The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

Working Papers

www.cesifo.org/wp

Health Insurance and Diversity of Treatment: A Policy Mix Perspective

David Bardey
Bruno Jullien
Jean-Marie Lozachmeur

CESIFO WORKING PAPER NO. 4111
CATEGORY 1: PUBLIC FINANCE
FEBRUARY 2013

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 4111*

Health Insurance and Diversity of Treatment: A Policy Mix Perspective

Abstract

We determine the optimal health policy mix when the average utility of patients increases with the supply of drugs available in a therapeutic class. Health risk coverage rely on two instruments, copayment and reference pricing, that affect the supported risk composed by health expenses and diversity of treatment. For a fixed supply of drugs, the reference pricing policy aims at minimizing expenses in which case, the equilibrium price of drugs is independent of the copayment rate. However, with endogenous supply of drugs, diversity of treatment may substitute for insurance so that the reference pricing may depart from maximal cost-containment to promote entry. We then analyse the determinants of the optimal policy. While an increase in risk aversion or in the side effect loss increases diversity and decreases the copayment rate, an increase in entry cost both decreases diversity and the copayment rate.

JEL-Code: I180, L110, L150, L510.

David Bardey
University of Los Andes – CEDE
Calle 19A
No. 1-37 Este Bloque W – Piso 9
Colombia – Bogotá
david.bardey@gmail.com

Bruno Jullien
Toulouse School of Economics
(IDEI & GREMAQ-CNRS)
Manufacture des Tabacs,
21 Allée de Brienne
France – 31015 Toulouse
bruno.jullien@tse-fr.eu

Jean-Marie Lozachmeur
Toulouse School of Economics
(IDEI & GREMAQ-CNRS)
Manufacture des Tabacs,
21 Allée de Brienne
France – 31015 Toulouse
jean-marie.lozachmeur@univ-tlse1.fr

January 26, 2013
We thank Kurt Brekke, Patricia Danzon, Marisa Miraldo and Luigi Siciliani for their excellent remarks on previous versions of this paper.

1 Introduction

Pharmaceutical expenditures have dramatically increased in most developed countries during the two last decades. Between 2000 to 2009, the annual growth rate of real pharmaceutical per capita spending has been 3.5% on average in OECD's countries. Pharmaceutical expenditures represent on average 1.5% of the GDP in OCDE's countries in 2010 (see OECD, 2011). This trend is mainly due to an explosion of expenditure in *R&D* and its risen cost (see Pammolli and Riccaboni, 2004). Increased expenditures raise serious concerns for both public and private health systems. As a result, regulators rely on various mechanisms to contain costs such as copayments, price control, rationing or dereferencing of some treatments. One controversial issue has long been whether costly incremental innovation leading to "me-too" or "follow-on" drugs (as opposed to major innovation leading to breakthrough products) has some value for the society as a whole. As argued by Wertheimer *et al.* (2001), me-too drugs have several advantages among which "differing dose delivery systems and dosage forms that enable extended uses with a variety of patient population; availability of choice when patient response to and tolerance of a particular agent is subject to great variation; the ability to tailor therapy to the needs and preferences of patients (pp 78)". In line with these comments, Di Masi and Paquette (2004) show that, among 72 therapeutic classes, one third of "me-too" drugs received a priority rating from the US FDA. Moreover, 57% of these classes include a "me-too" drug that received such a priority rating. As such, diversity of treatment itself has an insurance role in that it increases the probability that a patient finds the best (or the least worst-off) treatment.

Faced to budgetary concerns, the key question that a health system faces is thus not only how much financial insurance to provide but also what level of diversity of treatments to offer. Our goal is to analyze this issue in a context where health services are supplied in an imperfect competitive market, as it is the case for drugs' markets. In order to address the issue of the diversity of treatments offered to policyholders, we focus on two policy instruments that are used to maximize individuals' expected utility. The first one is a standard linear copayment rate while the second one is a therapeutic

reference pricing regulation - an instrument that has been adopted by several countries over the last years, including Germany, The Netherlands, Canada or Australia.

While modalities vary,[1] internal or therapeutic reference pricing (as opposed to external reference pricing) consists in determining a reference price as a weighted sum of drugs' prices adopted in the same therapeutic class. If the price of a drug is higher than this reference price, patients pay the full difference between the price of the drug and the reference price along with the insurance copayment on the reference price. This regulatory scheme is thus often perceived as a complement to copayment rates in order to encourage patients in consuming low price drugs. In a static view, where no innovation or drugs entries are taken into account, reference pricing mechanically generates a fall in drugs' reimbursements and prices. However, in the long run, where innovation by pharmaceutical firms is driven by the prospect of future profits, the welfare impact of such a regulation is far from being obvious.

We build a model where there are several pharmaceutical firms selling horizontally differentiated drugs that belong to the same therapeutic class. All the patients value differently the drugs because of their different side effects.[2] Risk averse consumers benefit from an insurance plan consisting in a premium and a linear copayment rate subject to internal reference pricing. In line with the German's modality, we consider a reference price that is defined as a linear combination of extreme value prices in the market. By choosing the weight attached to the lowest price in the reference pricing formula, the regulator modulates the competitive pressure on the equilibrium price of drugs and thus on total health expenses.[3] Our results show that in the short run, the copayment rate is chosen independently of the reference price and the reference pricing scheme aims at minimizing the price of drugs. Indeed, as long as the number of drugs in the therapeutic class is fixed, the health insurer is only willing to lower drugs prices since he cannot improve the drugs' diversity. However, in the long run, there may

[1] See Lopez-Casanovas and Puig-Junoy (2000), Danzon (2001) and Danzon and Ketcham (2004) for more details on reference pricing and its applications.

[2] Side effects accomodate adverse drug reactions, drug-drug interactions, dosing schedules as well as delivery systems.

[3] While the details of the policy depend on the formulation, our key results apply more generally provided that the policy mix allows to control entry prices as well as copayments.

be room for a more lenient reference pricing policy accommodating some increase in health expenses in order to improve the diversity of treatments. The desirable level of the reference price is the result of a trade-off between a diversity effect - a new drug decreases average side effects - and the cost generated by drugs' entries. We particularly emphasize the role of policyholders' risk aversion by showing that the higher is the level of risk aversion, the more likely the regulator chooses a more lenient reference pricing policy. While an increase in risk aversion leads to a lower copayment rate and a higher diversity of treatments, an increase in innovation costs implies a lower copayment and a lower diversity of treatments.

While mainly normative, our article borrows from the positive literature dealing with drugs' price regulation. We are close to Brekke *et al.* (2007) who develop a general set up containing horizontal and vertical differentiations in drugs' market. They consider two reference pricing modalities, namely internal and external reference pricing as well as price cap regulation, both regulatory schemes being associated to an exogenous copayment rate. Concerning the internal reference pricing, they consider laboratories' best reply function and they provide comparative statics on the equilibrium allocation. They show that an increase in the weight attached to the lowest price leads to a reduction in prices and increase the market share of the cheapest drug (the generic drug in their context). In a model that distinguishes breakthrough and follow-on drugs, Bardey *et al.* (2010) evaluate the long run impact of reference pricing on pharmaceutical innovation, delays of introduction, patients' health and expenditures. They show that reference pricing regulation allows to lower prices and therefore delays the entry of pioneer drugs and me-toos. Nevertheless, as me-toos are more delayed, it may favor costly breakthroughs and may increase health expenditures in the long run.[4] Miraldo (2007) compares the equilibrium allocations in an environment of horizontal and vertical differentiations under two alternative regimes: copayment and reference pricing schemes.

[4]This effect is also mentioned in Pammolli and Riccaboni (2004) who explore the interplay between technological advances and cost containment policy. They conclude that reference pricing regulation reduces the rents coming from horizontal differentiation and favors arrivals of major breakthroughs. This result is empirically corroborated by Kyle (2007). She explores the empirical puzzle of the pharmaceuticals entry patterns and finds that laboratories which invest in R&D in countries that use price control reach fewer markets with longer delays.

In contrast, our model is more narrow on the reference pricing modality as we focus our attention on internal reference pricing but both the copayment rate and the drug's price regulation are optimally determined.

On the normative side, very few papers study health insurance in the context of pharmaceutical markets. Lakadawalla and Sood (2009) show that encouraging health insurance can be welfare improving as it lowers static deadweight loss (*i.e.* implies more efficient utilization) without altering incentives for innovation. They also show in an older paper (2006) that a competitive health insurance market can combine static and dynamic efficiency in the drugs' market: a premium-financed fixed fee is offered to drugs' monopolists which ensures second best utilization and extracts the full surplus so that incentives for innovation are optimal. Their model however does not allow to treat the optimal diversity of treatment since individuals can only benefit from at most one innovation.

The next section presents the set up. Section 3 is devoted to the short-term analysis assuming a fixed number of drugs. Section 4 characterizes the optimal regulation in the long run, with an endogenous number of drugs. Section 5 discusses some possible extensions. Lastly, section 6 concludes.

2 The Set Up

Consider a collection of policyholders $J \equiv (0,1)$ who can be ill with probability π and healthy with probability $1-\pi$. We normalize the size of the population to 1. In case of illness, each patient chooses a drug among N (≥ 2) treatments, denoted by a subscript $i \in I \equiv (1,N)$. We refer to N as the *diversity of treatment*. When choosing a drug i in case of illness, a patient's net income is $w_s^i = w - \rho - p_i$, where w is an exogenous income, ρ denotes the premium paid to the health insurer and p_i is the out-of-pocket price paid for consuming drug i. When healthy, the net income is $w_h = w - \rho$. In the state of illness, treatments are horizontally differentiated: when consuming drug i, a policyholder $j \in J$ is affected by a side effect which depends on an individual's stochastic variable $\tilde{x}_j^i \in X$, observed by the policyholders before consumption takes

place.[5] The shocks $\tilde{x}_j^i$ are identically and independently distributed across policyholders and drugs over $\mathbb{R}_+$ and follow an exponential distribution with a mean normalized to 1, i.e. with a distribution function $F(x) = 1 - \exp(-x)$. This assumption implies that the mass of individuals with large adverse effects is always lower than the mass of individuals with small adverse effects.[6]

The policyholder j's expected utility when consuming drug i is:

$$U = \pi u\left(w_s^i - y - \theta\tilde{x}_j^i\right) + (1-\pi)u(w_h),$$

where $y \geq 0$ is a monetary equivalent fixed effect of illness. θ is the differentiation parameter so that $\theta\tilde{x}_j^i$ represents the monetary equivalent loss incurred by a patient j consuming drug i. Throughout the paper, we use an exponential VNM utility function of the CARA form:[7]

$$u(\omega) = -\exp(-\sigma\omega),\ k = h, s,$$

where σ represents the absolute risk aversion parameter.

At a symmetric equilibrium where $p_i = p$ for every $i \in (1, N)$, patients minimize the loss so that the aggregate expected utility function writes:

$$EU = \pi E\left[u\left(w_s - y - \theta \min_{i\in I}\tilde{x}_j^i\right)\right] + (1-\pi)u(w_h), \tag{1}$$

where E is the expectation operator over $J \otimes X^N$.

Our model can be interpreted in two ways. Either individuals know their type j before the risk occurrence and the aggregate utility represents a utilitarian objective or individuals do not know their type before the risk occurrence and the aggregate utility represents the representative individual's expected utility.

For a CARA utility function, equation (1) yields:

$$EU = \pi\frac{\eta}{Q}u(w_s) + (1-\pi)u(w_h), \tag{2}$$

[5] We implicitly assume that the doctor perfectly observes the state of the patient and delivers the good information on possible adverse effects of any drug to ensure the better match between each drug with each patient.

[6] We study a more general formulation of side effects' distribution in section 5.2.

[7] This assumption is made for analytical tractability. It allows us in particular to perform simple comparative statics with respect to the risk aversion parameter.

where

$$\eta = \exp(\sigma y) \geq 1 \text{ and } Q = \frac{1}{E\left\{\exp\left(\sigma\theta \min_{i\in I} \tilde{x}_j^i\right)\right\}} < 1.$$

The parameter η measures the disutility that a patient incurs if he obtains the best treatment and supports no side effect. It is referred as the *severity of the pathology*. The parameter Q is a measure of the average side effects of available treatments which impacts the average utility when ill. Given that u is negative (due to the CARA specification), the expected utility increases when Q increases. To ensure finiteness of the expected utility for all N we suppose that the following assumption holds:[8]

ASSUMPTION 1 The market is such that $N > \sigma\theta$.

When this assumption is fulfilled, one has:[9]

$$Q(N) = 1 - \frac{\sigma\theta}{N}. \tag{3}$$

Our model thus relates the average side effects to the diversity of available drugs in the therapeutic class. We thus refer $Q(N)$ as the *welfare index of diversity* (*WID* hereafter). *Ceteris paribus*, the higher is the number of drugs available in the therapeutic class, the higher is the perceived utility of the health service as patients benefits from a decrease in the average side effect. On the contrary, the negative consequences of side effects on patients' welfare both increase in the differentiation and in the risk aversion parameters.

3 Insurance Policy and the Short-run Equilibrium of the Pharmaceutical Market

In this section, we first determine the market equilibrium in the therapeutic class and then analyze the short-run insurance trade-off when diversity is given. We endogenize the diversity of treatment in the next section.

[8] In the long-run analysis, the assumption constrains an endogenous variable N. However, assumption 3 below ensures that it holds in any equilibrium.

[9] Since $\tilde{x} = \min_{i\in I} \tilde{x}_j^i$ is exponentially distributed with mean $1/N$ (density $N\exp(-Nx)$) $1/Q(N) = N\int_0^\infty \exp((\sigma\theta - N)x)$.

3.1 Insurance and Drugs Market Equilibrium

We consider one specific therapeutic class in which pharmaceutical firms each own one drug and compete in the market for drugs by setting prices. While prices are free, the equilibrium is shaped by the insurance policy which is based on two instruments: a copayment rate $a \in [0,1]$ and reference pricing.[10] We model reference pricing as a rule that associates to any vector of manufacturers' prices $(P_1, ...P_N)$ a reference price $R(P_1, ..., P_N)$. A patient buying drug i is thus reimbursed $(1-a)\min\{P_i, R\}$, that is a share $1-a$ of his expense up to R, but obtains no reimbursement for the excess payment above the reference price.

Several formula for the reference price can be used. We follow the German internal reference pricing regulation by assuming that the reference price is a weighted sum of the minimal and the maximal prices of the drugs in the market. Formally, the reference price is given by:

$$R = r \inf_i \{P_i\} + (1-r) \max_i \{P_i\}, \tag{4}$$

where $r \in [0,1]$ is the policy variable, referred to as the "strength" of the drug price regulation.[11] When $r = 1$, there is "maximal" reference pricing. This is the most stringent policy from the perspective of pharmaceutical firms. On the contrary, when $r = 0$, there is no reference pricing. In such a case, all the treatments are reimbursed at a rate $1 - a$.

In the main body of the paper, we assume that an equilibrium exists where the market is fully covered. Roughly speaking, prices and adverse effects are supposed to be small enough in comparison with treatments benefits so that all policyholders choose a drug in the therapeutic class when ill.[12] The case in which patients can forgo any treatment is analyzed in section 5.1. The demand for one drug is thus the mass of individuals who prefer this drug to any other. We then focus on symmetric equilibria where firms set the same prices.

[10] Alternative mechanisms to reference pricing could be price-cap or direct price negotiation with producers.

[11] For instance, in Germany r is set to 2/3.

[12] This assumption is quite realistic in countries characterized by high levels of coverage, *i.e.* low levels of out-of-pockets.

Consider any drug $i \in (1, N)$ and suppose that the net out-of-pocket paid by patients for all other drugs is p. Our assumptions on the utility functions (in particular, the fact that the utility of ill patients are symmetric and quasi-linear in prices) imply that the demand D_i, when the out-of-pocket price paid for consuming drug i is p_i, can be expressed as follows at any symmetric equilibrium:

$$D_i = D\left(p_i - p, N\right),$$

where D is a function independent of i with $\partial D/\partial p_i < 0$ for any p_i, p and N. At a symmetric equilibrium, each pharmaceutical firm's market share is $D\left(0, N\right) = \pi/N$.[13]

By definition, when all other firms charge P and firm i charges P_i, the out-of-pocket price paid by patients is defined by:

$$p_i = P_i - (1-a)\min\left[P_i, R\right]$$

so that,

$$p_i - p = P_i - P - (1-a)\left(\min\left[P_i, R\right] - \min\left[P, R\right]\right).$$

Using the definition of R in equation (4), this yields:

$$p_i - p = \left[a + (1-a)r\right]\left(P_i - P\right),$$

where $\alpha = a + (1-a)r$ measures the distortion introduced by the insurance system on the market for drugs. We refer to α as the "*pass-through rate*" resulting from the joint regulation of insurance coverage and reference pricing. This rate determines the fraction of the price which is actually paid by individuals and is increasing (resp. decreasing) in the copayment rate (resp. the reference pricing strength). Note that our reference pricing specification does not generate any kink in the demand function as opposed to Danzon (2001). This is due to the fact that we focus on symmetric equilibria in fully covered markets. In section 5.1, which considers non-fully covered markets, such a kink emerges.

[13] More formally $\frac{1}{\pi} D\left(p_i - p, N\right)$ is the probability that the "total loss" $\theta \tilde{x}_j^i + p_i$ be less than $\theta \min_{l \neq i} \tilde{x}_j^l + p$.

For-profit pharmaceutical firms compete in prices and have identical marginal costs that are normalized to 0. Therefore, we have:

$$P_i \in \arg\max P_i D\left(\alpha\left(P_i - P\right), N\right).$$

As the symmetric equilibrium yields $D_i = \pi/N$, we obtain:

Proposition 1 *The drug price in a symmetric equilibrium is*

$$P = \frac{\varphi(N)}{\alpha},$$

where $\varphi(N) = \theta/(N-1)$ *is the inverse semi-elasticity of the residual demand evaluated at* $p_i = p$ *(i.e.* $-D(0,N)/(\partial D(0,N)/\partial p_i)$*).*

Proof. See Appendix A. ■

The resulting price equilibrium is identical to the one in a standard general model of horizontal differentiation in which the differentiation parameter would be θ/α. Here, the pass-through rate α increases in the strength of reference pricing and in the copayment rate. In other words, increasing the insurance coverage lowers the competition intensity. This feature of the model illustrates what Danzon (2012) refers to as "supplier moral hazard". Equivalently, relaxing the strength of reference price regulation lowers the competition intensity.

At a symmetric equilibrium, the pharmaceutical firms' profit is then:

$$\Pi(N) = \frac{\pi\varphi(N)}{\alpha N}. \tag{5}$$

The closed form solution is useful to point out in a more obvious way the different effects at work. *Ceteris paribus*, the differentiation parameter θ increases the drugs' price and the profit obtained by laboratories. However, the overall differentiation $\varphi(N)$ decreases in the diversity of available treatments in the therapeutic class.

3.2 Fixed Drugs' Diversity

We determine here the optimal policy in the short-run, considering the market structure as given. Since the diversity N of available drugs in the therapeutic class is fixed, we drop the argument N in the functions of this section.

The social objective is to maximize the policyholder's expected utility as defined in (2).[14] Note that the optimum can be decentralized in two institutional *scenarii*. The first one corresponds to a regulator who designs simultaneously the reference price and the copayment rate. The second one is a two-stage game where the insurer (either public or private) first chooses the copayment rate and in a second stage, the regulator fixes the reference price. This first stage may correspond to the result of a perfect competition on the health insurance market with a positive loading factor or to the social insurance policy subject to a budget constraint in the presence of distortionary taxes.[15]

The resource constraint imposes $\rho = \pi(1-a)(1+\lambda)P$ where λ is to be interpreted as a loading factor if the insurer is private or as a shadow cost of public fund if the insurer is public. For a fixed diversity of treatments, the optimal policy is the solution of:

$$\max_{0\leq a,r\leq 1} EU = (1-\pi)u(w-\rho) + \pi\frac{\eta}{Q}u\left(w-\rho-aP\right)$$
$$s.t.\ \rho = \pi(1+\lambda)(1-a)P, \quad P = \frac{\varphi}{a+(1-a)r}.$$

For any $r < 1$, a decrease in a exacerbates the extent of supplier moral hazard by increasing the price paid by patients. Nevertheless, this problem can be dealt with by adjusting the strength of reference pricing: for a fixed supply of drugs, reducing the price P can only benefit consumers because the *WID* is fixed and the price is always above marginal cost. As increasing the weight r in the reference price decreases the equilibrium price, it follows that it is always optimal to choose maximal reference pricing $r = 1$. The price is then at φ, the level that would prevail without any insurance scheme. In other words, full reference pricing allows to cancel out the supplier moral hazard effect.

Given that the price is fixed at $P = \varphi$, the problem reduces to a standard insurance problem. The slope of the objective function with respect to a is:

$$\frac{\partial EU}{\partial a} = \pi EU'(1+\lambda)P - \frac{\pi\eta}{Q}u'\left(w_s\right)P, \tag{6}$$

[14] Laboratories'profits are not taken into account in the objective. This assumption is not a problem as long as insurers are private. This also helps comparing our results with the ones in the following section where profits will be nill.

[15] The second scenario is equivalent to the first one as long as insurers offer health insurance contracts before patients know their type.

where $EU' = (1-\pi)\,u'(w_h) + (\pi\eta/Q)\,u'(w_s)$ is the policyholders' expected marginal utility of income. Rearranging (6) for a solution with positive insurance coverage ($a < 1$) gives:

$$EU' \leq \frac{\eta u'(w_s)}{Q\,(1+\lambda)}\ (=0 \text{ if } a>0). \tag{7}$$

The optimal policy is then characterized in the following proposition:

Proposition 2 *For a fixed diversity of treatments, the optimal strength of reference pricing is $r=1$ and the copayment rate is non-decreasing in λ.*
If $\lambda \leq \lambda_0 = (\eta-1)(1-\pi)/(1-\pi+\pi\eta)$, full insurance is optimal for all N while if $\lambda \geq (1-\pi)/\pi$, no insurance is optimal. If $\lambda \in (\lambda_0, (1-\pi)/\pi)$, there exists $\bar{N}$ such that no insurance is optimal if $N \geq \bar{N}$, while the optimal copayment rate is implicitly determined by (7) if $N < \bar{N}$.

Proof. See Appendix B. ■

For λ small, the increase of the premium is not sufficient to require a positive copayment rate. This is due to the fact that the side effects renders the marginal utility of income dependent of the state, and in particular larger when being sick.[16] For λ larger than λ_0, the risk when N is large is not sufficient to grant insurance. Notice that when N is large, the price is close to zero and the average side effect is also close to zero, but there may be some residual risk due to the direct consequence of illness on utility (as measured by η). Thus when η/Q is small, the optimal policy is not to provide any insurance at all. Partial insurance occurs when the supply of drugs is insufficient and in this case, the optimal copayment rate, determined by condition (7), is the result of a trade-off between risk pooling (*ex-ante* efficiency) and an increase in the size of the risk (*ex-post* efficiency).[17]

The proposition is illustrated on Figure 1 which represents the choice in the space $(Q(N), \lambda)$.[18] There are two curves such that full insurance is optimal below the lower

[16] In this case a policy that would provide over-insurance would be optimal but this option is ruled out.

[17] See Geoffard (2006) for a complete analysis dealing with this trade-off in a context where *ex-post* moral hazard is due to a quantity effect.

[18] It is drawn for $\sigma = 1$, $\theta = 1.5$, $\pi = 1/3$ and $\delta = 1.2$. Thus the diversity index is $Q = 0.5$ for $N = 3$.

curve while no insurance is optimal above the upper one. The two curves coincide at zero and 1, so that there is no need for insurance for $\lambda \geq (1-\pi)/\pi$ and there is full insurance for $\lambda \leq \lambda_0$.

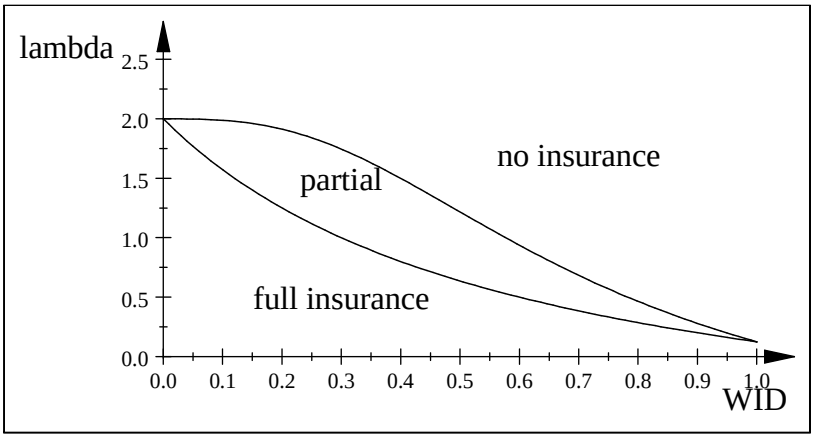

Figure 1: Optimal insurance scheme with fixed supply.

4 Endogenous Drugs' Diversity

In this section, we endogenize the market structure in the therapeutic class and we focus on the free-entry equilibrium that occurs in the long run. Laboratories enter into the market as long as the profits earned in the drug market outweigh a uniform sunk cost k generated by the $R\&D$ activities.[19] For conciseness we ignore the fact that N is an integer and treat N as a continuous variable (see footnote 21 for more details).

Equalizing the sunk cost to equilibrium profits given by equation (5), the diversity for a free-entry equilibrium is implicitly given by (up to the integer):

$$k = \frac{\pi\varphi(N)}{(a+(1-a)r)N}. \tag{8}$$

[19] This assumption fits well the reality as incremental innovation has often been generated by unintentional research aimed towards producing breakthrough products (see Wertheimer *et al.*, 2001). One can however argue that, starting from a certain number of drugs existing in a therapeutic class, the fixed cost of innovation is decreasing since mee-too drugs enter the market with less clinical-trial experiments (see Frothingham, 2004). A decreasing cost function would ultimately reinforce our result stating that less than maximal reference pricing would be optimal.

Accounting for endogeneity of supply changes the nature of the analysis compared to the case with fixed supply. Indeed, when N is fixed, the only scope for price policy is cost containment. Endogenizing the supply now creates a linkage between total expenses and diversity. The free-entry condition implies an increasing relationship between diversity of treatment and expenses:

$$\pi P = kN.$$

While the health policy may raise diversity, it has to account for this cost. This suggests that we view the choice of health policy as the joint choice of insurance and diversity of treatment with a cost function that is determined by the free-entry condition.

Given that $a+(1-a)r \leq 1$, the minimal number of laboratories entering the market is $\underline{N}$ given implicitly by the free entry condition $k = \pi\varphi(N)/N$ or:

$$\underline{N}(\underline{N}-1) \simeq \frac{\pi\theta}{k}. \tag{9}$$

Remind that Q is defined only for $N > \sigma\theta$. To ensure this condition we assume (defining $\bar{N} = +\infty$ if $\lambda \leq \lambda_0$):

ASSUMPTION 2 $\underline{N} \geq 2$ and $\sigma\theta < \underline{N} < \bar{N}$.

Assumption 2 guarantees that there are at least two firms in the market and the *WID* is well defined ($Q > 0$). Moreover, the condition $\underline{N} < \bar{N}$ ensures that there will be some insurance in equilibrium, *i.e.* the copayment will be less than 1.[20]

In order to provide more intuitions about optimal regulation, we proceed in two steps. First, we assume that the copayment rate a is exogenously fixed. In a second part, a is endogenously determined.

4.1 Reference Pricing and Diversity for Fixed Copayment

Even though we consider this step in order to decompose more easily the different effects at work, it is worth stressing that such a situation may occur when the copayment rate is set by another regulator or at a high level of aggregation of pathologies/therapeutic

[20] To see that notice that $N = \underline{N}$ if $a = 1$. The condition implies that $a < 1$ is optimal for $N = \underline{N}$, and therefore it is impossible that $a = 1$ is an optimal choice.

class. For a given copayment rate, the strength of reference pricing determines the equilibrium price for any level of entry. Increasing r results in lower equilibrium prices and consequently in less entry. Hence, the optimal policy reflects a trade-off between the total cost of treatment and the *WID*.

Notice that the free-entry equilibrium (8) induces a decreasing relationship between r and N. Hence, for a given a, choosing r is equivalent to choosing the diversity N under the relations:

$$0 \leq r = \frac{\pi\varphi\left(N\right)/kN - a}{1-a} \leq 1 \text{ and } \pi P = kN.$$

Thus, the problem reduces to the choice of diversity within a feasible range and the optimal policy is the solution of the following program:[21]

$$\begin{aligned} &\max_N EU = (1-\pi)u\left(w_h\right) + \frac{\pi\eta}{Q}u\left(w_h - a\frac{kN}{\pi}\right), \\ &s.t. \quad w_h = w - (1+\lambda)(1-a)kN, \\ &\qquad N \geq \underline{N}, \\ &\qquad akN - \pi\varphi\left(N\right) \leq 0. \end{aligned}$$

When $r = 1$, the level of entry is always equal to $\underline{N}$. Hence, at any copayment level, it is possible to implement the minimal *WID* at the minimal cost. One may raise diversity above this level by reducing the strength r of reference pricing. However, this would imply some more entry costs. The first question to answer is then whether it is optimal to minimize cost and to rely only on insurance, or to accommodate some cost increase. The gain of raising diversity is captured by the slope of EU with respect to N, which is equal to:

$$\frac{\partial EU}{\partial N} = -\frac{\pi\eta}{Q^2}\frac{\partial Q}{\partial N}u(w_s) - k\left(1+\lambda\right)\left(1-a\right)EU' - \frac{\eta}{Q}aku'\left(w_s\right),$$

where EU' is defined as in equation (6). Using equation (7) and $u' = -\sigma u$, this slope can be rewritten as:

$$\frac{\partial EU}{\partial N} = \left[\frac{\pi}{\sigma Q}\frac{\partial Q}{\partial N} - k\right]\frac{\eta}{Q}u'\left(w_s\right) - \left(\frac{1-a}{N}\right)\left.\frac{\partial EU}{\partial a}\right|_N, \tag{10}$$

[21] When N is treated as a discrete variable, several values of r generate the same N. Clearly it is optimal to choose the maximal strength compatible with N, hence r is such that $P = kN/\pi$. The problem is thus identical but restricted to integer values of N. The comparative statics are the same.

where $\frac{\partial EU}{\partial a}\big|_N$ is given by the LHS of equation (6). The first term in the RHS of equation (10) captures the trade-off between a higher *WID* (associated with more entry) and the cost of entry. The last term captures the extent to which reference pricing can correct for insufficient or excessive copayment to provide some insurance. Indeed, when there is insufficient coverage (*i.e.* $\partial EU/\partial a|_N < 0$), additional entry acts a substitute to insurance since it reduces the difference of marginal utilities between healthy and ill individuals. Thus, entry is more valuable when the copayment is too large. Whether it is optimal to reduce the strength of reference pricing to induce some entry then depends on the level of the copayment rate.

Proposition 3 *Suppose the copayment is fixed and let r_a^* denote the optimal strength of reference pricing.*
If $\sigma\theta \geq 1$, then $r_a^* < 1$ and the optimal diversity is above $\underline{N}$ for $a \in (\underline{a}, 1)$ where $\underline{a} < 1$. Moreover $r_a^* < 1$ for all a if full insurance is optimal in the short-run at $N = \underline{N}$.
If $\sigma\theta < 1$, either diversity is minimal for all a, or $r_a^* < 1$ for $a \in (\underline{a}, \bar{a})$ with $0 \leq \underline{a} < \bar{a} < 1$.

Proof. See Appendix C. ■

When risk aversion and the differentiation parameter are not too small, it is optimal to induce entry by relaxing the intensity of competition at least for large levels of coverage. This conclusion contrasts with the welfare analysis of entry in standard models of horizontal differentiation, which conclude that there is always excessive diversity at an unregulated free entry equilibrium (*e.g.* see Anderson *et al.*, 1992). The key difference between our set up and these models is that diversity has an additional benefit in the case of pharmaceutical drugs: improving the welfare of ill individuals improves the overall insurance performance. As a result, entry is desirable if insurance is not too excessive. This is the case in particular when full insurance is optimal as alternative means of providing insurance help reducing the gap between marginal utilities. On the contrary, for low levels of differentiation and risk aversion, the insurance benefit is not so important and the traditional result of excessive entry may apply. In this case,

relaxing the strength of reference pricing may be desirable only for intermediate levels of coverage or never.

We show in appendix D that the objective is quasi-concave which allows to perform simple comparative statics.

Proposition 4 *For a fixed copayment, the optimal diversity:*
i) is non-decreasing in the copayment if there is excessive coverage, i.e. if $\partial EU/\partial a \geq 0$*;*
ii) is non-increasing in λ *and non-decreasing in* σ *and* θ*.*

Proof. See Appendix D. ■

Increasing the copayment has two effects. It reduces the level of insurance and the extent of supplier's moral hazard. Since entry induces some indirect form of insurance through a higher WID, the two effects make entry more attractive if there is not enough insurance. In such a case, the optimal diversity of treatment increases in the copayment. However, if there is excessive insurance, the two effects work in opposite direction and no conclusion can be derived. The other results enunciated in Proposition 4 are rather intuitive and state that the policy becomes more oriented toward cost containment when the cost of public fund increases, risk aversion decreases or the differentiation parameter decreases.

4.2 Optimal Policy Mix

We now consider the complete set of instruments and discuss the optimal choice of coverage and strength of reference pricing. As already mentioned, assumption 2 rules out the possibility of no insurance. Hence, we only need to consider full insurance or partial insurance. Before determining which regime prevails, we discuss the strength of reference pricing in each regime.

Proposition 3 implies that for $\sigma\theta$ larger than 1, a full insurance policy is necessarily associated with some incentive to entry and $r^* < 1$. To see this, notice that if $a^* = 0$ and $r^* = 1$, then $N = \underline{N}$. But this would contradict proposition 3 which implies that $r_0^* < 1$ (as full insurance must be optimal for $a^* = 0$).

Let us now consider the case where partial insurance is optimal, the first-order conditions with respect to N give (using equations (6) and (10) for $\partial EU/\partial a = 0$):

$$EU' = \frac{\eta u'(w_s)}{Q(1+\lambda)}; \tag{11}$$

$$\frac{\pi}{\sigma Q}\frac{\partial Q}{\partial N} \leq k \ (= \text{ if } N > \underline{N}). \tag{12}$$

Equation (11) describes the same trade-off between risk pooling and size of the risk as in the benchmark section, but obviously taken at a different value of N. The second equation captures the trade-off between the social benefits and costs generated by the entry of a new drug, respectively described by the LHS and the RHS of equation (12). The social marginal cost is simply the welfare effect of an additional sunk cost, accounting for the optimal allocation of this cost across health states. The social marginal benefit captures the increase in the *WID* through the marginal reduction of average adverse effects generated by a new drug entry. In the following, we analyze whether the optimal diversity of treatment is interior or not, *i.e.* whether there is less than maximal reference pricing or not.

Remind that the minimal diversity $\underline{N}$ is given implicitly by the free entry condition $k = \pi\varphi(\underline{N})$ and is independent of the level of absolute risk aversion σ. Substituting (12) in (11), a necessary condition for an interior solution for r is

$$\frac{\pi}{\sigma Q(\underline{N})}\frac{\partial Q(\underline{N})}{\partial N} - k > 0. \tag{13}$$

We can now state our main proposition:

Proposition 5 *The optimal copayment a^* is non-decreasing in λ and diversity N^* is non-increasing in λ.*
i) If $\sigma\theta \geq 1$, then $r^* < 1$ and diversity is above $\underline{N}$ for all λ;
ii) If $\sigma\theta < 1$, then $r^* = 1$ when $a^* > 0$.

Proof. See Appendix E. ■

Increasing λ always leads to less coverage and less diversity as both are costly. Still, whenever $\sigma\theta$ is large enough, it is optimal to raise diversity above the minimum by

relaxing the strength of reference pricing and inducing some entry. Conversely, if $\sigma\theta$ is smaller than 1, cost minimization tends to prevail and minimal diversity will be chosen for λ not too small. Therefore, when $\sigma\theta$ is lower than 1, only $\underline{N}$ (and thus $r^* = 1$) is possible when full insurance is optimal (hence λ small).

The proposition also reveals that there exists a threshold $\underline{\lambda}$ such that there is partial insurance if $\lambda > \underline{\lambda}$ and in this case the policy departs from cost minimization when the combined value of risk aversion and disutility from side effects is large enough. When $\lambda < \underline{\lambda}$, the optimal policy coincides with the policy analyzed in the previous section for $a = 0$. On the range $\lambda > \underline{\lambda}$, we can analyze how the optimal level of copayment and the optimal diversity of treatment offered in the therapeutic class vary with preferences and the fixed cost.

Proposition 6 *For $\lambda > \underline{\lambda}$ (partial insurance) and $\sigma\theta > 1$:*
i) An increase in σ or θ increases diversity and decreases the copayment;
ii) An increase in k decreases both diversity and the copayment;
iii) An increase in η decreases the copayment but does not affect diversity.

Proof. See Appendix F. ■

The results first state that more risk aversion (σ) or more risk (θ and η) implies more coverage. This is in line with standard insurance analysis. The result on k follows the same logic: as providing a given level of WID becomes more costly, the treatment expense increases and the optimal policy compensates the increase in risk by a reduction in the copayment rate. The result *i)* shows that insurance and the diversity of treatment are complementary instruments to accommodate an increase in risk aversion or in the heterogeneity of side-effects. However, the result *ii)* states that they evolve as substitute when the cost of entry increases. The last result is interesting as it states that the optimal diversity is independent of the severity of the pathology. What matters is the marginal increase in utility that diversity of treatment induces, and not the absolute value of disutility.

5 Robustness Checks

In this section, we explore two alternative issues related to our set up. First, we develop the model when the drug market is not fully covered. Second, we check whether our main results hold with a more general specification of the distribution of adverse effects.

5.1 Partially Covered Market

This section extends the previous analysis to the case where individuals may forgo any treatment. We are in particular interested in the extent to which this possibility implies more or less diversity at the optimum.

Assume now that there is an outside alternative which is not to consume any drug. Let us denote $\theta_0 \tilde{x}_j^0$ the monetary equivalent loss incurred by an ill individual of type j when not consuming any drug. $\tilde{x}_j^0$ is assumed to be i.i. exponentially distributed along with the $\tilde{x}_j^i$. One can view it as the individual severity of illness. Note that $\theta_0 \to +\infty$ corresponds to the case of fully covered market as studied in the preceding sections.

Define $\tau_0 = \theta/\theta_0$ the measure of the opportunity benefit of forgoing a treatment. The symmetric equilibrium policyholder j's expected utility now becomes:

$$EU = \pi E\left[u\left(w_h - y - \min\left\{\theta \min_{i\in I} \tilde{x}_j^i + p;\ \frac{\theta}{\tau_0}\tilde{x}_0^i\right\}\right)\right] + (1-\pi)u(w_h),$$

where now the agent has access to one more option.

Suppose that all agents face an out-of-pocket price p except for drug $k \in I$. The demand function for drug k is given by the mass of sick individuals who prefer drug k over the others and whose severity of illness justifies a treatment:

$$D_k = \pi \int_0^\infty f(x)\left(1 - F\left(\frac{p_k - p}{\theta} + x\right)\right)^{N-1}\left(1 - F\left(\tau_0 \frac{p_k}{\theta} + \tau_0 x\right)\right) dx.$$

We then have

$$D_k = \frac{\pi}{N + \tau_0} \exp\left(-\frac{(N-1)(p_k - p) + \tau_0 p_k}{\theta}\right).$$

Using the definition of the reference price mechanism and $\alpha = a + r(1-a)$, one thus has:

$$D_k = \frac{\pi}{N + \tau_0} \exp\left(-\frac{(N-1)\alpha(P_k - P) + \tau_0(P_k - (1-a)\min[P_k, R])}{\theta}\right),$$

where R is given by (4). The demand function is thus defined by:

$$\begin{aligned} D_k\left(P_k,P\right) &= \frac{\pi}{N+\tau_0}\exp\left(-\frac{\left(N-1\right)\alpha\left(P_k-P\right)+\tau_0 aP_k}{\theta}\right) \quad if \quad P_k \leq P; \\ D_k\left(P_k,P\right) &= \frac{\pi}{N+\tau_0}\exp\left(-\frac{\left(N-1+\tau_0\right)\alpha\left(P_k-P\right)+\tau_0 aP}{\theta}\right) \quad if \quad P_k \leq P. \end{aligned}$$

As expected, with an elastic aggregate demand for drugs, the reference pricing system with $r > 0$ yields a kink in the residual demand that a pharmaceutical firm faces as pointed out in Danzon (2001). This kink implies that the demand is less elastic for a price P_k below the market price than for a price P_k above. As a result there exist multiple equilibria.

Proposition 7 *When $\tau_0 > 0$, there exists a continuum of symmetric price equilibria $P \in \left(\underline{P},\bar{P}\right)$ where*

$$\underline{P} = \frac{\theta}{\left(N-1+\tau_0\right)\alpha} \text{ and } \bar{P} = \frac{\theta}{\left(N-1\right)\alpha+\tau_0 a}$$

Proof. See appendix G. ■

We can measure the size of the equilibrium set by the following ratio:

$$\frac{\bar{P}-\underline{P}}{\bar{P}} = \frac{\tau_0}{N-1+\tau_0}\frac{r\left(1-a\right)}{a+r\left(1-a\right)}.$$

This is increasing with τ_0 and decreasing with N. Hence, the range increases when the treatments become less useful either because the option of no treatment becomes less detrimental or because the diversity of drugs decreases. The range also increases with the strength of reference pricing r. Figure 2 illustrates the price equilibria in the (r,P) space. Every price between the upper and lower curves are possible price equilibria in the market with an elastic demand.

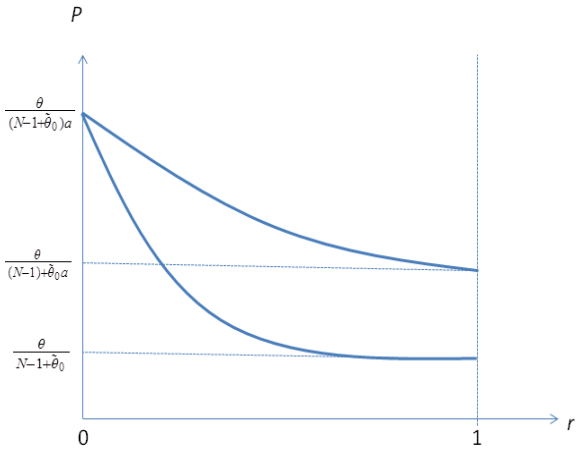

Figure 2: Price equilibria with elastic demand

Let us now turn to the welfare analysis. In this model, we have to consider three situations for individuals. An individual may be healthy, sick with treatment or sick without treatment. This implies that we adjust our notion of index of diversity and that we introduce a *WID* that accounts for the probability of not using a treatment. Considering that policyholders' preferences are represented by a CARA utility function and that assumption 1 holds, one can rewrite the expected utility function as:

$$EU = \pi\eta\left(\frac{1}{Q_1}u\left(w_s\right) + \frac{1}{Q_0}u\left(w_h\right)\right) + \left(1-\pi\right)u\left(w_h\right) \tag{14}$$

where:

$$\frac{1}{Q_1} = \frac{N}{N+\tau_0-\sigma\theta}\exp\left(-\frac{\tau_0 aP}{\theta}\right), \tag{15}$$

$$\frac{1}{Q_0} = \frac{\tau_0}{N+\tau_0-\sigma\theta}\exp\left(\frac{NaP}{\theta}\right). \tag{16}$$

Q_1 is the *WID* accounting both for the probability of consuming a drug and the expected side effect conditional on the fact that individuals consume a drug. This *WID* depends on prices as this affects the decision to consume. In contrast, Q_0 measures the value of

the new option of not consuming drugs when the severity of illness is low. Contrary to Q_1, it is obviously decreasing in τ_0 since it decreases the disutility of not consuming any drug. As in the preceding section, both Q_1 and Q_0 are decreasing (resp. increasing) in the risk aversion parameter (resp. the treatment diversity).

Notice that both Q_1 and Q_0 are functions of $p = aP$. This reflects the effect of the out-of-pocket payment on the probability of using a treatment and on the conditional expected utility. Indeed as p increases, it is less likely that the patient uses a treatment. This in turn also implies that a patient using a treatment supports lower side effects on average.

To conduct our policy analysis we need to make an assumption on the relevant price that emerges given the policy mix (a, r). Given that we are primarily interested by the conditions under which the optimal policy involves $r < 1$ under some insurance $a < 1$, and for conciseness, we develop the analysis only for $a = 0$ and $P = \bar{P}$, the largest possible equilibrium price. Then, we have $\bar{P} = \theta/(N-1)r$ and the average demand for a drug is $D = \pi/(N + \tau_0)$. In the long run, the free entry condition $k = \bar{P}D/N$ yields the diversity as an implicit function r as follows:

$$(N + \tau_0)(N - 1)r = \frac{\pi\theta}{k}. \tag{17}$$

The equilibrium number of firms is thus lower when the opportunity benefit of forgoing a treatment is high. Using this zero profit free entry condition, the actuarial premium must just cover the total cost of entry so that one has $\rho = (1 + \lambda)Nk$. It follows that the socially optimal level of r is obtained by maximizing welfare (14) subject to $w_h = w - (1 + \lambda)kN$ and condition (17).

Corollary 1 *Suppose that $a = 0$, then the optimal diversity N is decreasing with τ_0.*

Proof. See appendix H. ■

As the opportunity cost of forgoing a treatment decreases, entry becomes less desirable. Note that all the comparative statics in proposition 4 still hold.

5.2 Adverse Effects Distribution

In this section, we extend the model to test the robustness of our results to the assumptions on the distribution of side effects across the drugs. In particular, we want to investigate whether there is room for less than maximal reference pricing if the risk aversion parameter σ is set above some level. This conclusion holds provided that the condition (13) is increasing with σ. We maintain the assumption that the shocks $\tilde{x}_j^i$ are i.i. distributed across policyholders and drugs over $(0, \bar{x})$ but allow for an arbitrarily distribution $F(x)$. The demand for a drug i is now given by

$$D_i(p_i - p, N) = \pi \int_0^{\bar{x}} f(x) \left(1 - F\left(\frac{p_i - p}{\theta} + x\right)\right)^{N-1},$$

where p is the out of pocket price of any drugs $k \neq i$. Profit maximization by laboratories yields a symmetric equilibrium price given by:

$$P = \frac{\theta}{\alpha N(N-1)} \int_0^{\bar{x}} f^2(x) (1 - F(x))^{N-2}.$$

The relationship between the price and the policy is thus unchanged. Similarly, the free entry condition is unchanged and reduces to $P = kN/\pi$. The welfare analysis is similar provided the diversity index Q is defined as:

$$Q(N) = \int_0^{\bar{x}} \exp(\sigma\theta x) N (1 - F(x))^{N-1} f(x) dx. \tag{18}$$

As a result, equilibrium conditions (11) and (12) remain valid. Under quasi-concavity, a condition to have less than maximal reference pricing is thus that property (13) holds. The next proposition exhibits conditions for which the function (13) is monotonic in the risk aversion parameter.

Proposition 8 *Assume that $\tilde{x}_j^i$ are identically and independently distributed across policyholders and drugs, with atomless cumulative distribution function F on a support $[0, \bar{x}]$ (where $\bar{x}$ can be infinite) and density f. Assume $E(\exp(\sigma\theta x))$ exists for σ small. Then there exists σ_N (possibly infinite) such that $Q(N)$ is defined for $\sigma \in [0, \sigma_N[$. If $1 - F(x)$ is log-log concave, then $(\pi/\sigma Q(\underline{N}))\, \partial Q(\underline{N})/\partial N$ is increasing to $+\infty$ in σ.*

Proof. See appendix I. ■

The proposition shows that under quite general conditions, the inequality (25) holds for a level of risk aversion that is above some threshold. In this case, we maintain the conclusion that the reference price should not be indexed on the minimum price when agents are sufficiently risk averse.

6 Conclusion

Our analysis sheds light on various mechanisms and underline several results related to drug price regulation, such as reference pricing. First, on the short run, it is always optimal for the health insurer to implement the strongest policy in order to lower health care expenditure. In the long run, the regulator takes into account the drug entry process and the overall diversity of treatment offered to patients. Our model points to the complementarity between copayment policy and reference pricing, or other cost containment policy that work by limiting supply. Indeed, diversity of treatment is an indirect form of insurance that may substitute for or complement monetary insurance. In terms of policy making, our model points out the importance of having good estimations of the impact of prices on the supply and of the social value of drugs diversity (which depends positively on the level of policyholders' risk aversion) before deciding on a reference pricing rule.

Our model could be extended in several ways. It could be interesting to introduce some vertical differentiation between drugs. In this case, following Brekke *et al.* (2007), we could compare more explicitly the two modalities of reference pricing *i.e.* generic *versus* therapeutic. Another extension would be to open the black box of the entry process by modeling completely the innovation process (Aghion and Tirole, 1994).

In this paper, we determine the optimal public regulation in short and long run equilibrium analysis. Nevertheless, as in Bardey *et al.* (2010), it would be useful to complete this analysis by introducing delays in order to take into account the impact of health insurance and drug price regulation on drugs introduction delay. Indeed, these delays of introduction may have a huge impact in terms of welfare.

References

[1] Aghion P and Tirole J, 1994, Opening the Black Box of Innovation, *European, Economic Review.*

[2] Anderson S., de Palma A. J. Thisse, 1992, Discrete Choice Theory of Product Differentiation, MIT Press.

[3] Bardey D., Bommier A. and B. Jullien, 2010, Retail Price Regulation and Innovation: Reference Pricing in the Pharmaceutical Industry, *Journal of Health Economics*, Elsevier, vol. 29(2), 303-316.

[4] Brekke K., Konigbauer I. and O. Straume, 2007, Reference Pricing of Pharmaceuticals, *Journal of Health Economics*, 26, 613-642.

[5] Danzon P. M., 2001, Reference Pricing: Theory and Evidence, in Reference Pricing and Pharmaceutical Policy, G. Lopez-Casasnovas and B. Jonsson (eds.), New York: Springer.

[6] Danzon P.M., 2012, Regulation of Price and Reimbursement for Pharmaceuticals, in The Oxford Handbook of The Economics of the Biopharmaceutical Industry, Oxford University Press, Oxford.

[7] Danzon P. and J. Ketchman, 2004, Reference Pricing of Pharmaceuticals for Medicare: Evidence from Germany, the Netherlands and New Zealand, *Frontiers in Health Policy Research 7:2,* Bepress.

[8] Danzon, P.M. and H. Liu., 1996, Reference Pricing and Physician Drug Budgets: The German experience in controlling pharmaceutical expenditures, Working Paper, The Wharton School, University of Pennsylvania.

[9] DiMasi JA and C. Paquette, 2004,. The Economics of Follow-On Drug Research and Development, *Pharmacoeconomics*, 22 Supl 2: 1-14.

[10] Frothingham R., 2004, Me-Too" Products — Friend or Foe?, *New England Journal of Medicine* 350(20), 2100-2101.

[11] Geoffard P.-Y., 2006, Incentive and Selection Effects in Health Insurance, in Elgar Companion to Health Economics, Andrew Jones (ed.), Elgar.

[12] Kyle M., 2007, Pharmaceutical Price Controls and Entry Strategies, *Review of Economics and Statistics,* 89(1), 88-99.

[13] Lakdawalla D. and N. Sood, 2006, Health Insurance as a Two-Part Pricing Contract, NBER Working Paper No. 12681.

[14] Lakdawalla D. and N. Sood, 2009, Innovation and the Welfare Effects of Public Drug Insurance, *Journal of Public Economics*, 93(3-4), 541-548.

[15] Lopez-Casanovas G. and J. Puig-Junoy, 2000, Review of the Literature on Reference Pricing, *Health Policy*, 54, 87-123.

[16] Miraldo M., 2007, Reference Pricing Versus Co-Payment in the Pharmaceutical Industry: Firms' Pricing Strategies, Center for Health Economics, University of York, Research Paper 27.

[17] OECD, 2011, Health at a Glance, OECD indicators, http://www.oecd.org/health/healthpoliciesanddata/49105858.pdf.

[18] Pammolli F. and M. Riccaboni, 2004, Market Structure and Drug Innovation, *Health Affairs*, 23(1), 48-50.

[19] Wertheimer A., Levy R. and T. O'Connor, 2001, Too Many Drugs? The Clinical and Economic Value of Incremental Innovations, Irina Farquhar, Kent Summers, Alan Sorkin, in (ed.) Investing in Health: The Social and Economic Benefits of Health Care Innovation (Research in Human Capital and Development, Volume 14), Emerald Group Publishing Limited, 77-118.

Appendix

A Proof of Proposition 1

The equilibrium condition is

$$\frac{\pi}{N} + \frac{\partial D}{\partial p_i}\left(a + (1-a)r\right) P = 0. \tag{19}$$

Given prices p_i and p, the probability that drug i is chosen by a patient j is determined by

$$\begin{aligned} \Pr\left(p_i + \theta x_j^i = \min_l \left\{p + \theta \tilde{x}_j^l\right\} \mid x_j^i\right) &= \Pr\left(\min_{l \neq j}\left\{\tilde{x}_j^l\right\} \geq \frac{p_i - p}{\theta} + x_j^i \mid x_j^i\right) \\ &= \left(1 - F\left(\frac{p_i - p}{\theta} + x_j^i\right)\right)^{N-1}. \end{aligned}$$

Integrating over all possible realizations of x_j^i, the expected demand of drug i is thus:

$$D = \pi \int_0^\infty f(x)\left(1 - F\left(\frac{p_i - p}{\theta} + x\right)\right)^{N-1} dx = \frac{\pi}{N} \exp\left(-\frac{N-1}{\theta}\left(p_i - p\right)\right).$$

It yields

$$\frac{\partial D}{\partial p_i} = -\left(\frac{\pi}{N}\right)\frac{N-1}{\theta}.$$

Substituting this expression in (19) thus yields:

$$P = \frac{1}{a + (1-a)r}\frac{\theta}{N-1}.$$

B Proof of Proposition 2

For a fixed price $P = \varphi(N)$ the slope $\partial EU/\partial a$ increases in λ implying that the optimal copayment is non-decreasing in λ. The first order condition at $a = 0$ can be rewritten (using $w_h = w_s = w - \pi(1+\lambda)P$)

$$(1-\pi)\,u'(w-\rho) + \pi\frac{\eta}{Q(N)}u'(w-\rho) \leq \frac{\eta u'(w-\rho)}{Q(N)(1+\lambda)},$$

or

$$\lambda \leq \frac{\eta - Q(N)}{(1-\pi)\,Q(N) + \pi\eta}(1-\pi). \tag{20}$$

Similarly, we have $a = 1$ if

$$(1-\pi)\,u'(w) + \pi\frac{\eta}{Q(N)}u'(w-P) \geq \frac{\eta u'(w-P)}{Q(N)(1+\lambda)},$$

or

$$\lambda \geq \frac{\eta\exp(\sigma\varphi(N)) - Q(N)}{(1-\pi)Q(N) + \pi\eta\exp(\sigma\varphi(N))}(1-\pi). \tag{21}$$

The RHS of (20) and (21) decrease in N and take value λ_0 when $Q = 1$. Hence, (20) holds for all Q if $\lambda \leq \lambda_0$. When $(1-\pi)/\pi > \lambda > \lambda_0$ on the contrary, the condition (21) is satisfied if $N \geq \bar{N}$ where $\bar{N}$ solves (21) with equality. Finally, when $\lambda \geq (1-\pi)/\pi$, the condition holds for all N.

C Proof of Proposition 3

We first prove that the objective is quasi-concave in lemma 1 and then prove proposition 3.

Lemma 1 *The objective is quasi-concave in N.*

Proof. The slope

$$\frac{\partial EU}{\partial N} = -\pi\frac{\partial}{\partial N}\left(\frac{\eta}{Q}\right)u(w_s) - k(1+\lambda)(1-a)\left[(1-\pi)u'(w_h) + \frac{\eta}{Q}u'(w_s)\right] - \frac{\eta}{Q}aku'(w_s)$$

can be written as $u'(w_s)\,\Lambda$ where

$$\Lambda = -\frac{\pi}{\sigma}\frac{\partial}{\partial N}\left(\frac{\eta}{Q}\right) - k(a+(1-a)\pi(1+\lambda))\frac{\eta}{Q} - k(1+\lambda)(1-a)(1-\pi)\exp\left(-\sigma a\frac{kN}{\pi}\right) \tag{22}$$

We have

$$\begin{aligned}\frac{\partial\Lambda}{\partial N} &= -\frac{\pi}{\sigma}\frac{\partial}{\partial N^2}\left(\frac{\eta}{Q}\right) - k(a+(1-a)\pi(1+\lambda))\frac{\partial}{\partial N}\left(\frac{\eta}{Q}\right) + \\ &\quad a\sigma\frac{k}{\pi}\left[(1-a)k(1-\pi)\exp\left(-\sigma a\frac{kN}{\pi}\right)(1+\lambda)\right].\end{aligned}$$

Hence:

$$\begin{aligned}\left.\frac{\partial\Lambda}{\partial N}\right|_{\Lambda=0} &= -\frac{\pi}{\sigma}\frac{\partial}{\partial N^2}\left(\frac{\eta}{Q}\right) - k(a+(1-a)\pi(1+\lambda))\frac{\partial}{\partial N}\left(\frac{\eta}{Q}\right) + \\ &\quad a\sigma\frac{k}{\pi}\left[-\frac{\pi}{\sigma}\frac{\partial}{\partial N}\left(\frac{\eta}{Q}\right) - k(a+(1-a)\pi(1+\lambda))\frac{\eta}{Q}\right].\end{aligned}$$

Using

$$\begin{aligned}\frac{\partial}{\partial N}\left(\frac{\eta}{Q}\right) &= -\frac{\sigma\theta\eta}{(N-\sigma\theta)^2} = -\frac{\sigma\theta}{N(N-\sigma\theta)}\frac{\eta}{Q},\\ \frac{\partial}{\partial N^2}\left(\frac{\eta}{Q}\right) &= \frac{2\sigma\theta\eta}{(N-\sigma\theta)^3} = \frac{2\sigma\theta}{N(N-\sigma\theta)^2}\frac{\eta}{Q},\end{aligned}$$

we obtain

$$\begin{aligned}\left.\frac{\partial\Lambda}{\partial N}\right|_{\Lambda=0} = & -\frac{\pi}{\sigma}\frac{2\sigma\theta}{N(N-\sigma\theta)^2}\frac{\eta}{Q} + k\left(a+(1-a)\pi(1+\lambda)\right)\frac{\sigma\theta}{N(N-\sigma\theta)}\frac{\eta}{Q} + \\ & a\sigma\frac{k}{\pi}\left[\frac{\pi}{\sigma}\frac{\sigma\theta}{N(N-\sigma\theta)}\frac{\eta}{Q} - k\left(a+(1-a)\pi(1+\lambda)\right)\frac{\eta}{Q}\right],\end{aligned}$$

so

$$\frac{Q}{\eta}\left.\frac{\partial\Lambda}{\partial N}\right|_{\Lambda=0} = \begin{pmatrix} -\frac{2\pi}{\sigma(N-\sigma\theta)} \\ +k\left(2a+(1-a)\pi(1+\lambda)\right)\end{pmatrix}\frac{\sigma\theta}{N(N-\sigma\theta)} - a\frac{\sigma k}{\pi}k\left(a+(1-a)\pi(1+\lambda)\right). \tag{23}$$

Notice that $\Lambda = 0$ implies that

$$\begin{aligned}-\frac{\pi}{\sigma}\frac{\partial}{\partial N}\left(\frac{\eta}{Q}\right) - k\left(a+(1-a)\pi(1+\lambda)\right)\frac{\eta}{Q} &= \\ \left(\frac{\pi\sigma\theta}{\sigma N(N-\sigma\theta)} - k\left(a+(1-a)\pi(1+\lambda)\right)\right)\frac{\eta}{Q} &> 0\end{aligned}$$

implying (because $N > \sigma\theta$):

$$k\left(a+(1-a)\pi(1+\lambda)\right) < \frac{\pi}{\sigma(N-\sigma\theta)}.$$

Hence, we have $-2\pi/\left(\sigma(N-\sigma\theta)\right) + k\left(2a+(1-a)\pi(1+\lambda)\right) < 0$ so that the first term in brackets in the RHS of (23) is negative hence $\partial\Lambda/\partial N|_{\Lambda=0} < 0$ at any point where $\Lambda = 0$. This implies that $\partial^2 EU/\partial N^2 < 0$ when $\partial EU/\partial N = 0$ so that the objective is quasi-concave. ■

By quasi-concavity, $r < 1$ if and only if the RHS of (22) is positive *i.e.* if:

$$\frac{\pi}{\sigma}\frac{\eta}{Q^2}\frac{\partial Q(\underline{N})}{\partial N} - k\frac{\eta}{Q(\underline{N})} - (1-a)k\left(\begin{bmatrix}(1-\pi)\exp\left(-\sigma a\frac{k\underline{N}}{\pi}\right) \\ +\pi\frac{\eta}{Q(\underline{N})}\end{bmatrix}(1+\lambda) - \frac{\eta}{Q(\underline{N})}\right) > 0 \tag{24}$$

where the LHS of (24) is concave in a. This implies that $r < 1$ on an interval $(\underline{a}, \overline{a})$.

One has

$$\frac{\eta}{Q^2}\frac{\partial Q}{\partial N} = \frac{\sigma\theta}{N\left(N-\sigma\theta\right)}\frac{\eta}{Q},$$

so that

$$\frac{\pi}{\sigma}\frac{\eta}{Q(\underline{N})^2}\frac{\partial Q(\underline{N})}{\partial N} - k\frac{\eta}{Q(\underline{N})} = \left(\frac{\pi\theta}{k\underline{N}} - \underline{N} + \sigma\theta\right)\frac{k}{(\underline{N}-\sigma\theta)}\frac{\eta}{Q(\underline{N})}.$$

But the free entry condition yields

$$k\underline{N} = \frac{\pi\theta}{(\underline{N}-1)} \Longrightarrow -1 = \frac{\pi\theta}{k\underline{N}} - \underline{N}.$$

Therefore, we obtain

$$\frac{\pi}{\sigma}\frac{\eta}{Q(\underline{N})^2}\frac{\partial Q(\underline{N})}{\partial N} - k\frac{\eta}{Q(\underline{N})} = (-1+\sigma\theta)\frac{k\phi(\underline{N})}{\underline{N}-\sigma\theta}.$$

Substituting this last expression in (24) , it yields $r < 1$ if

$$\left(\frac{\sigma\theta-1}{\underline{N}-\sigma\theta}\right)k\frac{\eta}{Q(\underline{N})} > \left(\frac{1-a}{\underline{N}u'\left(w_s\right)}\right)\left.\frac{\partial EU}{\partial a}\right|_{\underline{N}}, \tag{25}$$

where $\left.\frac{\partial EU}{\partial a}\right|_{\underline{N}}$ is evaluated at $r = 1$.

$$\begin{aligned}\left(\frac{1-a}{\underline{N}u'\left(w_s\right)}\right)\left.\frac{\partial EU}{\partial a}\right|_{\underline{N}} &= \left(\frac{1-a}{\underline{N}}\right)\left(\left(\frac{\left(1-\pi\right)u'(w_h)}{u'\left(w_s\right)} + \left(\pi\frac{\eta}{Q}\right)\right)(1+\lambda) - \frac{\eta}{Q}\right)\pi\varphi\left(\underline{N}\right) \\ = \quad &\left(\frac{1-a}{\underline{N}}\right)\left(\left(\left(1-\pi\right)\exp\left(-\sigma a\varphi\left(\underline{N}\right)\right) + \left(\pi\frac{\eta}{Q}\right)\right)(1+\lambda) - \frac{\eta}{Q}\right)\pi\varphi\left(\underline{N}\right),\end{aligned}$$

which is decreasing with a when positive.

Suppose that $\sigma\theta > 1$. Then as $\left(\frac{1-a}{\underline{N}u'(w_s)}\right)\left.\frac{\partial EU}{\partial a}\right|_{\underline{N}}$ decreases when positive, condition (25) holds on an interval $(\underline{a}, 1)$. This interval is non-empty as the RHS vanishes at $a = 1$. At $a = 0$, the condition is verified if $\partial EU/\partial a|_{\underline{N},a=0} < 0$ (*i.e.* full insurance is optimal), in which case $\underline{a} = 0$.

For $\sigma\theta = 1$, the same conclusion holds because our assumption that $\underline{N} < \bar{N}$ ensures that $\left(\frac{1-a}{\underline{N}u'(w_s)}\right)\left.\frac{\partial EU}{\partial a}\right|_{\underline{N}} < 0$ for a close to 1.

If $\sigma\theta < 1$, the condition is violated at $a \simeq 1$. The function $\left(\frac{1-a}{\underline{N}u'(w_s)}\right)\left.\frac{\partial EU}{\partial a}\right|_{\underline{N}}$ is convex. Thus, the condition holds on some interval $(\underline{a}, \overline{a})$ which may be empty.

D Proof of Proposition 4

At any point where N_a^* is interior, the effect of a parameter on N_a^* is given by the sign of the derivative of Λ with respect to this parameter. Differentiating the LHS of (22) with respect to a and λ respectively yields:

$$\begin{aligned}\frac{\partial \Lambda}{\partial a} &= \left(\frac{1}{Nu'(w_s)}\right)\frac{\partial EU}{\partial a}\bigg|_N + (1-a)\,k\sigma\frac{kN}{\pi}(1-\pi)\exp\left(-\sigma a\frac{kN}{\pi}\right)(1+\lambda),\\ \frac{\partial \Lambda}{\partial \lambda} &= -(1-a)\,k\left[(1-\pi)\exp\left(-\sigma a\frac{kN}{\pi}\right)+\pi\frac{\eta}{Q}\right]<0,\end{aligned}$$

so that $\partial\Lambda/\partial a > 0$ if $\partial EU/\partial a|_N > 0$ and $\partial\Lambda/\partial\lambda < 0$. Then, using

$$\frac{\partial}{\partial N}\left(\frac{\eta}{Q}\right) = -\frac{\sigma\theta}{N(N-\sigma\theta)}\frac{\eta}{Q}$$

we have

$$\begin{aligned}\frac{Q}{\eta}\Lambda &= \frac{\pi\theta}{N(N-\sigma\theta)} - (a+(1-a)\,\pi(1+\lambda))\,k - \\ &\quad (1-a)\,k\,(1-\pi)\exp\left(-\sigma a\frac{kN}{\pi}\right)\frac{N-\sigma\theta}{N}(1+\lambda),\end{aligned}$$

which is increasing in σ and θ. Thus $\partial\Lambda/\partial\sigma > 0$ and $\partial\Lambda/\partial\theta > 0$ at $\Lambda = 0$.

E Proof of Proposition 5

Suppose a^* decreases in λ for some value. Then it must be the case that $a^* > 0$. Moreover $N^* > \underline{N}$ because a^* is non-decreasing in λ for a fixed N from proposition 2. At such a point, the first-order conditions are:

$$\begin{aligned}EU'(.) - \frac{\frac{\eta}{Q}u'(w_s)}{1+\lambda} &= 0,\\ \pi\frac{\partial}{\partial N}\left(\frac{\eta}{Q}\right)u(w_s) - k\frac{\eta}{Q}u'(w_s) &= 0,\end{aligned}$$

which can be rewritten (up to a proportional factor) as:

$$\nabla_a = (1+\lambda)(1-\pi)\frac{(N^*-\sigma\theta)}{\eta N^*}\exp\left(-\sigma a^*\frac{kN^*}{\pi}\right)+\pi(1+\lambda)-1=0, \quad (26)$$

$$\nabla_N = \pi\theta - kN^*(N^*-\sigma\theta) = 0. \quad (27)$$

We have at $\nabla_a = \nabla_N = 0$ (using $(N^* - \sigma\theta)\,\sigma k/\pi = \sigma\theta/N^*$ for ∇_{aN}):

$$\begin{aligned}
\nabla_{aa} &= -(1+\lambda)(1-\pi)(N^* - \sigma\theta)\,\sigma\frac{k}{\eta\pi}\exp\left(-\sigma a^*\frac{kN^*}{\pi}\right) < 0, \\
\nabla_{aN} &= (1+\lambda)(1-\pi)\exp\left(-\sigma a^*\frac{kN^*}{\pi}\right)\frac{\sigma\theta}{\eta N^{*2}}(1-a^*) > 0, \\
\nabla_{Na} &= 0, \quad \nabla_{NN} = -2kN^* + k\sigma\theta < 0.
\end{aligned}$$

First, notice that the determinant of the modified Hessien is $\det = \nabla_{aa}\nabla_{NN} > 0$, hence the matrix is semi-definite negative. Differentiation of a^* with respect to λ gives

$$\frac{\partial a^*}{\partial \lambda} = \frac{\begin{vmatrix} -\nabla_{a\lambda} & \nabla_{aN} \\ -\nabla_{N\lambda} & \nabla_{NN} \end{vmatrix}}{\det} = \frac{-\nabla_{a\lambda}\nabla_{NN}}{\det} > 0,$$

because $\nabla_{a\lambda} > 0$ and $\nabla_{N\lambda} = 0$. But this contradicts the assumption made that a^* decreases in λ. Thus, a^* is non-decreasing with λ and $a^* = 0$ for $\lambda \leq \underline{\lambda}$. Notice also that at an interior solution

$$\frac{dN^*}{d\lambda} = 0.$$

Hence, N^* is constant in λ when it is larger than $\underline{N}$ on $\lambda > \underline{\lambda}$. Since N^* decreases in λ for $\lambda \leq \underline{\lambda}$ from proposition 4, we conclude that it is non-increasing everywhere. Moreover, notice that $\nabla_N(\underline{N}) = k\underline{N}(\underline{N} - 1) - k\underline{N}(\underline{N} - \sigma\theta) = k\underline{N}(\sigma\theta - 1)$ is positive if and only if $\sigma\theta > 1$. Hence $r^* < 1$ for all a^* if $\sigma\theta > 1$. Finally, if $\sigma\theta < 1$, $r^* = 1$ if $a^* > 0$.

F Proof of Proposition 6

First, from equation (27), it is immediate that N^* increases in σ, θ and π, and N^* decreases in k. Applying the Cramer's rule yields

$$\frac{da^*}{d\sigma} = \frac{\begin{vmatrix} -\nabla_{a\sigma} & \nabla_{aN} \\ -\nabla_{N\sigma} & \nabla_{NN} \end{vmatrix}}{\det}$$

where

$$\begin{aligned}
\nabla_{a\sigma} &= -(1+\lambda)(1-\pi)\exp\left(-\sigma a^*\frac{kN^*}{\pi}\right)\frac{\theta}{\eta N^*}(1+a^*) \\
\nabla_{aN} &= (1+\lambda)(1-\pi)\exp\left(-\sigma a^*\frac{kN^*}{\pi}\right)\frac{\sigma\theta}{\eta N^{*2}}(1-a^*) \\
\nabla_{N\sigma} &= -\theta k N^* \\
\nabla_{NN} &= -k(2N^*-\sigma\theta)
\end{aligned}$$

so that

$$\frac{da^*}{d\sigma} = -\frac{2(1+\lambda)(1-\pi)\theta k\exp\left(-\sigma a^*\frac{kN^*}{\pi}\right)}{\eta N^*\det}\left[N^*-\sigma\theta+a^*N^*\right] < 0.$$

Moreover,

$$\frac{da^*}{d\theta} = \frac{\begin{vmatrix} -\nabla_{a\theta} & \nabla_{aN} \\ -\nabla_{N\theta} & \nabla_{NN} \end{vmatrix}}{\det}$$

where

$$\begin{aligned}
\nabla_{a\theta} &= -(1+\lambda)(1-\pi)\frac{\sigma}{\eta N^*}\exp\left(-\sigma a^*\frac{kN^*}{\pi}\right) \\
\nabla_{N\theta} &= \pi + kN^*\sigma
\end{aligned}$$

so that

$$\frac{da^*}{d\theta} = \frac{\sigma(1+\lambda)(1-\pi)\exp\left(-\sigma a^*\frac{kN^*}{\pi}\right)}{\eta N^*\det}\left(-k(2N^*-\sigma\theta)+(\pi+kN^*\sigma)\frac{\theta}{N^*}(1-a^*)\right).$$

Using $\pi\theta = kN^*(N^*-\sigma\theta)$, we obtain

$$\frac{da^*}{d\theta} = -\frac{\sigma(1+\lambda)(1-\pi)\exp\left(-\sigma a^*\frac{kN^*}{\pi}\right)}{\eta N^*\det}k\left(N^*(1+a^*)-\sigma\theta\right) < 0.$$

Finally, we have

$$\frac{da^*}{dk} = \frac{\begin{vmatrix} -\nabla_{ak} & \nabla_{aN} \\ -\nabla_{Nk} & \nabla_{NN} \end{vmatrix}}{\det} < 0,$$

as ∇_{ak}, ∇_{NN} and ∇_{Nk} are negative while ∇_{aN} is positive. Similarly, $\nabla_{a\eta} < 0$ and $\nabla_{N\eta} = 0$ imply $da^*/d\eta < 0$.

G Proof of Proposition 7

Since

$$P_k \in \arg\max P_k D_k\left(P_k, P\right)$$

every price P_k that satisfy

$$\begin{aligned} D^k\left(P_k, P\right) + P^k \left.\frac{\partial D_k\left(P_k, P\right)}{\partial P_k}\right|_{P_k \leq P} &\leq 0 \text{ for } P^k = P \\ D^k\left(P_k, P\right) + P^k \left.\frac{\partial D_k\left(P_k, P\right)}{\partial P_k}\right|_{P_k \geq P} &\geq 0 \text{ for } P^k = P \end{aligned}$$

are candidates for symmetric equilibria. Since

$$D^k + P^k \left.\frac{\partial D_k\left(P, P\right)}{\partial P_k}\right|_{P_k \leq P} = 0$$

leads to

$$P_k^* = \frac{\theta}{\left(N-1\right)\alpha + \tau_0 a}$$

and

$$D^k + P^k \left.\frac{\partial D_k\left(P_k, P\right)}{\partial P_k}\right|_{P_k \geq P} = 0$$

leads to

$$P_k^{**} = \frac{\theta}{\left(N-1\right)\alpha + \tau_0 \alpha},$$

and that $P_k^{**} < P_k^*$, any price belonging to the interval $\left(P_k^{**}, P_k^*\right) = \left(\underline{P}, \bar{P}\right)$ are possible candidates for the symmetric price competition game.

H Proof of Corollary 1

The optimal diversity N solves

$$\begin{aligned} &\max_N EU = \pi\eta u\left(w_h\right)\left(\frac{N+\tau_0}{N+\tau_0-\sigma\theta}\right) + \left(1-\pi\right)u\left(w_h\right) \\ &s.t.\ w_h = w - (1+\lambda)kN \text{ and } \left(N+\tau_0\right)\left(N-1\right) \leq \frac{\pi\theta}{k}. \end{aligned}$$

The gain of raising diversity is now captured by:

$$\begin{aligned} \frac{\partial EU}{\partial N} &= \pi\eta\left(\frac{-\sigma\theta}{\left(N+\tau_0-\sigma\theta\right)^2}\right)u\left(w_h\right) \\ &\quad - \left(1+\lambda\right)k\left(\left(1-\pi\right) + \pi\frac{N+\tau_0}{N+\tau_0-\sigma\theta}\right)u'\left(w_h\right). \end{aligned}$$

Thus at any interior equilibrium one has

$$\pi\theta - (1+\lambda)\, k\,(N+\tau_0-\sigma\theta)\,(N+\tau_0-(1-\pi)\,\sigma\theta) = 0.$$

The LHS is decreasing with N and τ_0, which shows that N is unique and $\partial N/\partial\tau_0 < 0$ when N is interior.

I Proof of Proposition8

Note first that if $1-F(x)$ is log-log concave then

$$\frac{d^2\left[\log\log\left(1-F(x)\right)\right]}{dx^2} = \frac{d}{dx}\left[\frac{\frac{f(x)}{1-F(x)}}{-\log\left(1-F(x)\right)}\right] < 0, \tag{28}$$

for any $x \in [0,\bar{x}]$. Since $1/Q(N)$ increases with σ, it follows immediately that $Q(N) < NE(\exp(\sigma\theta x))$ so that $Q(N)$ exists for σ small. Notice also that $\lim_{\sigma\to\sigma_N} Q(N) = +\infty$. Integrating (18) by part yields**:**

$$Q(N) = -\left\{(1-F(x))^N \exp\sigma\theta x\right\}_0^{\bar{x}} + \sigma\theta\int_0^{\bar{x}} (1-F(x))^N \exp(\sigma\theta x)\, dx.$$

If the integral exists $(1-F(x))^N \exp\sigma\theta x$ tends to zero at the upper boundary so that

$$Q(N) = 1 + \sigma\theta\int_0^{\bar{x}} (1-F(x))^N \exp[\sigma\theta x]\, dx.$$

Suppose that

$$\sigma\theta\int_0^{\bar{x}} \log\left(1-F(x)\right)(1-F(x))^N \exp[\sigma\theta x]\, dx$$

exists, then

$$-\frac{\pi}{\sigma}\frac{\partial\log(Q(N))}{\partial N} = \pi\theta\frac{\int_0^{\bar{x}} -\log\left(1-F(x)\right)(1-F(x))^N \exp[\sigma\theta x]\, dx}{\int_0^{\bar{x}} \exp(\sigma\theta x)\, N\,(1-F(x))^{N-1} f(x)\, dx}.$$

This increases with σ if

$$\frac{\int_0^{\bar{x}} -\theta x \exp[\sigma\theta x]\log\left(1-F(x)\right)(1-F(x))^N dx}{\int_0^{\bar{x}} -\exp[\sigma\theta x]\log\left(1-F(x)\right)(1-F(x))^N dx} > \frac{\int_0^{\bar{x}} \theta x \exp(\sigma\theta x)\,(1-F(x))^{N-1} f(x)\, dx}{\int_0^{\bar{x}} \exp(\sigma\theta x)\,(1-F(x))^{N-1} f(x)\, dx}$$

or if

$$\int_0^{\bar{x}} \theta x h(x)\, dx > \int_0^{\bar{x}} \theta x l(x)\, dx, \tag{29}$$

where $h(x)$ and $l(x)$ are two densities defined by:

$$\begin{aligned} h(x) &= \frac{-\exp(\theta x)\log(1-F(x))(1-F(x))^N}{\int -\exp(\sigma\theta x)\log(1-F(x))(1-F(x))^N dx}, \\ l(x) &= \frac{\exp(\sigma\theta x)(1-F(x))^{N-1} f(x)\,dx}{\int \exp(\sigma\theta x)(1-F(x))^{N-1} f(x)\,dx}. \end{aligned}$$

Since θx is increasing, a sufficient condition for (29) to hold is that the distribution with density $h(x)$ first order stochastically dominates (FOSD) the distribution with density $l(x)$ *i.e.*

$$\int_0^z h(x)\,dx < \int_0^z l(x)\,dx,$$

for any $z \in [0, \bar{x}]$. Using the Rothchild-Stiglitz single crossing property, this is true if

$$\frac{d}{dz}\left.\frac{l(z)}{h(z)}\right|_{l=h} < 0 \text{ for every } z \in [0, \bar{x}].$$

Since

$$\frac{l(z)}{h(z)} = \frac{\frac{f(z)}{1-F(z)}}{-\log(1-F(z))} * \frac{\int_0^{\bar{x}} -\exp(\sigma\theta x)\log(1-F(x))(1-F(x))^N dx}{\int_0^{\bar{x}} \exp(\sigma\theta x)(1-F(x))^{N-1} f(x)\,dx},$$

straightforward differentiation with respect to z yields

$$\operatorname{sign}\left(\frac{d}{dz}\left.\frac{l(z)}{h(z)}\right|_{l=h}\right) = \operatorname{sign}\left[\frac{d}{dz}\left(\frac{\frac{f(z)}{1-F(z)}}{-\log(1-F(z))}\right)\right],$$

which is negative if $(1-F(x))$ is log-log concave as shown by (28).