σ. The expression for μ_k follows from $\mu_k = (\sigma - c_k)/\sigma$.

From equations (12) and (16) we find the expression for λ_l with $q_k^\iota = 1/n$ in symmetric equilibrium. *Q.E.D.*

Proof of Lemma 4 Equation (19) clearly holds if $\varepsilon_l < 1$ since by assumption $\varepsilon_h > \varepsilon_l$. If $\varepsilon_h > \varepsilon_l > 1$, then the inequality can be written as (21). *Q.E.D.*

Proof of Lemma 5 Differentiating equation (17) with respect to ε_h, we find

$$\begin{aligned}\frac{d\lambda_l}{d\varepsilon_h} &\propto (c_h(1-\varepsilon_l) + c_l\varepsilon_l)(\phi\varepsilon_l c_l + (1-\phi)\varepsilon_h c_h) \\ &\quad - (1-\phi)c_h(c_h\varepsilon_h(1-\varepsilon_l) - c_l\varepsilon_l(1-\varepsilon_h)) \\ &= c_l\varepsilon_l(c_h - \phi\varepsilon_l(c_h - c_l)) > 0\end{aligned}$$

where the inequality follows from the following considerations. First, this derivative is positive if and only if $\mu_h > 0$ in equation (16). Since we are focusing here on the case where $\lambda_l > 0$, equation (12) implies that $\mu_h\varepsilon_h - 1 > 0$. Consequently, $\mu_h > 0$.

Next, consider the case where both elasticities are multiplied by a positive constant $\zeta > 0$: $\zeta\varepsilon_h, \zeta\varepsilon_l$. Substituting this into equation (17), it is routine to verify that

$$\frac{d\lambda_l}{d\zeta} \propto \frac{d(c_h\varepsilon_h(1-\zeta\varepsilon_l) - c_l\varepsilon_l(1-\zeta\varepsilon_h))}{d\zeta}$$

Hence we see that

$$\frac{d\lambda_l}{d\zeta} \propto -\varepsilon_h\varepsilon_l(c_h - c_l) < 0$$

From this it follows that λ_l is decreasing in ε_l, because the term is increasing in ε_h and decreasing in ζ.

Differentiating λ_l with respect to c_h gives

$$\frac{d\lambda_l}{dc_h} \propto \varepsilon_h(1-\varepsilon_l)(\phi c_l\varepsilon_l + (1-\phi)c_h\varepsilon_h) - (1-\phi)\varepsilon_h(\varepsilon_h c_h(1-\varepsilon_l) - \varepsilon_l c_l(1-\varepsilon_h))$$
$$= \varepsilon_h\varepsilon_l c_l(\phi(1-\varepsilon_l) + (1-\phi)(1-\varepsilon_h)) < 0$$

because of equation (20).

Similarly, we find for c_l that

$$\frac{d\lambda_l}{dc_l} \propto -\varepsilon_l\varepsilon_h c_h(\phi(1-\varepsilon_l) + (1-\phi)(1-\varepsilon_h)) > 0$$

Q.E.D.

Proof of lemma 6 Combining a binding IC_l constraint with the expressions for $u_{h,l}$ in equation (4), we find that

$$\sigma_h = \sigma_l + \sum_{i\in P}\psi_{il}v_i(x_{ih} - x_{il}) + \sum_{j\in O}\psi_{jl}v_j(x_{jh} - x_{jl})$$

Using the results in lemma 1 we can write this equation as (22).

Using $\sigma_l = \sigma$ and $\sigma_h = \sigma + \Delta\sigma$ in equation (14), we can write this as

$$\phi\left((1-\frac{c_l}{\sigma})\varepsilon_l - 1\right) + (1-\phi)\left((1-\frac{c_h}{\sigma}\frac{\sigma}{\sigma+\Delta\sigma})\varepsilon_h - 1\right) = 0 \tag{32}$$

From this it follows that (15) needs to be adjusted to (23) in case $\Delta\sigma \neq 0$. *Q.E.D.*

Proof of lemma 7 $\mu_l = (\sigma - c_l)/\sigma$ while

$$\mu_h = \frac{\sigma + \Delta\sigma - c_h}{\sigma + \Delta\sigma} = \frac{\sigma - (c_h - \sum_{i\in E_h}\psi_{il}v_i)}{\sigma + \Delta\sigma}$$

Hence, there are two reasons why $\mu_l > \mu_h$. First, consider the denominator: $\sigma + \Delta\sigma > \sigma$. Second, in the numerator we have

$$c_h - \sum_{i\in E_h}\psi_{il}v_i = \sum_{i\in P\setminus E_h}\psi_{ih}v_i + \sum_{i\in E_h}(\psi_{ih}p_i - \psi_{il}v_i) > \sum_{i\in P\setminus E_h}\psi_{il}v_i = c_l$$

because $\psi_{ih}p_i > \psi_{il}v_i$ as $p_i \geq v_i$ and $\psi_{ih} > \psi_{il}$. *Q.E.D.*

Proof of proposition 1 Note that the right hand side of the condition in the proposition is positive. We know from equation (13) that $\mu_l\varepsilon_l < 1$ because we consider the case where $\lambda_l > 0$ (see below). This we can write as $1 - \varepsilon_l(\sigma - c_l)/\sigma > 0$.

Further note that $\sigma - c_l > 0$. We prove this by contradiction. Assume that $c_l \geq \sigma$. That is, the l contract is loss making. Consider the h contract:

$$\sigma + \Delta\sigma - c_h \leq c_l + \sum_{i\in E_h}\psi_{il}v_i - c_h = \sum_{i\in E_h}\psi_{il}v_i - \Delta c \leq \sum_{i\in E_h}\psi_{il}v_i - \sum_{i\in E_h}\psi_{ih}p_i < 0$$

where the first inequality follows from the assumption (that we want to contradict) that $c_l \geq \sigma$, the second from the fact that there can be other differences in costs between h and l in addition to the h coverage of treatments in E_h (e.g. treatments priced at $p_k = v_k$ and $\psi_{kh} > \psi_{kl}$) and the final inequality follows from $p_i > v_i$ and $\psi_{ih} > \psi_{il}$ for $i \in E_h$.

From equation (24), we can write the optimal $p_i > v_i$ as

$$p_i = v_i \left(1 + \lambda_l \frac{\Delta\psi_i}{(1-\phi)\psi_{ih}/n}\right)$$

Firm i can set $p_i = v_i$ and sell to both types or set $p_i > v_i$ and only sell to type h. The latter is more profitable if

$$(\phi\psi_{il} + (1-\phi)\psi_{ih})v_i \leq (1-\phi)\psi_{ih}v_i \left(1 + \lambda_l \frac{\Delta\psi_i}{(1-\phi)\psi_{ih}/n}\right)$$

Using the expression for λ_l in equation (13) written as $\lambda_l = \phi/n(1 - \varepsilon_l(1 - c_l/\sigma))$, this expression can be written as the inequality in the proposition. If the inequality holds, equation (24) can be written as (25).

Finally, note that our assumption in the main text is also a sufficient condition for $\lambda_l > 0$ in this separating equilibrium. This can be seen as follows. Equation (13) implies that

$$\lambda_l > 0 \text{ if and only if } 1 - \varepsilon_l \left(1 - \frac{c_l}{\sigma}\right) > 0$$

Hence, if

$$1 - \varepsilon_l \left(1 - \frac{c_l}{\sigma}\right) > 0 \tag{33}$$

holds for an upper-bound on σ_l, we find that $\lambda_l > 0$ for each σ below this upper-bound. Comparing the expressions in (15) and (23), we see that $\Delta\sigma > 0$ implies $\sigma/(\sigma + \Delta\sigma) < 1$ and σ in (23) is lower than σ in (15) – see the proof of lemma 8 on how to solve the implicit equation (23). Hence, (15) provides the upper-bound for σ in (23). Inequality (19) ensures that (33) holds for the upper-bound of σ given by (15). Hence, (33) also holds for the solution to (23). *Q.E.D.*

Proof of lemma 8 Deriving comparative statics from implicit equation (23), we view this equation as looking for a fixed point of the expression on its right hand side (which we denote rhs).

We will show that $\partial rhs/\partial\sigma \in \langle 0, 1\rangle$ and hence a change in a variable that shifts rhs up (down) will increase (decrease) σ in (23); see Figure 2. Hence, we want to show that

$$\frac{\partial rhs}{\partial\sigma} = \frac{\Delta\sigma}{(\sigma + \Delta\sigma)^2} \frac{(1-\phi)\varepsilon_h c_h}{\phi\varepsilon_l + (1-\phi)\varepsilon_h - 1} < 1$$

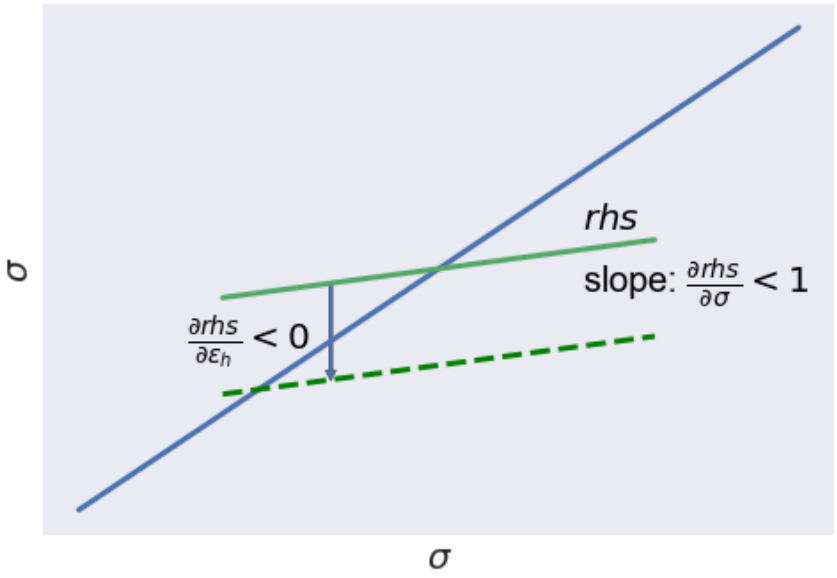

Figure 2: Solving equation (23) for σ.

while it is obvious that $\partial rhs/\partial\sigma > 0$. Equivalently

$$\frac{\Delta\sigma}{\sigma+\Delta\sigma}\frac{c_h}{\Delta\sigma+\sigma}(1-\phi)\varepsilon_h < \phi\varepsilon_l + (1-\phi)\varepsilon_h - 1$$

Combining this with equation (32) we find

$$\frac{\Delta\sigma}{\sigma+\Delta\sigma}\frac{c_h}{\sigma+\Delta\sigma}(1-\phi)\varepsilon_h < \phi\frac{c_l}{\sigma} + (1-\phi)\varepsilon_h\frac{c_h}{\sigma+\Delta\sigma}$$

which clearly holds because $\Delta\sigma/(\Delta\sigma+\sigma) \leq 1$. Hence, we have found that $\partial rhs/\partial\sigma < 1$.

To derive the effect of ε_h on σ, we need to find $\partial rhs/\partial\varepsilon_h$. It is routine to verify that

$$\frac{\partial rhs}{\partial\varepsilon_h} = (1-\phi)\frac{\phi\varepsilon_l(c_h\frac{\sigma}{\sigma+\Delta\sigma} - c_l) - c_h\frac{\sigma}{\sigma+\Delta\sigma}}{(\phi\varepsilon_l + (1-\phi)\varepsilon_h - 1)^2} < 0 \tag{34}$$

To prove this inequality, first consider the case of $\Delta\sigma = 0$, i.e. we are back in the pooling contract. We know from equation (12) and $\lambda_l > 0$ that $\mu_h > 0$. Equation (16) then implies the inequality in (34). Next consider $\Delta\sigma > 0$. For $\Delta\sigma$ big enough, the expression in (34) is also negative. Finally, note that the sign of the derivative of the right hand side of (34) with respect to $\Delta\sigma$ does not vary with $\Delta\sigma$ nor with σ. That is, it cannot feature a maximum where the numerator of (34) is positive. Hence, irrespective of the sign of this derivative, we see that (34) is negative for all $\Delta\sigma > 0$.

As ε_h increases, rhs shifts downwards in Figure 2 and σ falls. Hence, we find that

$$\frac{d\left(\frac{p_i-v_i}{v_i}\right)}{d\varepsilon_h} = -\frac{\phi}{1-\phi}\frac{\Delta\psi_i}{\psi_{ih}}\frac{\varepsilon_l c_l}{\sigma^2}\frac{d\sigma}{d\varepsilon_h} > 0$$

To find the effect of ε_l, we use equation (12) for λ_l and hence we write

$$\frac{p_i - v_i}{v_i} = \frac{\Delta\psi_i}{\psi_{ih}}\left(\varepsilon_h\left(1 - \frac{c_h}{\sigma + \Delta\sigma}\right) - 1\right)$$

To find $d\sigma/d\varepsilon_l$, we consider $\partial rhs/\partial\varepsilon_l$. It is routine to verify that the sign of this derivative equals the sign of the following expression:

$$(1-\phi)\varepsilon_h\left(c_l - c_h\frac{\sigma}{\sigma+\Delta\sigma}\right) - c_l \tag{35}$$

This expression is negative, as can be seen as follows. We know from lemma 7 that $\mu_l > \mu_h$ which can be written as

$$1 - \frac{c_l}{\sigma} > 1 - \frac{c_h}{\sigma + \Delta\sigma}$$

or equivalently

$$\frac{c_l}{\sigma} < \frac{c_h}{\sigma + \Delta\sigma}$$

from which it follows that (35) is negative. Thus we find that $d\sigma/d\varepsilon_l < 0$.

Next, consider c_h. We find that

$$\frac{d\left(\frac{p_i-v_i}{v_i}\right)}{dc_h} = -\frac{\phi}{1-\phi}\frac{\Delta\psi_i}{\psi_{ih}}\frac{\varepsilon_l c_l}{\sigma^2}\frac{d\sigma}{dc_h}$$

Since it is obvious that $\partial rhs/\partial c_h > 0$, we find that $d\sigma/dc_h > 0$ and hence supra profits fall with c_h.

For c_l, we define $\tilde{\sigma} = \sigma/c_l$ and we write equation (23) as

$$\tilde{\sigma} = \frac{\phi\varepsilon_l + (1-\phi)\varepsilon_h c_h\frac{\tilde{\sigma}}{\tilde{\sigma}c_l+\Delta\sigma}}{\phi\varepsilon_l + (1-\phi)\varepsilon_h - 1}$$

It is clear that $\partial rhs/\partial c_l < 0$ and hence $\partial\tilde{\sigma}/\partial c_l < 0$. This implies that $\partial(c_l/\sigma)\partial c_l > 0$ and therefore:

$$\frac{d\left(\frac{p_i-v_i}{v_i}\right)}{dc_l} > 0$$

The effect of ζ_i is obvious: since we assume that there are many treatments covered by insurance so that the effect of ζ_i on σ is negligible, we have $d(\Delta\psi_i/\psi_{ih})/d\zeta_i > 0$. *Q.E.D.*

Proof of Corollary 2 We find the following derivative

$$\frac{d\pi_i}{d\zeta_i} = (1-\phi)\frac{d\psi_{ih}}{d\zeta_i}v_i + \phi\frac{d\Delta\psi_i}{d\zeta_i}v_i(1-\varepsilon_h(1-\frac{c_l}{\sigma})) > 0$$

for ϕ close enough to 1 because $d\psi_{ih}/d\zeta_i < 0$ and $d\Delta\psi_i/d\zeta_i > 0$. *Q.E.D.*