# Health Insurance using a Cost-Effectiveness Threshold in the Presence of Market Power

Javad Moradpour and Aidan Hollis

Department of Economics, University of Calgary

November 8, 2019

**Abstract**

Public health insurers typically use a cost-effectiveness threshold to determine which health products and services should be insured. We challenge the convention of a single threshold. For competitively provided products and services, prices are determined by cost; but for products with market power, patentees will increase the price to the threshold. As a result, a change in the threshold for patented products affects the prices of all, including inframarginal ones. The insurer can increase efficiency by reducing the threshold for patented products, even accounting for the effect on innovation. In a multi-country world, thresholds may however be too low.

keywords: cost-effectiveness threshold; market power; public insurance.

## 1 Introduction

National health insurance is one of the most important functions of many governments, and health expenditures have grown to absorb an increasing share of national budgets. Thus the question of which medical services and products should be insured has grown to be an increasingly pressing problem. For many innovative products, getting included in the national formulary (or list of products to be insured) is a key to profitability. The market for innovative medicines, however, is an unusual one, with state insurers dominating demand, and setting prices through specialized measures of value.

The mechanism that insurers have increasingly come to favour is the use of "cost-effectiveness" assessments to determine which products should be

included in the formulary. Each product or service is assessed for the cost per unit of health benefit. Typically, health benefits are measured in terms of "Quality-Adjusted Life-Years" (QALYs) or some similar measure that allows for comparability across different diseases, conditions, and interventions. With a fixed budget, it is generally believed that the insurer can obtain the most QALYs by choosing the products or services with the greatest cost-effectiveness until its budget is used up. Of course, there are typically some exceptions to this rule. Some countries, for example, allow higher prices for drugs treating rare diseases. These are, however, exceptions to the general rule that insurers will spend their scarce resources paying for the interventions that do the most good per dollar spent.

The implementation of this rule is typically through the establishment of a cost-effectiveness threshold, which may be explicit or implicit. The UK's National Institute for Health and Care Excellence, for example, has set a threshold in the range of £20,000 to £30,000 per QALY, and rarely funds technologies costing above this threshold. Claxton et al. (2015) argue that this threshold is too high, since they estimate the marginal cost of a QALY in the UK's health care system to be under £13,000. Ireland has an explicit threshold of €41,000 (O'Mahony and Coughlan, 2016). Whatever the threshold is, the core idea is that interventions with a lower cost per QALY than the threshold should be funded, while those with a higher cost per QALY represent a poor use of scarce resources (Culyer, 2016). Given a fixed budget, paying for products and services that cost more than the threshold requires the insurer to exclude other services with lower cost per QALY. In these circumstances, the net impact of paying more than the threshold will be a reduction in population health.

In this paper, we demonstrate that the simple, widely used rule of a single threshold, despite its intuitive appeal, fails to engage with the strategic behaviour by companies or to use countervailing monopsony power appropriately. As Ellison and Snyder (2010) note, insurers that are willing to use restrictive formularies can effectively exercise countervailing power. Welfare maximization, in these circumstances, requires that monopolized products and services should face a lower national insurance threshold than products and services for which there is a competitive supply. The monopoly price will in general be either above the threshold (in which case the product is excluded from the formulary and sold only to cash-paying consumers) or at the threshold. It will never be profit-maximizing for a monopolist to set a price below the threshold, since insurance essentially fixes the demand for the product: insured consumers are insensitive to the price paid by the insurer to the monopolist. The implication is that lowering the threshold for monopolized products will have two different effects. It will reduce the number of

monopolized products available at the margin, but it will also generate cost savings on inframarginal products. These cost savings may be reallocated to the supply of competitively sourced products.

One of the arguments for paying more for monopolized products is the value of innovation. Our model fully incorporates the value of innovation; if the value of innovation is not accounted for, then the optimal threshold for patented products falls yet further. For example, if a country deems itself too small to affect investments in pharmaceutical R&D, it may, like New Zealand, choose a very low effective threshold for patented products. Our model also includes consideration of the limited duration of market power enabled by a patent. When the expected patent life is short enough, it is possible that the optimal threshold for patented products may be above that of unpatented products. We also show, in a world with many countries acting uncooperatively, that privately optimal thresholds will be too low, compared to the globally efficient threshold.

In this paper, we begin by providing a summary of the literature on setting the threshold for cost-effectiveness decisions. We then introduce a simple model of public drug insurance and compare the optimal thresholds for products with and without market power. We show that these thresholds must be different, and that this difference becomes larger for insurers that do not consider the effect of the threshold on incentives for innovation. We also show that even when the threshold is based endogenously on a fixed budget, the optimal threshold is lower for products with market power.

## 1.1 The cost-effectiveness threshold as a decision tool for insurers

In order to make decisions across different health technologies and services to be funded, it is necessary for the insurer to be able to compare interventions using a common measure of health benefit such as the QALY. This measure is clearly imperfect and does not necessarily represent marginal willingness to pay in the same way that an effective market would, but the insurer has little alternative, since it is an agent for many individuals with varying, unknown preferences. QALYs have attracted extensive criticism, but continue to be used because the alternatives are not obviously better.

The idea of using cost per QALY as a guide to what insurers should cover has a long history (Weinstein and Zeckhauser, 1973; Doubilet et al., 1986; Birch and Gafni, 1992; Garber and Phelps, 1997; George et al., 2001; Towse et al., 2002; Devlin and Parkin, 2004; Culyer et al., 2007; McCabe et al., 2008; Eckermann and Pekarsky, 2014; Newall et al., 2014; Danzon et al.,

2011, 2015; Marseille et al., 2015; Culyer, 2016). While there are considerable challenges to implementation, many state insurers use this approach explicitly (e.g. U.K, Thailand); others use it implicitly. Indeed, an insurer that failed to choose the most cost-effective interventions, while covering high-cost, ineffective interventions, would not be a very good insurer.

The "cost-effectiveness" threshold can be based on the opportunity cost in the health system or on the marginal willingness to pay in the population. Claxton et al. (2015) assess the opportunity cost, finding that the marginal cost per QALY in the UK's National Health Service (NHS) is approximately £13,000. This implies that, given a limited budget, funding any new technologies costing more than £13,000 per QALY will displace funding for more efficient interventions, thus reducing population health. Determining a threshold based on population willingness to pay is perhaps more challenging given the considerable differences between individuals, but is at least theoretically possible. In either case, the threshold can be seen as a tool for identifying which interventions should be funded, and for signalling to investors the financial value of innovations (Culyer, 2016).

National insurance systems do not cover the entire set of possible drugs that have been approved by the national regulatory authority, since many drugs do not meet the standards of cost-effectiveness required to be covered. Drugs excluded from the formulary have a limited market, with sales paid for out of pocket or through private insurance plans. In Canada, for example, private spending makes up more than half of the market for prescription drugs. We call out of pocket and private insurance plans the "free market."

# 2 A simple public insurance system

We investigate a system in which there is public health insurance alongside a free market.[1] We assume that the goal of the public insurer is to achieve the maximum total QALYs given its budget. For drugs submitted to to be reimbursed by public health insurance, there is no free-market mechanism to induce price revelation. The standard solution to this problem, as discussed above, is for the insurer to set a threshold for cost per QALY and to cover all the drugs with cost per QALY below that threshold. The public insurer is assumed to offer insurance to the entire population for any covered product or service with a zero copayment.

We adopt, for the purposes of modelling, an unusual convention. Rather

[1] Public insurance never covers every drug with regulatory approval, and private insurers, such as BUPA in the UK, typically fund a wider set of drugs. Patients may also pay out of pocket for drugs not covered by their insurance plans.

than counting quantity through the number of pills or patients treated, we count the QALYs created by the drug. In essence, the drug's quantity is measured through the services that it provides to health, rather than through the number of pills. This makes all drugs comparable in their quantity dimension. This modelling approach embodies the implication that the insurer is indifferent to the way that health is improved and is interested only in the number of QALYs gained. Note, however, that drugs are not necessarily substitutes for a given individual since an individual will need only those drugs indicated for his condition. For example, a treatment for malaria will provide QALYs to a patient only if she has malaria.

The effect of public insurance on demand for a single product or service is illustrated in Figure 1. In the absence of public insurance, the demand curve has a typical downward slope. However, once price is at or below the threshold $T_p$, the insurer adds the product to the formulary, effectively reducing the price to all consumers to zero. The presence of the public insurer changes the behaviour of the firms in the free market: firms which charge slightly more than $T_p$ in the free market without public insurance know that if they decrease the price of the drugs to $T_p$ in a system with public insurance, the quantity demanded will jump to $q^0 \equiv q(0)$. (If the insurer requires copayments by patients, the quantity would adjust, without changing the direction of the results in this paper.) Similarly, the threshold allows any firm with market power to increase the price of its product to $T_p$ with no reduction in quantity. This implies that a monopolist will never set a price below $T_p$. When there is competition in supply, in contrast, the competitive supply curve will determine the price.

## 2.1 Public insurance for patented health products

Pharmaceuticals are the most significant class of patented health products. Innovative drugs and biologics are usually protected by numerous patents on the constituent molecules, manufacturing processes and the like (Hemphill and Sampat, 2011). Following preliminary research, a product will undergo lengthy and expensive clinical trials. If the firm successfully develops the drug, it can sell it in the free market at a profit-maximizing price. If the firm submits the drug to the public health insurance system, it will offer a price per QALY designed to obtain inclusion in the formulary, *i.e.*, at the threshold. The price of patented drugs is generally much higher than the cost of producing them and for simplicity, we assume the marginal cost of production for patented products is equal to zero. For convenience, we will think of patented products as being interchangeable with patented drugs, while recognizing that not every patented drug may benefit from market

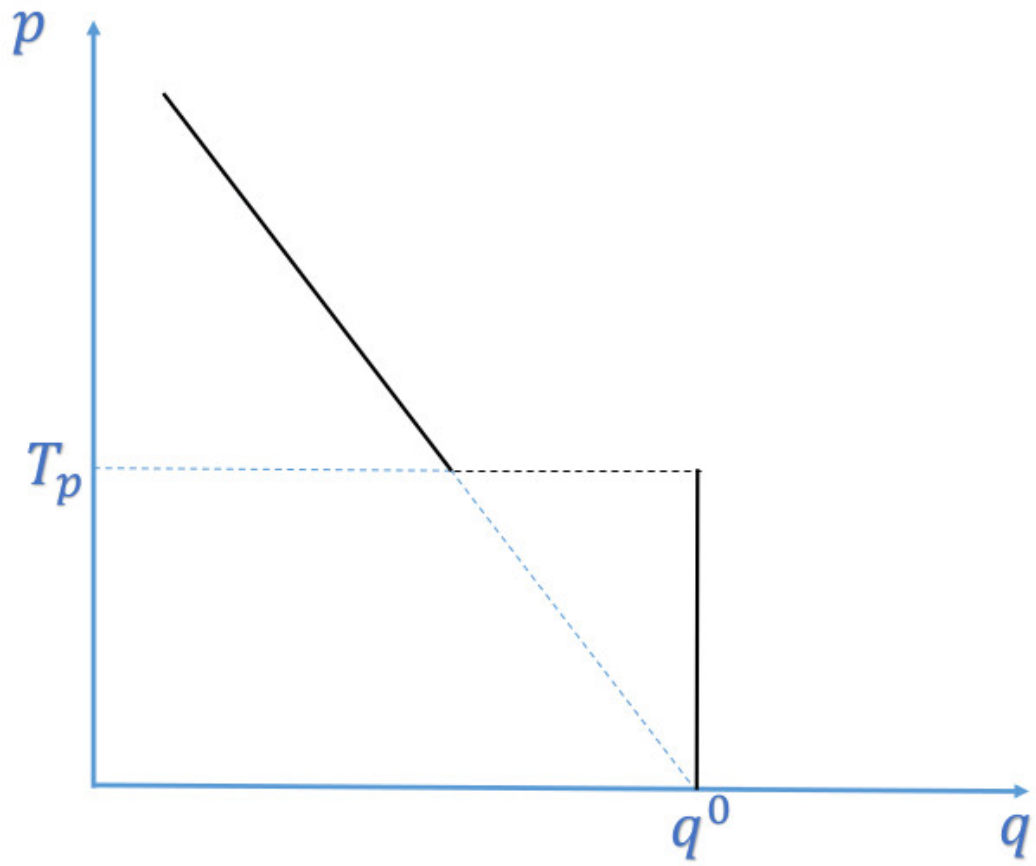



Figure 1: Demand curve for a drug in a system with free market and public insurance system.

power, and not every patented product is a drug or biologic.

We denote the profit of drug $i$, if developed, by $\pi_i$. The probability of discovery $\phi(\pi_i)$ is an increasing function of its profit (Dubois et al., 2015). For simplicity, we assume this function is the same for all potential drugs. We note that $\phi(\pi_i)$ can be zero, depending on how small the expected profit is. We assume that the process of innovation is monopolistically competitive – thus, firms will invest in drug innovation as long as the expected profits are positive.

We begin by examining the situation in which there is a "free market" into which firms sell their drugs. We define $q_i(p_i)$ as the demand for QALYs from drug $i$ in a given country, if the firm charges $p_i$ per QALY. (To be clear, we ignore the possibility of arbitrage between countries.) The firm's maximization problem is $\max_{p_i} p_i q_i(p_i)$. If we define $p_i^* \equiv \operatorname{argmax}_{p_i} p_i q_i(p_i)$ and

$q_i^* \equiv q_i(p_i^*)$, then the firm's maximum profit from selling the drug in that country would be:

$$\pi_i = p_i^* q_i^*. \tag{1}$$

To facilitate comparisons between the free market and the public insurance system, we calculate the minimum price per QALY that a public health insurer would have to pay per QALY to generate at least as much profit as the free market. The public insurer in our model provides covered drugs to patients at no cost to the patients. Thus, if we define $q_i^0 \equiv q_i(0)$, a public insurer buys $q_i^0$ of the drug $i$ from the firm. Here $q_i^0$ can be interpreted as the total potential QALYs for drug $i$. We assume that only patients who need drug $i$ buy it, and only in the amount prescribed, so $q_i^0$ is not infinite. If we define $p_i^0 \equiv p_i^* \frac{q_i^*}{q_i^0}$, the public health insurer would have to pay $p_i^0$ per unit of QALY to the manufacturer to make it indifferent between the free market and the public insurance system. Given these definitions, we can now write the profit of the drug $i$ in the free market as $\pi_i = p_i^0 q_i^0$.

Recall that $q_i^0$ is the total potential QALYs for drug $i$ when sold at a zero price, which is a function only of the underlying characteristics of the drug in terms of its therapeutic value per patient and the number of patients who would benefit from it. Thus $\frac{q_i^*}{q_i^0}$ is the proportion of the total potential QALYs which would be achieved, given monopoly pricing instead of marginal cost pricing. Note that $p_i^0$ is likely to vary across drugs because (a) the willingness to pay of patients with a given disease varies and (b) the shape of the demand curve can affect both $p_i^*$ and $\frac{q_i^*}{q_i^0}$. For a disease suffered mainly by relatively poor people, for example, a relatively low value of $p_i^0$ will make the firm indifferent between the free market and public insurance.

Thus each drug is characterized by the two variables $p^0$ and $q^0$ and can therefore be represented in continuous $p^0 - q^0$ space. We define the probability density function of candidate drugs as $\omega_p(p^0, q^0)$. This means the number of candidate drugs in this space is $\int\int \omega_p(p^0, q^0)\phi(\pi(p^0, q^0))dq^0 dp^0$.

As discussed above, the public insurer sets a threshold for cost per QALY. We denote the threshold cost per QALY for patented drugs by $T_p$. Firms seeking coverage by the public insurer will set the price per QALY of their drugs equal to that threshold. Then the profit of a drug at $(\tilde{p}^0, \tilde{q}^0)$ in the public health system is $T_p\tilde{q}^0$, while its profit in the free market is $\tilde{p}^0\tilde{q}^0$. If for a drug $p^0 < T_p$, firms will submit it to the public insurance system and otherwise they will sell it in the free market. Therefore the public health insurer ends up covering the drugs with $p^0 < T_p$, and the profit for these drugs would be $\pi(p^0, q^0) = T_p q^0$.

Since $T_p$ is the threshold for cost per QALY in the public insurance system, the total QALYs for all the patented drugs covered by public insurance must be:

$$Q_p(T_p) = \int_0^{T_p} \int_0^{\infty} \omega_p(p^0, q^0)\phi(T_p q^0) q^0 dq^0 dp^0. \tag{2}$$

Note that $Q_p$ is an increasing function of $T_p$ and the first derivative can be expressed as:

$$\frac{\partial Q_p(T_p)}{\partial T_p} = \psi_p(T_p) + \psi_\phi(T_p).$$

where

$$\psi_p(T_p) \equiv \int_0^{\infty} \omega_p(T_p, q^0)\phi(T_p q^0) q^0 dq^0,$$

$$\psi_\phi(T_p) \equiv \int_0^{T_p} \int_0^{\infty} \omega_p(p^0, q^0) \frac{\partial \phi}{\partial \pi} (q^0)^2 dq^0 dp^0,$$

Both $\psi_p$ and $\psi_\phi$ are positive because all the terms in the integrals are positive. Here $\psi_p$ represents the rate of QALYs that switch to the public insurance system at the boundary of the insurance system and the free market. Similarly, $\psi_\phi$ represents the increase in the probability of development for all drug candidates that, if developed, will be covered by public health insurance.

We can interpret $Q_p(T_p)$ as the patented drug supply function for QALYs; if a public insurer set the price of QALYs for patented products equal to $T_p$, then $Q_p(T_p)$ of QALYs would be supplied to the system. So we just showed that this supply function is increasing in term of $T_p$, which is what we would expect. We assume that this supply function is weakly concave, so that $\partial^2 Q_p(T_p)/\partial T_p^2 \leqslant 0$.

## 2.2 Public insurance for unpatented health products

Most services covered under health insurance are not patented pharmaceuticals. Health insurance covers hospital services, doctors, other health professionals, and unpatented products including generic drugs. We assume perfect competition for these products so that $p_i = c_i$ in both the free market and the public health insurance system, where the cost per QALY of a product or service is denoted by $c_i$. We note that marginal cost for these unpatented products may be much greater than zero. Then the profit is $\pi_i = (p_i - c_i)q_i = 0$. As before, the QALYs of a product in the public insurance system will be $q^0$. Without coverage, fewer people would be able to purchase the product, and the QALYs would be smaller than $q^0$. Each

unpatented health product can be characterized by two variables $c$ and $q^0$ and can therefore be represented in the continuous $c - q^0$ space.

We denote the cost threshold per QALY for unpatented products by $T_u$. All products and services with $c < T_u$ will be covered by public health insurance; any more expensive products may have some sales in the free market. We define the probability density function of possible products and services as $\omega_u(c, q^0)$. Thus, the total QALYs of unpatented health products, covered by public health insurance will be:

$$Q_u(T_u) = \int_0^{T_u} \int_0^{\infty} \omega_u(c, q^0) q^0 dq^0 dc. \tag{4}$$

Note that $Q_u$ is an increasing function of $T_u$. If we increase $T_u$, the threshold would be higher and more drugs would switch from free market to the public health insurance:

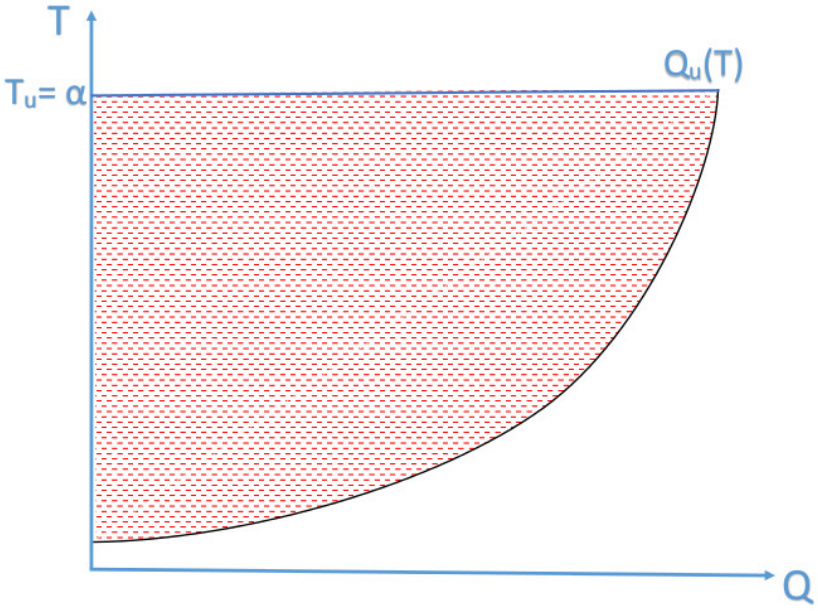
$$\frac{\partial Q_u(T_u)}{\partial T_u} = \int_0^{\infty} \omega_u(T_u, q^0) q^0 dq^0.$$

All of the terms in the integral are positive, so $\partial Q_u(T_u)/\partial T_u > 0$.

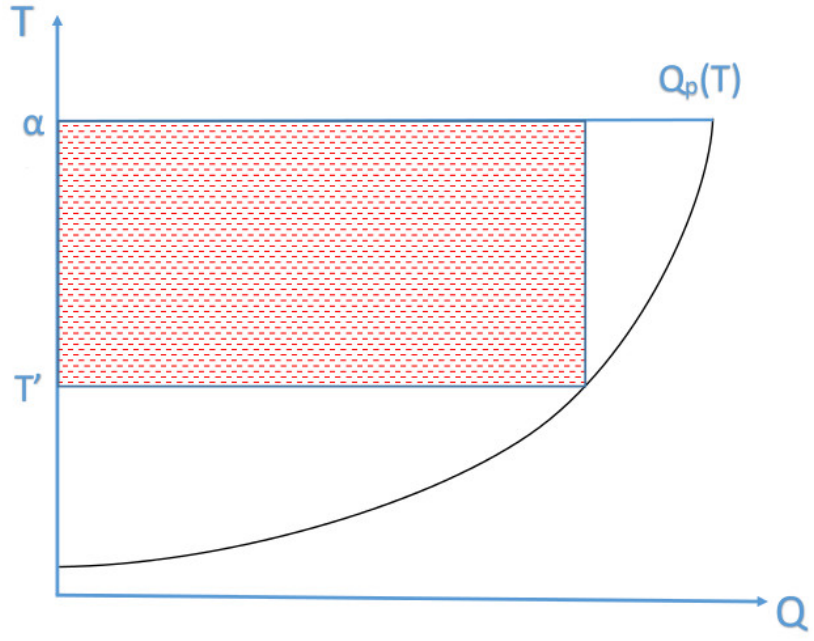
# 3 Optimal thresholds for a public health insurer with a constant willingness to pay

In this section we assume that the public insurer has a known willingness to pay per QALY. For expositional simplicity, we assume that this willingness to pay is constant and equal to $\alpha$. (At least, the following discussion requires that within a relevant range, the demand for QALYs is perfectly elastic.) As in the previous section, we denote supply functions for patented and unpatented health products by $Q_p(T_p)$ and $Q_u(T_u)$. Since the demand for QALYs is perfectly elastic, we can investigate the surplus from patented and unpatented products separately, as illustrated in Figure 2. This figure shows the aggregate supply in QALYs of health products; the upper figure shows supply for unpatented, and the lower figure supply for patented, products.

These figures clarify the position of the public insurer as a monopsony, which seeks to maximize its surplus. For unpatented, competitively-supplied products, the public insurer can perfectly price discriminate between products, in that each product is purchased at its marginal cost. Therefore the surplus in the unpatented market is the entire shaded area, and the optimal threshold is $T = \alpha$. In contrast, producers of patented products, because of their monopoly power, will price up to the threshold and the public insurer




(a)




(b)

Figure 2: Supply and demand of QALYs for (2a) unpatented and (2b) patented health products in the public health insurance system. The shaded areas show the surplus for the public insurer.

cannot price discriminate between products. In this case, as shown in the lower panel, given any threshold $T'$, the surplus will be equal to the shaded area above $T'$. This implies the following proposition:

**Proposition 1.** *The optimal threshold cost per QALY for patented products is below the willingness to pay, while the optimal threshold for unpatented health products is equal to willingness to pay. Specifically:*

$$T_p^* = \alpha - \frac{Q_p}{(\partial Q_p/\partial T_p)}, \tag{5a}$$

$$T_u^* = \alpha. \tag{5b}$$

*Proof.* If the willingness to pay of the public insurer per QALY is $\alpha$ and it pays $p$ per QALY, the surplus it receives from that QALY is $\alpha - p$. As we discussed, the price per QALY for patented drugs in the public health insurance system is $T_p$, so the surplus per QALY would be $\alpha - T_p$. Thus, given Equation 2, the total surplus of the insurer can be written as:

$$\mathcal{S}_p = (\alpha - T_p)Q_p(T_p).$$

The first-order condition with respect to $T_p$ is given by

$$\frac{\partial \mathcal{S}_p}{\partial T_p} = -Q_p(T_p^*) + (\alpha - T_p^*)(\partial Q_p/\partial T_p) = 0 \tag{6}$$

$$\Rightarrow T_p^* = \alpha - \frac{Q_p}{(\partial Q_p/\partial T_p)} = \alpha - \frac{Q_p(T_p^*)}{\psi_p(T_p^*) + \psi_\phi(T_p^*)}. \tag{7}$$

To ensure that this solution is a maximum, the second order condition requires $\partial^2 \mathcal{S}_p/\partial T_p^2 < 0$ at $T_p^* = \alpha - \frac{Q_p}{(\partial Q_p/\partial T_p)}$:

$$\frac{\partial^2 \mathcal{S}_p}{\partial T_p^2} = -2\frac{\partial Q_p}{\partial T_p} + (\alpha - T_p)\frac{\partial^2 Q_p}{\partial T_p^2}$$

If we set $T_p = T_p^* = \alpha - \frac{Q_p}{(\partial Q_p/\partial T_p)}$,

$$\frac{\partial^2 \mathcal{S}_p}{\partial T_p^2} = -2\frac{\partial Q_p}{\partial T_p} + \frac{Q_p}{(\partial Q_p/\partial T_p)}\frac{\partial^2 Q_p}{\partial T_p^2}$$

We know $Q_p > 0$ and $\partial Q_p/\partial T_p > 0$, and assumed that $\partial^2 Q_p/\partial T_p^2 \leqslant 0$. Therefore $\partial^2 \mathcal{S}_p/\partial T_p^2 < 0$.

For unpatented products,

$$\mathcal{S}_u = \int_0^{T_u} \int_0^{\infty} (\alpha - c)\omega_u(c, q^0)q^0 dq^0 dc.$$

The first-order condition with respect to $T_u$ is given by

$$\frac{\partial \mathcal{S}_u}{\partial T_u} = \int_0^{\infty} (\alpha - T_u)\omega_u(T_u, q^0)q^0 dq^0 = (\alpha - T_u^*)\frac{\partial Q_u}{\partial T_u} = 0,$$

implying that

$$T_u^* = \alpha.$$

For the second order condition to be satisfied, $\partial^2 \mathcal{S}_u / \partial T_u^2 < 0$ at $T_u = T_u^* = \alpha$:

$$\frac{\partial^2 \mathcal{S}_u}{\partial T_u^2} = -\frac{\partial Q_u}{\partial T_u} + (\alpha - T_u)\frac{\partial^2 Q_u}{\partial T_u^2}$$

Given $T_u = T_u^* = \alpha$,

$$\frac{\partial^2 \mathcal{S}_u}{\partial T_u^2} = -\frac{\partial Q_u}{\partial T_u} < 0.$$

■

The first term in Equation 6 corresponds to strategic behavior by patentees: if the threshold is increased, the price of inframarginal patented drugs will increase to the new threshold. While it is understood that monopolists will price up to the threshold, it has not been noted in the literature that this will change the optimal threshold for these products. If we do not consider this term, we would have: $\partial \mathcal{S}_p / \partial T_p = (\alpha - T_p)(\psi_p + \psi_\phi)$. In this case the optimal threshold would be $T_p^* = \alpha$, which is the same as $T_u^*$. So we would have one threshold for patented and unpatented products. In effect, the reason that the optimal threshold for patented drugs should be different from the optimal threshold for unpatented products is the response to strategic behaviour by firms with market power.

Figure 2 gives some further intuition for why the optimal thresholds for patented and unpatented health products are different. The public health insurer is only concerned about maximizing QALYs. When the threshold for patented products is decreased from $\alpha$, the spending of the insurer decreases, and, assuming a monopolistically competitive industry, total costs of innovators decrease by the same amount. The change in surplus from decreasing the threshold from $\alpha$, as shown by Equation 6, is then equal to $Q_p(\alpha)$. As the threshold decreases further, the loss from decreased innovation gets bigger and the gain from saving gets smaller, so decreasing the threshold becomes

less attractive. At the optimal threshold, the loss from decreased innovation is balanced against the savings from further spending reductions. In essence, this is a standard monopsony situation, except that the commodity being purchased is healthcare innovation.

We assume a model of monopolistic competition for patented products in the health system. That is, while there are quasi-rents available to firms that are successful in bringing new products to market, overall the rate of economic profits in the industry is zero. If the expected profit for a given potential drug is positive, more firms will enter the market, increasing investment in that drug until the expected profits falls to zero. So, the total surplus of the public insurer is the same as total surplus of the society. Therefore the optimal threshold calculated in this section is the same as the optimal threshold for society.

What is unusual about this exercise of monopsony power by the insurer is that all the patented products also have market power. The insurer strips away some of this market power by value QALYs, – rather than the treatment offered by a particular drug – and thus forcing firms to compete against each other even though the products are not substitutes in their uses. It is, however, exactly this effort to put firms on an equal footing that creates the opportunity for firms to increase their prices up to the threshold.

## 3.1 Nash Equilibrium solution

We now consider a world with many countries, each able to set its own threshold cost per QALY. We are able to characterize the Nash Equilibrium threshold of each country for patented and unpatented drugs, and how those thresholds relate to the case with only one country described above. We assume symmetry across $N$ countries, each of which has the same willingness to pay for QALYs, $\alpha$.

If the total global potential QALYs of a drug is $q^0$, the potential QALYs of that drug in each country is $\frac{1}{N}q^0$. If country $n$ sets its threshold for patented drugs equal to $T_p$, while all other countries choose $T_{p,-n}$, the total profit of a drug is $\frac{N-1}{N}T_{p,-n}q_0 + \frac{1}{N}T_p q_0$. Then we can write the total QALYs of patented drugs in country $n$ as:

$$\bar{Q}_p(T_p, T_{p,-n}) = \int_0^{T_p}\int_0^{\infty} \omega_p(p^0, q^0)\phi(\frac{N-1}{N}T_{p,-n}q_0 + \frac{1}{N}T_p q_0)\frac{1}{N}q^0 dq^0 dp^0.$$

We can now calculate the Nash Equilibrium threshold for patented drugs:

**Proposition 2.** *The Nash Equilibrium threshold for patented drugs in each country is smaller than the globally optimal threshold for patented drugs.*

*Specifically if there are $N$ similar countries, the Nash Equilibrium solution for each country would be:*

$$T_p^{NE} = \alpha - \frac{Q_p}{\psi_p + (1/N)\psi_\phi}. \tag{8}$$

*Proof.* The surplus of the public insurer in country $n$ is:

$$\bar{\mathcal{S}}_p = (\alpha - T_p)\bar{Q}_p(T_p, T_{p,-n}).$$

The First Order Condition requires:

$$\frac{\partial \bar{\mathcal{S}}_p}{\partial \bar{T}_p} = -\bar{Q}_p(T_p^{NE}, T_{p,-n}^{NE}) + (\alpha - T_p^{NE})\frac{\partial \bar{Q}_p(T_p^{NE}, T_{p,-n}^{NE})}{\partial \bar{T}_p} = 0.$$

Given symmetry between countries, $T_{p,-n}^{NE} = T_p^{NE}$, so that

$$\bar{Q}_p(T_p^{NE}, T_{p,-n}^{NE}) = \frac{1}{N}Q_p(T_p^{NE}),$$

$$\frac{\partial \bar{Q}_p(T_p^{NE}, T_{p,-n}^{NE})}{\partial \bar{T}_p} = \frac{1}{N}(\psi_p(T_p^{NE}) + \frac{1}{N}\psi_\phi(T_p^{NE}))$$

Here $Q_p$, $\psi_p$ and $\psi_\phi$ are the same function as in the previous section. Thus the NE solution would be:

$$T_p^{NE} = \alpha - \frac{Q_p(T_p^{NE})}{\psi_p(T_p^{NE}) + \frac{1}{N}\psi_\phi(T_p^{NE})}$$

The next step is to show that $T_p^{NE} < T_p^*$, where $T_p^*$ is the solution from previous section. We use proof by contradiction. First, assume $T_p^{NE} = T_p^*$, from Equation 5a and Equation 8:

$$\alpha - \frac{Q_p(T_p^{NE})}{\psi_p(T_p^{NE}) + \frac{1}{N}\psi_\phi(T_p^{NE})} = \alpha - \frac{Q_p(T_p^*)}{\psi_p(T_p^*) + \psi_\phi(T_p^*)}$$

$$\Rightarrow \frac{Q_p(T_p^{NE})}{\psi_p(T_p^{NE}) + \frac{1}{N}\psi_\phi(T_p^{NE})} = \frac{Q_p(T_p^*)}{\psi_p(T_p^*) + \psi_\phi(T_p^*)}$$

If we substitute $T_p^{NE}$ for $T_p^*$ in the last equation, we obtain:

$$\Rightarrow \frac{Q_p(T_p^*)}{\psi_p(T_p^*) + \frac{1}{N}\psi_\phi(T_p^*)} = \frac{Q_p(T_p^*)}{\psi_p(T_p^*) + \psi_\phi(T_p^*)},$$

which is only true for $N = 1$. Next we assume that $T_p^{NE} > T_p^*$. We know that $Q_p(T_p)$ is an increasing function, so $Q_p(T_p^{NE}) > Q_p(T_p^*)$. Given our earlier assumption that $\partial^2 Q_p(T_p)/\partial T_p^2 \leqslant 0$:

$$\psi_p(T_p^{NE}) + \psi_\phi(T_p^{NE}) \leqslant \psi_p(T_p^*) + \psi_\phi(T_p^*)$$

$$\Rightarrow \psi_p(T_p^{NE}) + \frac{1}{N}\psi_\phi(T_p^{NE}) < \psi_p(T_p^*) + \psi_\phi(T_p^*)$$

Given $Q_p(T_p^{NE}) > Q_p(T_p^*)$ and the last equation, we would have:

$$\Rightarrow \frac{Q_p(T_p^{NE})}{\psi_p(T_p^{NE}) + \frac{1}{N}\psi_\phi(T_p^{NE})} > \frac{Q_p(T_p^*)}{\psi_p(T_p^*) + \psi_\phi(T_p^*)}$$

$$\Rightarrow \alpha - \frac{Q_p(T_p^{NE})}{\psi_p(T_p^{NE}) + \frac{1}{N}\psi_\phi(T_p^{NE})} < \alpha - \frac{Q_p(T_p^*)}{\psi_p(T_p^*) + \psi_\phi(T_p^*)} \Rightarrow T_p^{NE} < T_p^*.$$

We started with an assumption that $T_p^{NE} > T_p^*$ but conclude that $T_p^{NE} < T_p^*$, which is impossible. Therefore $T_p^{NE}$ cannot be equal to or greater than $T_p^*$. ■

When $1/N$ approaches zero, for each country $T_p^{NE}$ approaches $\alpha - Q_p/\psi_p$. That is the smallest threshold a country would set for patented drugs. This solution is intuitively sensible since small countries cannot influence innovation and therefore do not consider it in setting a threshold. The upper limit for the threshold for patented drugs is when $N = 1$, which is the globally optimal solution.

The Nash Equilibrium threshold for *unpatented* drugs is the same as the globally optimal solution, which is equal to willingness to pay $\alpha$. So as $N$ increases, the difference between the privately optimal thresholds for patented drugs and unpatented products will also increase.

## 3.2 Considering limited patent duration

One of the important features of patents is that the market power they enable is only temporary. For a given product, therefore, we can expect that there will be a period during which the patentee is able to price up to the threshold, and then a period during which price will be determined by marginal cost, which we approximate as zero. This, of course, allows the insurer to capture very different amounts of surplus. If the patent period covers $m$ years, the drug is sold for $M$ years in total, and the discount rate is $r$, we can multiply our one-year values by $\sum_{t=1}^{m}(1+r)^t$; and the

post-patent period as $\sum_{t=m+1}^{M}(1+r)^n$. To facilitate tractability, we define $\gamma \equiv \sum_{t=1}^{m}(1+r)^t / \sum_{t=m+1}^{M}(1+r)^n < 1$.

Because of the future post-patent price reduction, pharmaceutical companies argue that the threshold for patented products should be higher than that of unpatented products. To investigate the effects of considering the future price reduction of patented drugs, we assume that $\gamma$ portion of the benefit would be realized before the patent expiry and $1-\gamma$ portion of the benefit would be realized after the patent. Some drugs may exist and not be covered at all by the public insurance system during the patent, the profit from the private market being sufficient incentive for investment in their development. We denote the total QALYs of these drugs by $Q_{ex}$:

$$Q_{ex}(T_p) = \int_{T_p}^{\infty}\int_{0}^{\infty} \omega_p(p^0, q^0)\phi(p^0 q^0) q^0 dq^0 dp^0.$$

The public insurer will benefit from these drugs in the future, even though they are not included in the formulary before patent expiry. The insurer would pay marginal cost (which we assume is zero) for the drugs after patent expiry.

The surplus of the drugs which *are* covered in the public insurance system is $\gamma(\alpha - T_p)Q_p$ before the patent expiry and $(1-\gamma)\alpha Q_p$ after the patent expires. The surplus of the drugs not covered by public insurance system is $(1-\gamma)\alpha Q_{ex}$ after the patent expiry. Considering all of these, we now calculate the optimal threshold for patented drugs.

**Proposition 3.** *If the public insurer includes consideration of future price reductions for patented drugs, the optimal threshold for patented drugs increases, compared to the case without such consideration. The optimal threshold can even be greater than the willingness to pay $\alpha$:*

$$T_p^* = \alpha \frac{\psi_p + \psi_\phi/\gamma}{\psi_p + \psi_\phi} - \frac{Q_p}{\psi_p + \psi_\phi}$$

*Proof.* The total surplus of the insurer from patented drugs is:

$$\mathcal{S}_p = \gamma(\alpha - T_p)Q_p + (1-\gamma)\alpha Q_p + (1-\gamma)\alpha Q_{ex} = (\alpha - \gamma T_p)Q_p + (1-\gamma)\alpha Q_{ex}.$$

The first derivative of $\mathcal{S}_p$ is given by

$$\frac{\partial \mathcal{S}_p}{\partial T_p} = -\gamma Q_p + (\alpha - \gamma T_p)\frac{\partial Q_p}{\partial T_p} + (1-\gamma)\alpha \frac{\partial Q_{ex}}{\partial T_p}.$$

Next we calculate $\partial Q_{ex}/\partial T_p$:

$$\frac{\partial Q_{ex}}{\partial T_p} = -\int_{0}^{\infty} \omega_p(T_p, q^0)\phi(T_p q^0) q^0 dq^0 = -\psi_p.$$

We also know that $\partial Q_p/\partial T_p = \psi_p + \psi_\phi$. So we can write the first derivative of surplus for patented drugs as:

$$\frac{\partial \mathcal{S}_p}{\partial T_p} = -\gamma Q_p + (\alpha - \gamma T_p)(\psi_p + \psi_\phi) - (1-\gamma)\alpha\psi_p$$

$$\Rightarrow \frac{\partial \mathcal{S}_p}{\partial T_p} = -\gamma Q_p + (\gamma\psi_p + \psi_\phi)\alpha - \gamma(\psi_p + \psi_\phi)T_p. \tag{9}$$

Setting $\partial \mathcal{S}_p/\partial T_p = 0$ generates $T_p^*$. If $\gamma$ is small enough, then $T_p^* > \alpha$.

■

If the insurer ignores strategic behaviour by firms, instead of Equation 9, we would have:

$$\frac{\partial \mathcal{S}_p}{\partial T_p} = (\gamma\psi_p + \psi_\phi)\alpha - \gamma(\psi_p + \psi_\phi)T_p.$$

In this case, the threshold for patented drugs would be:

$$T_p^* = \alpha\frac{\psi_p + \psi_\phi/\gamma}{\psi_p + \psi_\phi}.$$

For unpatented products we still have $T_u^* = \alpha$, therefore the relationship between optimal threshold for patented and unpatented product should be:

$$T_p^* = T_u^*\frac{\psi_p + \psi_\phi/\gamma}{\psi_p + \psi_\phi}.$$

Ignoring strategic behavior by firms, the conventional wisdom that the optimal threshold for patented products should be greater than the optimal threshold for unpatented drugs would hold (since $\gamma < 1$). However, as we showed, adding strategic behavior by firms overturns this conventional wisdom.

Now we investigate the Nash Equilibrium threshold for patented drugs, if the public insurers consider the future price reductions for patented drugs. We use all the assumptions of the previous section. Additionally, we denote the QALYs of drugs which are not covered by public insurance system in this country during the patent period by $\bar{Q}_{ex}$:

$$\bar{Q}_{ex}(T_p) = \int_{T_p}^{\infty}\int_0^{\infty} \omega_p(p^0, q^0)\phi(p^0 q^0)\frac{1}{n}q^0 dq^0 dp^0 = \frac{1}{n}Q_{ex}(T_p).$$

We can now investigate the Nash Equilibrium solution, when we consider the future price reductions for patented drugs:

**Corollary 1.** *The Nash Equilibrium threshold for patented drugs is greater when the insurers consider the future price reductions for patented drugs, compared to the case without such consideration. But is it smaller than the globally optimal threshold. Specifically:*

$$T_p^{NE} = \alpha \frac{\psi_p + \psi_\phi/(N\gamma)}{\psi_p + \frac{1}{N}\psi_\phi} - \frac{Q_p}{\psi_p + \frac{1}{N}\psi_\phi}$$

*Proof.* The surplus of the insurer in country $n$ from patented drugs is:

$$\bar{\mathcal{S}}_p = (\alpha - \gamma T_p)\bar{Q}_p + (1-\gamma)\alpha\bar{Q}_{ex}.$$

To calculate the Nash Equilibrium solution, we use the First-Order Condition:

$$\frac{\partial \bar{\mathcal{S}}_p}{\partial T_p} = -\gamma\bar{Q}_p + (\alpha - \gamma T_p^{NE})\frac{\partial \bar{Q}_p}{\partial T_p} + (1-\gamma)\alpha\frac{\partial \bar{Q}_{ex}}{\partial T_p} = 0.$$

Next we calculate $\partial\bar{Q}_{ex}/\partial T_p$:

$$\frac{\partial \bar{Q}_{ex}}{\partial T_p} = -\int_0^\infty \omega_p(T_p, q^0)\phi(T_p q^0)\frac{1}{N}q^0 dq^0 = -\frac{1}{N}\psi_p.$$

If we substitute for $\bar{Q}_p$, $\partial\bar{Q}_p/\partial T_p$ and $\partial\bar{Q}_{ex}/\partial T_p$, the first order condition can be written as:

$$\frac{\partial \bar{\mathcal{S}}_p}{\partial T_p} = -\gamma\frac{1}{N}Q_p + (\alpha - \gamma T_p^{NE})\frac{1}{N}(\psi_p + \frac{1}{N}\psi_\phi) - (1-\gamma)\alpha\frac{1}{N}\psi_p = 0.$$

Rearranging completes the proof.

■

As discussed above, as $\frac{1}{N}$ approaches zero, the Nash Equilibrium threshold in Corollary 1 simplifies to:

$$T_p^{NE} = \alpha - \frac{Q_p}{\psi_p},$$

This is the same result as when there is no consideration of future price reductions. Small countries which cannot influence innovation do not set their threshold with a view to future price reductions. The intuition for this result is straightforward: a large country will invest more (proportionally) because its threshold affects the available set of innovations, while the small country acts as a free-rider. In this case we can clearly see that this solution is smaller than willingness to pay. The threshold for unpatented products is still equal to willingness to pay. Therefore for small countries, the threshold for patented drugs is always smaller than the threshold for unpatented products.

# 4 Optimal thresholds for a public health insurance with a fixed budget

In this section, we show that the same results hold when the insurer has a fixed budget, but not some pre-defined willingness to pay. We assume that the goal of the public health insurer is to maximize total QALYs, given its limited budget. The general assumption in the literature is that this goal can be achieved with a single threshold cost per QALY (Culyer, 2016). We show that the insurer should still set different thresholds for patented and unpatented health products.

**Proposition 4.** *The threshold cost per QALY for patented products should be smaller than the threshold for other unpatented health products. Specifically* $T_u^* - T_p^* = \dfrac{Q_p}{(\partial Q_p / \partial T_p)}$.

*Proof.* Denote the budget by $B$. Since each QALY will cost $T_p$ for patented products, the budget for these products in the public insurance system will be:

$$B_p(T_p) = \int_0^{T_p} \int_0^{\infty} T_p \omega_p(p^0, q^0) \phi(T_p q^0) q^0 dq^0 dp^0. \tag{10}$$

And since the public insurer pays $c_i$ per unit of QALY for unpatented products, the total budget for these products must be:

$$B_u(T_u) = \int_0^{T_u} \int_0^{\infty} c \omega_u(c, q^0) q^0 dq^0 dc. \tag{11}$$

We can write the maximization problem as:

$$\max_{T_p, T_u} \left( Q_p(T_p) + Q_u(T_u) \right)$$

$$\text{s.t. } B_p(T_p) + B_u(T_u) = B.$$

The Lagrangian for this problem is:

$$\mathcal{L} = Q_p(T_p) + Q_u(T_u) + \lambda \left( B - B_p(T_p) - B_u(T_u) \right)$$

The first-order condition with respect to $T_u$ is given by

$$\frac{\partial \mathcal{L}}{\partial T_u} = \frac{\partial Q_u}{\partial T_u} - \lambda \frac{\partial B_u}{\partial T_u} = 0$$

$$\Rightarrow \int_0^{\infty} \omega_u(T_u^*, q^0) q^0 dq^0 - \lambda \int_0^{\infty} T_u^* \omega_u(T_u^*, q^0) q^0 dq^0 = 0$$

$$\Rightarrow \frac{\partial Q_u}{\partial T_u} - \lambda T_u^* \frac{\partial Q_u}{\partial T_u} = 0$$

$$\Rightarrow \lambda = \frac{1}{T_u^*} \tag{14}$$

The first-order condition with respect to $T_p$ is similarly

$$\frac{\partial \mathcal{L}}{\partial T_p} = (1 - \lambda T_p^*)\frac{\partial Q_p}{\partial T_p} - \lambda Q_p = 0. \tag{15}$$

For the second order condition to be satisfied, the determinant of the bordered Hessian matrix must be positive at $T_u = T_u^* = 1/\lambda$ and $T_p = T_p^* = 1/\lambda - \frac{Q_p}{(\partial Q_p/\partial T_p)}$. The bordered Hessian matrix for this problem is:

$$HL = \begin{pmatrix} 0 & -\frac{\partial B_p}{\partial T_p} & -\frac{\partial B_u}{\partial T_u} \\ -\frac{\partial B_p}{\partial T_p} & \mathcal{L}_{T_pT_p} & 0 \\ -\frac{\partial B_u}{\partial T_u} & 0 & \mathcal{L}_{T_uT_u} \end{pmatrix}$$

So determinant of the bordered Hessian matrix would be:

$$det(HL) = -\left( (\frac{\partial B_p}{\partial T_p})^2 \mathcal{L}_{T_uT_u} + (\frac{\partial B_u}{\partial T_u})^2 \mathcal{L}_{T_pT_p} \right).$$

Then given $\mathcal{L}_{T_uT_u} < 0$ and $\mathcal{L}_{T_pT_p} < 0$, the second order condition is satisfied. First we show that $\mathcal{L}_{T_uT_u}$ is smaller than zero:

$$\frac{\partial \mathcal{L}}{\partial T_u} = (1 - \lambda T_u^*)\frac{\partial Q_u}{\partial T_u}$$

$$\Rightarrow \frac{\partial^2 \mathcal{L}}{\partial T_u^2} = -\lambda \frac{\partial Q_u}{\partial T_u} + (1 - \lambda T_u^*)\frac{\partial^2 Q_u}{\partial T_u^2}.$$

Setting $\lambda = \frac{1}{T_u^*}$,

$$\frac{\partial^2 \mathcal{L}}{\partial T_u^2} = -\frac{1}{T_u^*}\frac{\partial Q_u}{\partial T_u}.$$

We know $T_u^* > 0$ and $\partial Q_u/\partial T_u > 0$, so that $\partial^2 \mathcal{L}/\partial T_u^2 < 0$. The next step is to show $\mathcal{L}_{T_pT_p} < 0$:

$$\frac{\partial \mathcal{L}}{\partial T_p} = (1 - \lambda T_p^*)\frac{\partial Q_p}{\partial T_p} - \lambda Q_p$$

$$\Rightarrow \frac{\partial^2 \mathcal{L}}{\partial T_p^2} = -2\lambda \frac{\partial Q_p}{\partial T_p} + (1 - \lambda T_p^*)\frac{\partial^2 Q_p}{\partial T_p^2}.$$

Using Equations 14 and 15, we can rewrite this equation as:

$$\frac{\partial^2 \mathcal{L}}{\partial T_p^2} = -2\frac{1}{T_u^*}\frac{\partial Q_p}{\partial T_p} + \frac{1}{T_u^*}\frac{Q_p}{\partial Q_p/\partial T_p}\frac{\partial^2 Q_p}{\partial T_p^2}.$$

We know $T_u^*$, $Q_p$ and $\partial Q_p/\partial T_p$ are positive and $\partial^2 Q_p/\partial T_p^2$ is negative, therefore $\partial^2 \mathcal{L}/\partial T_p^2 < 0$. Thus, the second order condition is satisfied.

Now we rewrite Equation 15 as:

$$T_p^* = \frac{(\partial Q_p/\partial T_p) - \lambda Q_p}{\lambda(\partial Q_p/\partial T_p)} = \frac{1}{\lambda} - \frac{Q_p}{(\partial Q_p/\partial T_p)}.$$

If we substitute $\lambda$ from Equation 14 and rearrange, we obtain:

$$T_u^* - T_p^* = \frac{Q_p}{(\partial Q_p/\partial T_p)} = \frac{Q_p}{\psi_p + \psi_\phi}. \tag{16}$$

Since $(\partial Q_p/\partial T_p) > 0$, the optimal threshold for patented health products $(T_p^*)$ must be lower than the threshold for unpatented products $(T_u^*)$. ■

# 5 Discussion

In this section we do rough calculations to get some intuition about different propositions in this paper. We first propose a simple functional form: $Q_p(T_p) = bT_p$. This is, for example, consistent with the claims of Abbott and Vernon (2007). Implementing this functional form in Proposition 1, we obtain $T_p^* = \alpha/2$ and $T_u^* = \alpha$. Thus for this case we have $T_p^* = T_u^*/2$. This means if we do not consider that patents are temporary, the global optimal threshold for patented products is only half of the optimal threshold for unpatented products.

As we stated in Proposition 2, the Nash Equilibrium solution of optimal threshold for patented drugs is smaller than the global optimal threshold. But for unpatented products, the Nash Equilibrium solution is the same as globally optimal solution which is $T_u^* = \alpha$. So if each country considers only its own surplus and does not consider the fact the patents are temporary, it should set the threshold for patented drugs less than half of the threshold for unpatented products. Also from this proposition we can see why smaller countries want to set their threshold for patented drugs below the threshold for larger countries.

If we wish to consider the temporary nature of patents, we need to know how large $\gamma$ is likely to be in practice. Exclusivity periods for new drugs are around 10 years, on average. In the US, approximately 90% of prescriptions

are filled with generic drugs. Thus, we could think of $m$ as being 10 and $M$ as being 100. With a discount rate of say 5%, we will find that $\gamma$ is approximately 0.4. If the discount rate is 10%, $\gamma$ is approximately 0.6. Therefore for this section we assume $\gamma$ is about 0.5.

This allows us to explore the optimal threshold for patented drugs, with due consideration of the limited duration of patents. If we again assume $Q_p(T_p) = bT_p$, from Proposition 3 we find that $\alpha/2 < T_p^* < \alpha$. The optimal threshold for *unpatented* products is still the same as before: $T_u^* = \alpha$. Thus, we should have $T_u^*/2 < T_p^* < T_u^*$, meaning the global optimal threshold for patented drugs is greater than half of the optimal threshold for unpatented products, but it is still smaller than the optimal threshold for unpatented products.

As described in Corollary 1, the Nash Equilibrium threshold for each country is smaller than the global optimal threshold for patented drugs; and the smaller countries are, the lower the threshold. For very small countries which cannot influence innovation, we have:

$$T_p^{NE} = \alpha - \frac{Q_p}{\psi_p} < \alpha - \frac{Q_p}{\psi_p + \psi_\phi} = \alpha - T_p^*.$$

Since $T_u^*/2 < T_p^*$, $T_p^{NE} < \alpha - T_p^* < \alpha - T_u^*/2$. Substituting $\alpha$ with $T_u^*$, we will get $T_p^{NE} < T_u^*/2$. This means that for small countries, the optimal threshold for patented drugs is smaller than half of the optimal threshold for unpatented products, even after considering the limited duration of patents.

We see, to a very limited extent, that the small/large country effect appears to play out in the real world. The United States pays high prices, with the explicit goal of supporting innovation. New Zealand, as a small country, doesn't appear to see itself as a having a role in supporting innovation in therefore is aggressive in limiting the set of drugs covered by its national insurance program. Nevertheless, in all countries, there appears to be a lack of understanding about the fact that the optimal threshold for patented products should be lower than the optimal threshold for unpatented, competitively provided products, because of strategic behaviour by patentees.

Finally, it is instructive to understand the results of this paper in the context of global negotiations about how to share the burden of paying for innovative pharmaceuticals. Before the implementation of the TRIPS agreement in 1995, many countries did not grant effective patent rights for pharmaceutical inventions, and so the effective price was simply the marginal cost of production by generic producers. (This, for example, included Canada until 1991.) Once all countries agreed to patent rights, we entered a world that resembled the Nash Equilibrium outcome described in Section 3.1, in which

countries attempted to control costs through the use of appropriate thresholds. Many insurers had little consideration for innovation – which they could hardly effect – and so attempted to maximize health benefits within a limited budget. (As we have described above, it is likely that owing to a failure to understand the strategic position of suppliers with market power, the thresholds chosen for patented products were likely generally higher than was privately optimal.) This has led to a further effort by one large country, the United States, to push on pricing policies in other countries, and to move them towards the "global optimum", which may be substantially higher than the Nash Equilibrium thresholds, as we have shown above. These efforts to push for higher pricing, rather than just longer patents, have accelerated in recent years as global standards for patent protection have solidified (USTR, 2018).

# 6 Conclusion

In this paper, we have shown that in the presence of market power of patentees, public insurers with countervailing monopsony power should use a lower threshold for products with market power than for products that are competitively provided. The reduced threshold for products with market power is driven by the fact that companies will price those products up to the threshold, thus creating the opportunity for the insurer to capture significant inframarginal gains. This result includes the effects of a reduced threshold on investment in innovation. Innovative companies often argue that because patents are time-limited, payers over-estimate the cost of innovative products in cost-effectiveness analyses; taking this into account means that generally the willingness to pay for innovative medicines should be increased, though not up to the willingness to pay. However, in a world in which each country is relatively small, the privately optimal threshold for each country is likely to be considerably below the global optimum, and there may be a role for coordinating pricing strategies across larger groups of countries to ensure that the incentives for innovation match its value.

# 7 Appendix

Until now, we have assumed a zero excess burden of taxation, so that paying for the public insurance system did not create any inefficiency. We now consider the outcome if we assume that there is an excess burden created by taxation.

**Corollary 2.** *When we account for the excess burden of taxation, the optimal thresholds decline.*

*Proof.* We implement the burden of taxation by multiplying the costs of public insurance by a scalar $\Lambda > 1$. This changes the total surplus of the insurer from patented drugs to:

$$\mathcal{S}_p = (\alpha - \gamma\Lambda T_p)Q_p + (1-\gamma)\alpha Q_{ex}.$$

The first derivative of $\mathcal{S}_p$ is given by

$$\frac{\partial \mathcal{S}_p}{\partial T_p} = -\gamma\Lambda Q_p + (\alpha - \gamma\Lambda T_p)\frac{\partial Q_p}{\partial T_p} + (1-\gamma)\alpha\frac{\partial Q_{ex}}{\partial T_p}.$$

$$\Rightarrow \frac{\partial \mathcal{S}_p}{\partial T_p} = -\gamma\Lambda Q_p + (\gamma\psi_p + \psi_\phi)\alpha - \gamma\Lambda(\psi_p + \psi_\phi)T_p.$$

If we set the derivative of surplus for patented drugs equal to zero, we obtain:

$$T_p^* = \frac{\alpha}{\Lambda}\frac{\psi_p + \psi_\phi/\gamma}{\psi_p + \psi_\phi} - \frac{Q_p}{\psi_p + \psi_\phi}$$

For unpatented products,

$$\mathcal{S}_u = \int_0^{T_u}\int_0^{\infty}(\alpha - \Lambda c)\omega_u(c, q^0)q^0 dq^0 dc.$$

The first-order condition with respect to $T_u$ is given by

$$\frac{\partial \mathcal{S}_u}{\partial T_u} = \int_0^{\infty}(\alpha - \Lambda T_u^*)\omega_u(T_u^*, q^0)q^0 dq^0 = (\alpha - \Lambda T_u^*)\frac{\partial Q_u}{\partial T_u} = 0,$$

implying that

$$T_u^* = \frac{\alpha}{\Lambda}.$$

■

As we can see both optimal thresholds decrease after accounting for the excess burden of taxation. In essence, allowing for the excess burden of taxation is equivalent to having a smaller willingness to pay.

# References

# References

Abbott, T. A., Vernon, J. A., 2007. The cost of us pharmaceutical price regulation: a financial simulation model of r&d decisions. Managerial and Decision Economics 28 (4-5), 293–306.

Birch, S., Gafni, A., 1992. Cost effectiveness/utility analyses: Do current decision rules lead us to where we want to be? Journal of health economics 11 (3), 279–296.

Claxton, K., Martin, S., Soares, M., Rice, N., Spackman, E., Hinde, S., Devlin, N., Smith, P. C., Sculpher, M., 2015. Methods for the estimation of the national institute for health and care excellence cost-effectiveness threshold. Health technology assessment, 1–542.

Culyer, A., McCabe, C., Briggs, A., Claxton, K., Buxton, M., Akehurst, R., Sculpher, M., Brazier, J., 2007. Searching for a threshold, not setting one: the role of the national institute for health and clinical excellence. Journal of health services research & policy 12 (1), 56–58.

Culyer, A. J., 2016. Cost-effectiveness thresholds in health care: a bookshelf guide to their meaning and use. Health Economics, Policy and Law, 1–18.

Danzon, P., Towse, A., Mestre-Ferrandiz, J., 2015. Value-based differential pricing: Efficient prices for drugs in a global context. Health economics 24 (3), 294–301.

Danzon, P. M., Towse, A., Mulcahy, A. W., 2011. Setting cost-effectiveness thresholds as a means to achieve appropriate drug prices in rich and poor countries. Health Affairs 30 (8), 1529–1538.

Devlin, N., Parkin, D., 2004. Does nice have a cost-effectiveness threshold and what other factors influence its decisions? a binary choice analysis. Health economics 13 (5), 437–452.

Doubilet, P., Weinstein, M. C., McNeil, B. J., 1986. Use and misuse of the term "cost effective" in medicine.

Dubois, P., De Mouzon, O., Scott-Morton, F., Seabright, P., 2015. Market size and pharmaceutical innovation. The RAND Journal of Economics 46 (4), 844–871.

Eckermann, S., Pekarsky, B., 2014. Can the real opportunity cost stand up: displaced services, the straw man outside the room. Pharmacoeconomics 32 (4), 319–325.

Ellison, S. F., Snyder, C. M., 2010. Countervailing power in wholesale pharmaceuticals. The Journal of Industrial Economics 58 (1), 32–53.

Garber, A. M., Phelps, C. E., 1997. Economic foundations of cost-effectiveness analysis. Journal of health economics 16 (1), 1–31.

George, B., Harris, A., Mitchell, A., 2001. Cost-effectiveness analysis and the consistency of decision making. Pharmacoeconomics 19 (11), 1103–1109.

Hemphill, C. S., Sampat, B. N., 2011. When do generics challenge drug patents? Journal of Empirical Legal Studies 8 (4), 613–649.

Marseille, E., Larson, B., Kazi, D. S., Kahn, J. G., Rosen, S., 2015. Thresholds for the cost–effectiveness of interventions: alternative approaches. Bulletin of the World Health Organization 93 (2), 118–124.

McCabe, C., Claxton, K., Culyer, A. J., 2008. The nice cost-effectiveness threshold. Pharmacoeconomics 26 (9), 733–744.

Newall, A., Jit, M., Hutubessy, R., 2014. Are current cost-effectiveness thresholds for low-and middle-income countries useful? examples from the world of vaccines. Pharmacoeconomics 32 (6), 525–531.

O'Mahony, J. F., Coughlan, D., 2016. The irish cost-effectiveness threshold: does it support rational rationing or might it lead to unintended harm to ireland's health system? PharmacoEconomics 34 (1), 5–11.

Towse, A., Pritchard, C., Devlin, N. J., 2002. Cost-effectiveness thresholds: economic and ethical issues. King's Fund.

USTR, 2018. 2018 special 301 report.

Weinstein, M., Zeckhauser, R., 1973. Critical ratios and efficient allocation. Journal of Public Economics 2 (2), 147–157.