A model investigating incentives for illegal cultural goods sharing sites to become legal

Jacques Lévy Véhel, Pierre-Emmanuel Lévy Véhel, Victor Lévy Véhel

A MODEL INVESTIGATING INCENTIVES FOR ILLEGAL CULTURAL GOODS SHARING SITES TO BECOME LEGAL

Jacques Lévy Véhel
Regularity team, INRIA Saclay and MAS laboratory
jacques.levy-vehel@inria.fr
Pierre-Emmanuel Lévy Véhel
La Banque Postale
pierem2@hotmail.com
Victor Lévy Véhel
Student at ENS Lyon
victor.levy.vehel@gmail.com

December 16, 2014

Abstract

Illegal sharing of cultural goods on the Internet has become a massive reality in today's connected society. Numerous studies have been performed to try and evaluate the impact of these practices on the industry of cultural goods, and how much harm, if any, they have entailed. The effect of legal and technical responses to limit pirating has also been investigated, showing in general inconclusive effect. Instead of penalizing illegal actors - providers and/or consumers -, a totally different approach has been proposed recently by the french government agency Hadopi. The idea is to offer the possibility to sites that illegally share cultural goods to become legal in exchange of a retribution proportional to their activity. The present work builds on parts of the elegant model proposed in [1] to study the economic feasibility of such a scheme under various assumptions on the behaviour of the different actors involved. Our main finding is that, supposing that more popular goods are more prone to pirating, a retribution of the order of the increase in benefit per user gained by legalized sites does indeed lead to a win-win situation for both producers/sellers of cultural goods and willing-to-be-legalized sites.

1 Context

Digital cultural goods, including music, movies and books, have been subject to intense illegal circulation over the Internet for more than fifteen years. Technologies used for this purpose have greatly evolved over time, and nowadays a large array of possibilities are available. They include e.g. DDL, Streaming or BitTorrent. Legal responses to reduce such practices have shown limited efficiency, in spite of a few largely publicized cases. Legal actions targeting illegal providers are sometimes successful, but the long list of services of different natures which were closed or modified due to such actions, which includes e.g. Napster, Scour, Morpheus, Kazaa, Grokster, Limewire, BitTorrent, The Pirate Bay, Megaupload and

Hotfile, shows that technology seems to always keep one step ahead of judicial responses. Even when legal actions succeed in blocking an illegal offer and momentarily reduce pirate traffic, this one soon recovers its pre-blocking level, as witnessed for instance by the case of Megaupload.

Directly targeting and suing consumers sometimes makes the headlines, but does not seem to be more efficient, as this does not significantly modify the behaviour of pirate users in the medium term. For instance, in France, the governmental agency Hadopi [2] has tried to enforce legal sharing at the consumers' level by implementing a mechanism called "gradual response". As stated in [3], *targeting individual internet users is not likely to reverse the trend toward an online sharing culture, and there is an urgent need for independent verification of claims of harm to the creative industries as a result of individual copyright*. The second part of this statement raises an important question: independently of the questionable efficiency of legal actions, it is not clear how illegal downloading affects legal consumption of cultural goods. Indeed, despite a long list of informed works on the topic, there are no clear answers to such questions as (1) whether consumer who have a significant use of illegal download have reduced their legal consumption, (2) whether a notable proportion of illegal consumption would automatically convert into legal consumption were pirate downloading made impossible, (3) whether pirate downloading fosters legal consumption in a measurable way by allowing consumers to get to know artists they would not have come across otherwise. We refer the reader to the articles [4, 5, 6, 7, 8, 9, 10, 11, 12] (a small sample of the available literature), some of which conclude that pirating does hurt in a significant way the industry of cultural goods, while an equal number states that such an effect cannot be asserted. We note that some of ideas of, for instance, [7, 12], are captured in the model proposed in [1], which is our main source of inspiration.

Instead of pursuing this apparently endless play at cops and robbers between technology and copyright holders, one may seek alternative solutions that would respect the rights of authors and industry in an era of massive digital exchange. A path that has been proposed in France in 2001 by ADAMI [13] and SPEDIDAM [14] is based on the concept of a *global license* [15, 16]: the essential idea is to legalize circulation of cultural goods in exchange for a fee on broadband internet subscriptions, either in a "universal" fashion, that is, the fee applies to any individual possessing a broadband internet connection, or optionally, where it would concern only users that intend to download cultural goods. Advantages of such a license are that it avoids the complications related to suing pirates while providing a stable revenue to artists and the industry. Drawbacks include the fact that legal issues related to this scheme are not clear, as well as the difficulty to redistribute the money to artists in a fair way. Though global license almost made it to be voted by the French Assembly, it was finally dismissed, but the idea is still supported by various actors in France and around the world. Indeed, similar concepts have been proposed under various denominations such as *alternative compensation system, artistic freedom voucher* or *non-commercial use levy* in many countries including Australia, Canada, Germany, the UK and the USA.

The aim of this work is to investigate another path to let producers of cultural goods recover some of the retribution they are entitled to. The new and original idea, which has been proposed and is promoted by the Hadopi, is to provide incentive for (some) illegal forums to become legal in exchange of a monetary compensation, proportional to the amount of downloaded goods. In other words, the possibility would be offered to forums to become legal provided they pay a contribution computed as a fraction of the money they make from downloads. Such a scheme would be implemented by a public regulation authority: indeed, one may argue that society would benefit from the legalization of forums both in an immaterial way (a environment with less crime, more peaceful relations between forums and industry) and in a material one, with added fiscal incomes from industry revenue, or even avoiding extreme negative externalities such as the collapse of parts of the industry if pirating gets out of control. Note that, in a sense, this proposition is a sort of "global licence" with the twist that this fee is financed by advertisement.

In order to study the economic feasibility of this proposition[1], we set up a simple model for the circulation of cultural goods and the financial fluxes it produces. Under some assumptions on the behaviour of consumers, we compute the profits of different actors such as the industry of cultural goods or illegal/legalized forums, both in the current situation and in a framework where a proportional voluntary retribution would be paid to industry by forums willing to become legal. This allows us to highlight scenarios where both the industry and forums would financially benefit from such a shift. As is shown below, a win-win situation does exist under the reasonable assumption that more popular goods are more prone to pirating. A noteworthy fact is that, in the favourable case, the retribution will be of the order of the increase in benefit per user gained by legalized sites: thus the industry will absorb all the benefit coming from increased "price", while forums will make more money from the fact that they will have more users. See Sections 4.3 and 5 for more details.

The literature on economic models which evaluate in a quantitative way the effect of illegal download is as rich as the one dealing with empirical studies, and it is by no means possible to list it extensively. We content ourselves with mentioning that, besides [1], which is the starting point of our work, our approach has some common features in particular with [11, 17, 18, 19], Less related works include, among many others, [8] and Section 6.1 of [12], which use a welfare-theoretical approach, or [20], in which an economy with two types of consumers with different willingness to pay is considered.

Despite the impressive number of articles that propose and study models for the interplay between legal and illegal fluxes of cultural goods, and to the best of our knowledge, there is no structured work that tries to assess the impact of policies aiming at reducing pirate downloading, whether these policies are currently in use, such as legal suing of pirate sites or downloaders, or still discussed at governmental level, such as the global licence or the proposition of Hadopi. Such studies are however needed in order to assess the viability and efficiency of those responses, and to compare them in order to choose the most adequate one. We hope that the current work is a first step in this direction.

The remaining of this article is organized as follows. In Section 2, we present our model in an abstract setting: we formulate some general assumptions on the structure of the different sets of actors (industry, consumers, forums, ...) in Section 2.1, which enable us to compute in Section 2.2 the profits, depending on various relevant parameters such as for instance the price of the goods or the tendency for a given consumer to resort to pirate downloading. This analysis provides general bounds on the retribution that illegal forums would have to pay in exchange for legalisation. These bounds characterize the win-win situations where each actor sees its profits increase. A particular case, termed "separable scenario", where the formula simplify greatly, is treated in Section 2.3. In order to compute numerically the bounds, the parameters of the model must be estimated. A full study of this issue requires a considerable amount of work that goes beyond the aims of the present article and will be performed elsewhere. Nevertheless, we provide in Section 3 some rough but reasonable estimates of the parameters based on existing literature. Section 4 gives explicit computations using these estimates in three cases of interest describing more precisely the structure of both the set of consumers and the one of cultural goods. These numerical calculations allow us to determine the range of acceptable values for the retribution. Finally, Section 5

[1] we do not consider in this work technical or legal aspects. For legal aspects, see [25].

briefly discusses our results and proposes some extensions to our simple model that are desirable to make it more complete.

2 The model

2.1 General frame

Our model is an abstract and generalized version of a part of the one presented in [1], which deals with the specific case of the music industry.

We consider cultural goods of a given type, e.g. music, movies or books. Four to six kind of actors are involved:

1. the industry, which produces and sells the cultural goods;
2. consumers;
3. legal digital platforms, that propose physical or digital versions of the goods for purchase. In the case of the musical or book industry, this corresponds for instance to Amazon.com or Fnac.com;
4. illegal forums for downloading digital versions of the goods.

In some instances (e.g. music), one or two more types of actors are present:

5. platforms providing free and legal streaming services and paying subscriptions. Typical examples include Deezer or Spotify;
6. the artists.

In this preliminary study, only actors 1,2, and 4 are taken into account. In addition, for the sake of simplicity, we consider a monopolistic situation for actor 1, so that we will speak of "the" industry I.

Before we detail the structure of the set of actors, let us mention that all the quantities defined below are expressed in some abstract units whose exact meaning is irrelevant: indeed, in the end, we will characterize situations where proportional retribution is economically feasible, and express this retribution as a function of the “price” of cultural goods and the “earnings” of illegal forums.

Industry and cultural goods

We consider the situation at a given time T. The set G of cultural goods produced by industry at this particular time is modelled as a countable set of points distributed on $2\,\mathbb{N}$, the set of even integers[2]. Each cultural good is identified with its coordinate on the positive real line. These coordinates are denoted $g_i = 2i, i = 0, 1, \dots$ (i will always be used for indexing goods, so that a symbol like $\sum_i$ will mean that the sum runs over all goods g_i in G). All goods are assumed to be sold at the same price p.

Each good g_i is characterized by its “attractiveness” a_i, a real number in $[0, 1]$. The exact meaning of attractiveness and the choice of its range will become clear below. In [1], attractiveness is chosen to

[2] This choice is purely for notational convenience; it helps avoiding unsightly formulas later.

be of the form aq_i, where the subjective "quality" q_i perceived by consumers depends on the "effort" x_i put by industry to promote good g_i in the following way:

$$q_i = x_i^\alpha, \tag{1}$$

with $\alpha \in (0, 1)$ (since quality increases less and less when promotional effort increases). We will have no need of this expression below. Only the values of a_i will matter.

Consumers

Consumers are distributed on the infinite strip $[0, 1] \times \mathbb{R}$. Each consumer C is characterized by its coordinates (θ, x), which are interpreted as follows:

- the distance $\xi = \xi(x) := d(C, G) = \inf_i |x - g_i|$ measures the willingness of consumer $C(x, \theta)$ to consume a good: the smallest $d(C, G)$, the more the consumer is interested in the production. Note that $\xi \in [0, 1]$. In addition, each consumer buys or pirates only the good that is "closest" to him at time T (or chooses the one with smallest coordinate if there are two such goods);

- the coordinate θ characterizes the behaviour of C as regards pirating. We assume that all goods are available for free on illegal forums. In addition, following [1], we consider that pirating a good translates into an immaterial cost for the consumer. This immaterial cost consists of two components: a "comfort cost', that incorporates at least the following elements: (a) a pirated good will typically be of lower quality, (b) it may not be as easily accessible as the legal version, (c) cover and other information, such as subtitles for films, may be unavailable, (d) downloading from a forum generally requires some technical skills, (e) fear from computer viruses. The second component is a "legal cost", that accounts for the fact that typical consumers prefer to remain legal rather than facing the possibility of being prosecuted. Both costs are combined and their effect is taken into account as follows: we assume that the immaterial cost may be written as $\frac{\theta}{\tau}\, p$, where θ follows a distribution whose support is $[0, 1]$ and τ is a real number in $(0, 1)$. The parameter θ measures how reluctant a particular consumer is with regard to piracy: when $\theta < \tau$, the consumer is willing to use illegal downloading, since its immaterial cost is smaller than industry cost p, while values of θ larger than τ prevent piracy.

The behaviour of consumer C is described by the maximal amount $w = w(x)$ that he is prepared to pay to buy good g_i, where $i = i(x)$ is the element of G that is closest to him. [1] proposes that this amount depends on attractiveness a_i and on ξ as follows (we write i instead of $i(x)$ for simplicity)[3]

$$w = a_i - \xi. \tag{2}$$

Each consumer is thus faced with three choices:

- if $p < \frac{\theta}{\tau} p$ and $p < w$, then he will buy the good;
- if $\frac{\theta}{\tau} p = \min(p, \frac{\theta}{\tau} p, w)$, then he will illegally download the good;

[3]In [1], a term $t\xi$ instead of ξ is considered in (2), where t is a positive real. As t is a just a fixed proportionality factor, it may be incorporated into the distance ξ through an appropriate change of units.

- if $\min(p, \frac{\theta}{\tau}p, w) = w$, then he will pass.

These situations are equivalently characterized by the following conditions:

1. buy if $\tau < \theta \leqslant 1$ and $\xi < a_i - p$,
2. pirate if $0 \leqslant \theta < \tau$ and $\xi < a_i - \frac{\theta}{\tau}p$,
3. no consumption if $a_i - \min(1, \frac{\theta}{\tau})p < \xi$.

Note that case 1 is void whenever $a_i < p$; likewise, case 2 is void if $a_i < \frac{\theta}{\tau}p$.

As in [1], we assume that, when there is consumption, the consumer devotes a share $0 \leqslant \rho \leqslant 1$ of what he has left in consuming products related to the good he has bought or pirated[4]: this amounts to $\rho(w-p)$ in the case he has bought the good and to ρw if he used an illegal downloading site. From these amounts, industry collects a fraction σ (where $0 \leqslant \sigma \leqslant \rho$), that is $\sigma(w-p)$ when the good was bought and σw when it was pirated.

The distribution of consumers on $[0,1] \times \mathbb{R}$ is described by a measure $\mu = \mu(d\theta, dx)$ on this set. We will mainly consider three configurations that seem to be of particular interest:

1. Configuration 1 is when $\mu(d\theta, dx) = \mu_\theta(d\theta)\mu_x(dx)$ with μ_θ uniform on $[0,1]$ and μ_x uniform on $[-1, 2n+1]$ for some fixed positive integer n. More precisely, we set $\mu_\theta(d\theta) = d\theta \mathbf{1}_{\{\theta \in [0,1]\}}$ and $\mu_x(dx) = c\, \frac{dx}{2n+2}\mathbf{1}_{\{x \in [-1, 2n+1]\}}$, where $c > 0$ is a constant which represents the total number of consumers. In this model, consumers are thus only interested in the first n goods, and the proportion of pirates is equal to τ. If one assumes in addition, as we will, that $a_i = a$ for all i, then nothing depends on i: all goods are equally popular, and the proportion of consumers that will buy or pirate g_i is the same for all i. This is essentially the case considered in [1].
2. Configuration 2 takes again $\mu(d\theta, dx) = \mu_\theta(d\theta)\mu_x(dx)$ with μ_θ uniform on $[0,1]$, but with μ_x given by $\mu_x(dx) = c\, \frac{dx}{1+x^\gamma},\ c > 0, \gamma > 0$, and supported either on $[0, 2n+1]$ if consumers are only interested in the first n goods, or on $\mathbb{R}^+$ if they are interested in a potentially infinite number of goods (in which case one has to choose $\gamma > 1$). This form serves as rough model for the situation where there are a few popular goods with rapidly decreasing popularity, and "infinitely" many goods with very low popularity. Such Pareto type distributions are often used for this purpose.
3. Configuration 3 is the most realistic one: it allows for a coupling between willingness to pirate and popularity of a good. Assuming that more popular goods are more prone to pirating, we wish to design μ so that, for values of x such that $\mu(d\theta, dx)$ is large (*i.e.* the desired good is popular), then there is more mass when θ close to 0 (*i.e.*, the proportion of pirates increases). Writing $\mu(d\theta, dx) = \mu_x(dx)\ \mu_\theta(d\theta|x)$, a simple choice is to set $\mu_\theta(d\theta|x) = \left((1-\theta)\frac{\mu_x(dx)}{dx} + 1 - \frac{\mu_x(dx)}{2dx}\right)\ d\theta$, that is, the $\theta-$marginal density is a line with decreasing slope equal to $-\frac{\mu_x(dx)}{dx}$ such that $\int_0^1 \mu_\theta(d\theta|x) = 1$ for all x. Simple computations show that, if $\mu_x(dx)$ is chosen as in Configuration 2, then the proportion of pirates among consumers with first coordinate

[4] these can be live music, derived products for movies or video games, ...

between x and $x + dx$, that is $\frac{\int_0^\tau \mu_\theta(d\theta|x)}{\int_0^1 \mu_\theta(d\theta|x)}dx$, is equal to $\frac{\tau(2x^\gamma+3-\tau)}{2(1+x^\gamma)}dx$. For the most popular good, *i.e.* when $x = 0$, this is $\frac{\tau(3-\tau)}{2}$, which is indeed larger than τ, the proportion of pirates for the least popular goods (which corresponds to $x \to \infty$). In general, our choice for $\mu_\theta(d\theta|x)$ ensures that, if the support of μ_x extends to infinity, then the limiting proportion of pirates when x tends to infinity is exactly τ. Other, non linear, choices are of course possible.

Forums

We assume that there exist m forums $F_1, \ldots, F_m$, and that consumers that decide to use illegal download choose at random one of the F_k with probability z_k. Forum F_k (the index k will be reserved for forums) is characterized by a parameter λ_k measuring its reluctance to accept the proposition of the Hadopi. Indeed, some forums might agree to become legal under certain conditions, while others would never do so (in that case, λ is infinite). Without loss of generality, we assume $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_m$. The larger λ, the more reluctant is the forum to become legal.

We wish to model the revenue of illegal forums. A first very rough way to do so is as follows. Forums earn money from advertisements. We assume that they get paid a fixed amount $\tilde{p}$ for each download (although a more correct model would assume that they get paid for each unique visitor to their page). The total profit P_{IF} is then obtained by integrating over all pirate downloads.

Legalizing illegal forums could be performed by asking them to pay the industry a share of their revenue. We choose a form $r_i = r_i(g_i) > 0$ for each download of good g_i for this contribution. Since we assume that all goods are equally available on all forums, the contribution will essentially be equal $r = \sum_i pop(i) r_i$, where $pop(i)$ is the normalized popularity of g_i (normalized means that the sum of all $pop(i)$ is one), Under the assumption of a uniform distribution of x, $pop(i)$ does not depend on i.

Forums that will agree to become legal are then those for which λ_k is smaller that a threshold $\lambda(r)$, a function of the required contribution r. We let L denote the number of such forums, so that $\lambda_L < \lambda(r) \leqslant \lambda_{L+1}$. A forum that becomes legal will typically attract more advertisements, and, more importantly, from legal companies. This will likely increase the amount it gets paid for each download (visitor). We denote this new amount by $\hat{p}$, with $\hat{p} > \tilde{p}$.

For legal forums, there is no more a notion of pirating a good. Legalizing some forums would then seem to imply that this form of consumption will kill both buying cultural goods and pirating them. Indeed, consumers will have access for free to legal goods. This is however not the case. On the one hand, a consumer downloading from a legalized forum is still faced with the comfort cost part of the immaterial cost. This cost will make some consumers still reluctant to use forums, even though they are legalized, and prefer to buy the good. We choose to model the comfort cost in a manner similar to the one of the total immaterial cost of pirating, i.e. in the form $\frac{\theta}{\hat{\tau}}p$. We choose $\hat{\tau}$ in $(0, 1)$ with $\tau < \hat{\tau} \leqslant 1$ since the cost is here less than for the one for pirating (the legal cost has been removed). On the other hand, pirating will still occur, at least in a first phase for the following reason: if a consumer has a choice of downloading the same good with the same features (quality, ...) either on a legalized forum or on a pirate one, he will obviously choose the legalized forum (except in rare occasions that we choose to ignore). This simply translates the fact that the cost $\frac{\theta}{\hat{\tau}}p$ is always smaller than $\frac{\theta}{\tau}p$ since $\hat{\tau} > \tau$. However, as was assumed above, when searching for a cultural good, the consumer is redirected at random to one of the m forums. In our model, a consumer thus never has a choice between a legalized and a pirate forum: a proportion $\sum_{k=1}^{L} z_k$ of consumers will be redirected to a legalized forum, and thus will have

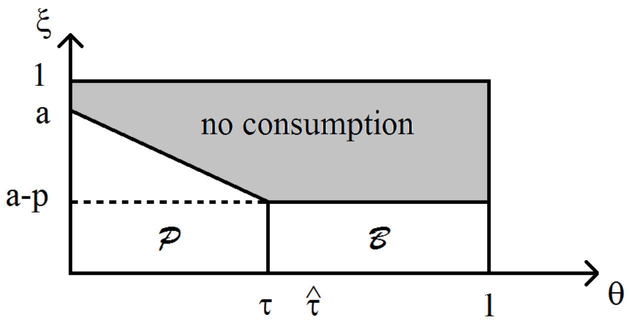

Figure 1: Sets of buyers $\mathcal{B}$, pirates $\mathcal{P}$, and non-consumers for one good with attractiveness a.

a choice between (a) this kind of consumption, (b) buying, and (c) no consumption, while a proportion $1 - \sum_{k=1}^{L} z_k$ of consumers will have a choice between (a) pirating, (b) buying, and (c) no consumption. It might be argued that, in time, the legalized offer will be better known and will attract more consumers. This can be accounted for in our model by adjusting dynamically the probabilities $(z_k)_k$. We will assume that, prior to legalization, all forums have equal probability of being chosen, *i.e.* $z_k = 1/m$ for all k, and that this remains the case for some time after legalization of the first L forums. This assumption is justified by the fact that it will take some time before consumers get to be acquainted to this new form of downloading and to adjust their practice. An extension to this work will examine the case where $z_1, \ldots, z_L$ become in time larger than $z_{L+1}, \ldots, z_m$.

After legalization, we thus have four non void classes of internet users : non-consumers, pirates, downloaders from legalized forums, and buyers. We assume that the distribution μ does not change after legalization: indeed, the preference of internet users as regards the goods, as encoded in the x variable, has no reason to be altered, while their changed behaviour with respect to pirating is already taken into account through the introduction of $\hat{\tau}$.

We now compute the profits of the different actors, prior to, and after legalization.

2.2 Computing profits

We shall make use of the following notations. $\mathcal{B}$ denotes the set of buyers, $\mathcal{P}$ the set of pirates and $\mathcal{C} = \mathcal{B} \cup \mathcal{P}$ the set of active consumers before legalization. After legalization of some forums, the set of active consumers $\overline{\mathcal{C}}$ is split as $\overline{\mathcal{C}} = \overline{\mathcal{B}} \cup \mathcal{D} \cup \overline{\mathcal{P}}$ where $\mathcal{D}$ is the subset of downloaders from legalized forums, $\overline{\mathcal{B}}$ and $\overline{\mathcal{P}}$ are respectively the new sets of buyers and pirates after legalization. The sets $\mathcal{C}, \overline{\mathcal{C}}$ and some of their subsets are pictured on Figures 1 to 3.

Let us express various measures and moments that will be needed in the sequel (for easy reference, the meaning of all the parameters is recalled at the beginning of Section 3). First,

$$\mu(\mathcal{B}) = \sum_i \int_\tau^1 \int_{2i-a_i+p}^{2i+a_i-p} \mu(d\theta, dx), \tag{3}$$

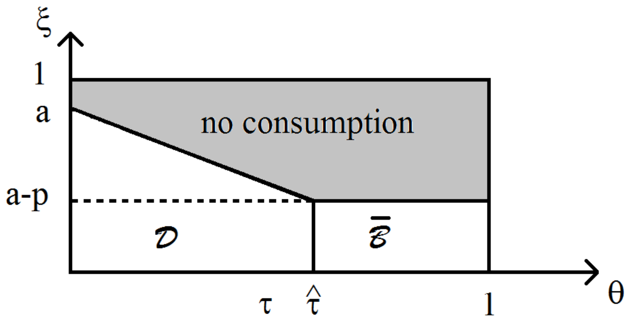

Figure 2: Sets of buyers $\overline{\mathcal{B}}$, downloaders from a legalized forum $\mathcal{D}$, and non-consumers for one good with attractiveness a for individuals that were redirected to a legalized forum.

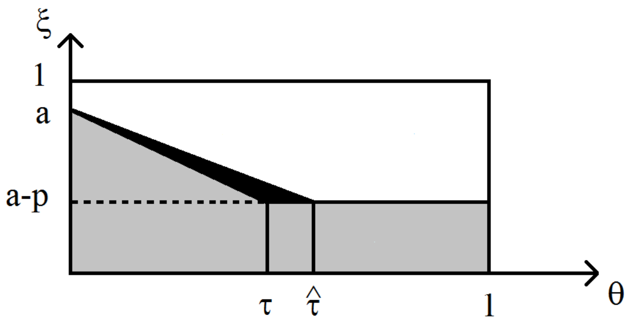

Figure 3: Sets $\mathcal{C}$ (gray) and $\overline{\mathcal{C}}$ in the case of a legalized forum (gray + black) for a particular good with attractiveness a.

with the convention that any integral whose lower bound is not smaller than its upper bound is considered to be 0 (remark also that the integrals with respect to x above are non-overlapping since $a_i + a_{i+1} \leqslant 2 \leqslant 2p+2$). Likewise,

$$\mu(\overline{\mathcal{B}}) = \left(1 - \frac{L}{m}\right)\mu(\mathcal{B}) + \frac{L}{m}\sum_i \int_{\hat{\tau}}^{1} \int_{2i-a_i+p}^{2i+a_i-p} \mu(d\theta, dx), \tag{4}$$

$$\mu(\mathcal{P}) = \sum_i \int_0^{\tau} \int_{2i-a_i+\frac{\theta}{\tau}p}^{2i+a_i-\frac{\theta}{\tau}p} \mu(d\theta, dx), \tag{5}$$

$$\mu(\overline{\mathcal{P}}) = \left(1 - \frac{L}{m}\right)\mu(\mathcal{P}), \tag{6}$$

$$\mu(\mathcal{D}) = \frac{L}{m}\sum_i \int_0^{\hat{\tau}} \int_{2i-a_i+\frac{\theta}{\hat{\tau}}p}^{2i+a_i-\frac{\theta}{\hat{\tau}}p} \mu(d\theta, dx), \tag{7}$$

$$\int_{\mathcal{B}} w\, \mu(d\theta, dx) = \sum_i \int_{\tau}^{1} \int_{2i-a_i+p}^{2i+a_i-p} (a_i - |x-2i|)\, \mu(d\theta, dx), \tag{8}$$

$$\int_{\overline{\mathcal{B}}} w\, \mu(d\theta, dx) = \left(1 - \frac{L}{m}\right)\int_{\mathcal{B}} w\, \mu(d\theta, dx) + \frac{L}{m}\sum_i \int_{\hat{\tau}}^{1} \int_{2i-a_i+p}^{2i+a_i-p} (a_i - |x-2i|)\, \mu(d\theta, dx), \tag{9}$$

$$\int_{\mathcal{P}} w\, \mu(d\theta, dx) = \sum_i \int_0^{\tau} \int_{2i-a_i+\frac{\theta}{\tau}p}^{2i+a_i-\frac{\theta}{\tau}p} (a_i - |x-2i|)\, \mu(d\theta, dx), \tag{10}$$

$$\int_{\overline{\mathcal{P}}} w\, \mu(d\theta, dx) = \left(1 - \frac{L}{m}\right)\int_{\mathcal{P}} w\, \mu(d\theta, dx) \tag{11}$$

$$\int_{\mathcal{D}} w\, \mu(d\theta, dx) = \frac{L}{m}\sum_i \int_0^{\hat{\tau}} \int_{2i-a_i+\frac{\theta}{\hat{\tau}}p}^{2i+a_i-\frac{\theta}{\hat{\tau}}p} (a_i - |x-2i|)\, \mu(d\theta, dx). \tag{12}$$

The profit P_I of industry prior to legalizing forums may be expressed as follows:

$$\begin{aligned} P_I &= \int p_u(\theta, \xi)\, \mu(d\theta, dx) - K_I \\ &= \int_{\mathcal{B}} [p + \sigma(w-p)]\ \mu(d\theta, dx) + \int_{\mathcal{P}} \sigma w\, \mu(d\theta, dx) - K_I \\ &= (1-\sigma)p\mu(\mathcal{B}) + \sigma \int_{\mathcal{C}} w\, \mu(d\theta, dx) - K_I \end{aligned} \tag{13}$$

where p_u is the profit per individual and K_I denotes the global cost incurred by industry to produce and promote all cultural goods.

Post-legalization industry profit reads:

$$\begin{aligned} \overline{P_I} &= \int_{\overline{\mathcal{B}}} [p + \sigma(w-p)]\ \mu(d\theta, dx) + \int_{\mathcal{D}\cup\overline{\mathcal{P}}} \sigma w\, \mu(d\theta, dx) + r\, \mu(\mathcal{D}) - K_I \\ &= (1-\sigma)p\mu(\overline{\mathcal{B}}) + \sigma \int_{\overline{\mathcal{C}}} w\, \mu(d\theta, dx) + r\, \mu(\mathcal{D}) - K_I. \end{aligned} \tag{14}$$

The incurred gain or loss induced by legalization for industry is

$$\overline{P_I} - P_I = (1-\sigma)p\left(\mu(\overline{\mathcal{B}}) - \mu(\mathcal{B})\right) + \sigma\left(\int_{\overline{\mathcal{C}}} w\,\mu(d\theta, dx) - \int_{\mathcal{C}} w\,\mu(d\theta, dx)\right) + r\,\mu(\mathcal{D}). \qquad (15)$$

Using (3) to (12), one computes:

$$\begin{aligned}\overline{P_I} - P_I &= -\frac{L}{m}\sum_i \int_{\tau}^{\hat\tau}\int_{2i-a_i+p}^{2i+a_i-p} \left[p(1-\sigma) + \sigma(a_i - |x-2i|)\right]\mu(d\theta, dx) \\ &+ \sigma\,\frac{L}{m}\left(\sum_i \int_0^{\hat\tau}\int_{2i-a_i+\frac{\theta}{\hat\tau}p}^{2i+a_i-\frac{\theta}{\hat\tau}p} (a_i - |x-2i|)\,\mu(d\theta, dx) - \sum_i \int_0^{\tau}\int_{2i-a_i+\frac{\theta}{\tau}p}^{2i+a_i-\frac{\theta}{\tau}p} (a_i - |x-2i|)\,\mu(d\theta, dx)\right) \\ &+ r\,\mu(\mathcal{D}). \qquad (16)\end{aligned}$$

$\overline{P_I} - P_I$ is a sum of three terms. The one on the first line of (16) is non-positive since $\hat\tau > \tau$. The one on the second line is also typically non-positive. Thus, industry profit will increase only if r is large enough, which is intuitively obvious. Note that in the limiting case where $\hat\tau = \tau$, industry profit increases by the amount of $r\,\mu(\mathcal{D})$, as is expected.

Prior to legalization, profit of forum k, $k = 1, \ldots, m$, is equal to

$$P_{F_k} = \frac{1}{m}\int_{\mathcal{P}} \tilde p\, d\mu = \frac{\tilde p}{m}\mu(\mathcal{P}).$$

For $k = 1, \ldots, L$, profit after legalization becomes (recall that we assume equal probabilities z_k, $k = 1, \ldots, m$)

$$\overline{P_{F_k}} = \frac{1}{L}\int_{\mathcal{D}} \hat p\, d\mu - \frac{r}{L}\int_{\mathcal{D}} d\mu = \frac{\hat p - r}{L}\mu(\mathcal{D}),$$

while, for $k = L+1, \ldots, m$, the new profit reads

$$\overline{P_{F_k}} = \frac{1}{m-L}\int_{\overline{\mathcal{P}}} \tilde p\, d\mu = \frac{\tilde p}{m-L}\mu(\overline{\mathcal{P}}).$$

Thus, the shift in profit is

$$\overline{P_{F_k}} - P_{F_k} = \frac{\hat p - r}{L}\mu(\mathcal{D}) - \frac{\tilde p}{m}\mu(\mathcal{P}) \qquad (17)$$

for $k = 1, \ldots, L$, and

$$\overline{P_{F_k}} - P_{F_k} = \tilde p\left(\frac{\mu(\overline{\mathcal{P}})}{m-L} - \frac{\mu(\mathcal{P})}{m}\right) = 0$$

for $k = L+1, \ldots, m$, where we have used (6). The fact that legalization has no impact on the profit of forums that choose to remain illegal is simply a consequence of our assumption that the z_k all remain equal after legalization.

Legalizing forums will be profitable to industry whenever (15) is positive, and to forums as soon as (17) is positive. This provides a range of values acceptable to both parties for the contribution r to be paid by forums to industry:

$$-\frac{1}{\mu(\mathcal{D})}\left((1-\sigma)p\left(\mu(\overline{\mathcal{B}}) - \mu(\mathcal{B})\right) + \sigma\left(\int_{\overline{\mathcal{C}}} w\,\mu(d\theta, dx) - \int_{\mathcal{C}} w\,\mu(d\theta, dx)\right)\right) \leqslant r \leqslant \hat p - \tilde p\,\frac{L}{m}\,\frac{\mu(\mathcal{P})}{\mu(\mathcal{D})}. \qquad (18)$$

Inequalities (18) constitute our master formula, from which all the analysis below derives.

At first sight, it seems that the upper bound on r in (18) is a decreasing function of L, which would be rather counter-intuitive. However, using (5) and (7), one sees that

$$\hat{p} - \tilde{p}\,\frac{L}{m}\,\frac{\mu(\mathcal{P})}{\mu(\mathcal{D})} = \hat{p} - \tilde{p}\,\frac{\sum_i \int_0^{\tau} \int_{2i-a_i+\frac{\theta}{\tau}p}^{2i+a_i-\frac{\theta}{\tau}p} \mu(d\theta, dx)}{\sum_i \int_0^{\hat{\tau}} \int_{2i-a_i+\frac{\theta}{\hat{\tau}}p}^{2i+a_i-\frac{\theta}{\hat{\tau}}p} \mu(d\theta, dx)}, \tag{19}$$

which shows that it depends neither on L nor on m. Inspection of (7) and (16) reveals that the lower bound also does not depend on L nor on m.

Inequalities (18) are meaningful only if they define a non-empty interval in $\mathbb{R}^+$, *i.e.* if

$$\max\left(0, (1-\sigma)p\left(\mu(\mathcal{B}) - \mu(\overline{\mathcal{B}})\right) + \sigma\left(\left(\int_{\mathcal{C}} w\,\mu(d\theta, dx) - \int_{\overline{\mathcal{C}}} w\,\mu(d\theta, dx)\right)\right)\right) < \hat{p}\,\mu(\mathcal{D}) - \tilde{p}\,\frac{L}{m}\,\mu(\mathcal{P}). \tag{20}$$

One checks that, in the limiting case $\hat{\tau} = \tau$, $\frac{\mu(\mathcal{P})}{m} = \frac{\mu(\mathcal{D})}{L}$ so that (18) becomes:

$$0 \leqslant r \leqslant \hat{p} - \tilde{p}.$$

In other words, would there be no legal costs, the upper bound on the contribution r would simply be the profit increase of legalized forums, a result which seems rather natural.

In the next section, we consider the special case where the measure μ is separable, a property that is satisfied in Configurations 1, 2 described above. In this situation, the expression of the admissible range for r as given by (18) simplifies significantly.

2.3 Separable scenario

In the case where $\mu(d\theta, dx) = h(x)g(\theta)dxd\theta$, and thus, in particular for Configurations 1 and 2, the various integrals computed above, and, as a consequence, (18), take a simpler form. Note first that we do not loose any generality in supposing then that $g = \mathbf{1}_{\{\theta\in[0,1]\}}$, since all we are interested in is the proportion of individuals with $\theta < \tau$, where $0 \leqslant \tau \leqslant 1$. Replacing $\mu(d\theta, dx)$ by $h(x)dxd\theta$ and denoting H a primitive of h, one computes

$$\begin{aligned}
\mu(\mathcal{P}) &= \sum_i \int_0^{\tau} \int_{2i-a_i+\frac{\theta}{\tau}p}^{2i+a_i-\frac{\theta}{\tau}p} h(x)\,dx\,d\theta \\
&= \sum_i \int_0^{\tau} \left(H(2i+a_i-\frac{\theta}{\tau}p) - H(2i-a_i+\frac{\theta}{\tau}p)\right)\,d\theta \\
&= \tau \sum_i \int_0^{1} \left(H(2i+a_i-\theta p) - H(2i-a_i+\theta p)\right)\,d\theta \\
&= \tau IS(p),
\end{aligned}$$

where we have set $IS(p) := \sum_i \int_0^1 \left(H(2i+a_i-\theta p) - H(2i-a_i+\theta p)\right)\,d\theta$. The same computations lead to

$$\mu(\mathcal{D}) = \frac{L}{m}\,\hat{\tau}\,IS(p),$$

so that the upper bound in (18) now just reads $\hat{p} - \frac{\tau}{\hat{\tau}}\tilde{p}$. The lower bound does not simplify so drastically, but one can compute

$$\begin{aligned}\mu(\mathcal{B}) &= \sum_i \int_\tau^1 \int_{2i-a_i+p}^{2i+a_i-p} h(x)\,dx\,d\theta \\ &= (1-\tau)\sum_i (H(2i+a_i-p) - H(2i-a_i+p)) \\ &= (1-\tau)S(p),\end{aligned}$$

where $S(p) := \sum_i (H(2i+a_i-p) - H(2i-a_i+p))$. Likewise,

$$\mu(\overline{\mathcal{B}}) = \left(\left(1-\frac{L}{m}\right)(1-\tau) + \frac{L}{m}(1-\hat{\tau})\right) S(p).$$

Let $\check{H}_i$ denote a primitive of the function $x \mapsto (a_i - |x-2i|)h(x)$. Then

$$\begin{aligned}\int_{\mathcal{B}} w\mu(d\theta, dx) &= \sum_i \int_\tau^1 \int_{2i-a_i+p}^{2i+a_i-p} (a_i - |x-2i|)h(x)\,dx\,d\theta \\ &= (1-\tau)\sum_i \left(\check{H}_i(2i+a_i-p) - \check{H}_i(2i-a_i+p)\right) \\ &= (1-\tau)\check{S}(p),\end{aligned}$$

where $\check{S}(p) = \sum_i \left(\check{H}_i(2i+a_i-p) - \check{H}_i(2i-a_i+p)\right)$. One obtains in the same way

$$\int_{\overline{\mathcal{B}}} w\mu(d\theta, dx) = \left(\left(1-\frac{L}{m}\right)(1-\tau) + \frac{L}{m}(1-\hat{\tau})\right) \check{S}(p),$$

$$\begin{aligned}\int_{\mathcal{P}} w\mu(d\theta, dx) &= \sum_i \int_0^\tau \int_{2i-a_i+\frac{\theta}{\tau}p}^{2i+a_i-\frac{\theta}{\tau}p} (a_i - |x-2i|)h(x)\,dx\,d\theta \\ &= \tau\sum_i \int_0^1 \left(\check{H}_i(2i+a_i-\theta p) - \check{H}_i(2i-a_i+\theta p)\right)\,d\theta \\ &= \tau\widecheck{IS}(p),\end{aligned}$$

where we have set $\widecheck{IS}(p) := \sum_i \int_0^1 \left(\check{H}_i(2i+a_i-\theta p) - \check{H}_i(2i-a_i+\theta p)\right)\,d\theta$, and finally

$$\int_{\mathcal{D}} w\mu(d\theta, dx) = \frac{L}{m}\hat{\tau}\widecheck{IS}(p).$$

Gathering the above expressions, one sees that, in the separable case, (18) reads

$$\left(1-\frac{\tau}{\hat{\tau}}\right)\frac{(1-\sigma)S(p) + \sigma(\check{S}(p) - \widecheck{IS}(p))}{IS(p)} \leqslant r \leqslant \hat{p} - \frac{\tau}{\hat{\tau}}\tilde{p}. \tag{21}$$

These bounds will be used in the next section in the cases of Configurations 1 and 2.

We note the remarkable fact that the upper bound does not depend on μ: when the reluctance for pirating is independent from the popularity of the goods, the distribution of consumers has no effect on fixing the upper limit for the contribution r.

3 Parameter estimation

In order to put our model to work and obtain meaningful numerical results, we need to set its parameters to realistic values. For the reader's convenience, we recall here the full set of the twelve parameters defined above, with some comments related to their practical use:

1. c : constant proportional to the total number internet users, that is, buyers + pirate downloaders + non-consumers. One easily sees from (21) that the bounds do not depend on c in the separable case, and thus we may ignore it when dealing with Configurations 1 and 2. It does however play an important role in Configuration 3, as we shall see below.

2. n: number of cultural goods of a given type (music, movies, series, ...) that generate "significant" traffic on the internet. In separable scenarios, the bounds do not depend on n.

3. a_i: attractiveness of good g_i. It does not seem easy at this time to estimate the full collection of a_i, and we will assume that $a_i = a$ for all i, that is, attractiveness does not depend on the good. Although a simplifying assumption, this still allows us to get an idea of the feasibility of the proposal; in separable configurations, only the lower bound in (18) depends on a.

4. p: industry price of a good. Again, in separable configurations, only the lower bound depends on p;

5. τ: parameter measuring the total cost of pirating, that is, the comfort cost plus the legal cost. In separable scenarios, τ is just the proportion of pirate downloaders. In general, this is still true in most cases of interests provided n is very large;

6. $\hat{\tau}$: parameter measuring the comfort cost of pirating, that is, the reluctance to download from legalized forums;

7. σ: factor that governs the industry share of the proportion of what consumers spend for products related to the good they either buy or download illegally. Only the lower bound in (18) depends on σ, and it does so in an affine way;

8. m: total number of pirate forums before legalization, or of pirate +legalized forums after legalization. Formula (18) shows that the bounds do not depend on m;

9. L: number of legalized forums. Formula (18) shows that the bounds do not depend on L;

10. $\tilde{p}$: remuneration of pirate forums. Only the upper bound in (18) depends on $\tilde{p}$;

11. $\hat{p}$: remuneration of legalized forums. Only the upper bound in (18) depends on $\hat{p}$;

12. γ: factor characterizing the speed of decay of the popularity of goods in Configurations 2 and 3.

Recall finally that r denotes the retribution that will be asked to illegal forums in exchange for legalization.

As one can see, the set of the ten parameters that needs to be estimated is $\{c, n, a, p, \tau, \hat{\tau}, \sigma, \tilde{p}, \hat{p}, \gamma\}$ (although not all parameters are relevant in each situation). This is not an easy task, as most of them are not directly observable. A full calibration requires a work of its own and it will be the topic of a subsequent study. For the current work, we adopt the following strategy: for each parameter, we set a sensible default value, estimated from various sources as explained below. In order to appreciate the impact of each parameter and what would happen if our estimations were significantly wrong, we let each parameter vary in its natural range while keeping the other ones fixed to their default value. That is, we draw below curves describing the evolutions of the lower and upper bounds in (18) separately as functions of each of the parameters. This allows us to recognize for which values proportional retribution will make economical sense.

We now comment on the way each default value is set. We mainly use the following sources:

- [21] is a study made by the Hadopi in June 2013. One aim of the Hadopi was to monitor the usage of Internet in France as concerns the circulation of cultural goods. In that view, the Hadopi has set up a large scale investigation to obtain data based on consumption logs filled by voluntary users. More precisely, users had to record on a daily basis their digital consumption of music, movies, series, video games and books/comics over a period of seven consecutive days. The investigation involved 4,740 individuals.

- [22] is another study performed by the Hadopi in September and October 2012, based on an investigation on 1,530 internet users of age at least fifteen chosen from the Survey Sampling International panel.

- [12] discusses the findings of a representative survey of 1,500 Dutch internet users in 2009, who were asked about their behaviour in relation to file sharing of music, films and games with the aim of identifying the economic and cultural effects of file sharing.

- [23] is a vast survey of practices regarding copying and downloading conducted in August and September 2011 based on a random phone survey of 1,000 Germans and 2,303 Americans.

- the web facilities *CEG TEK International* and *Torrent Freak*.

Based on these sources, we estimate the default values as follows:

1. c: as mentioned above, we need only consider c in the frame of Configuration 3. A Taylor expansion of the integral of $\mu(dx, d\theta)$ over $[0, 2n+1] \times [0, 1]$ when n tends to infinity, with $\gamma = 0.41$ (see item 8 below), yields that the total mass, that is, the number of internet users, is of the order of $2.53\, c\, n^{3/5}$. With $n = 500$ (see item 2 below), this amounts to $105\, c$. The number of internet users in France is estimated to be around 43.2 million (for comparison, p. 81 of [12] gives an estimate of 10.7 million users in the Netherlands). We thus take $c \approx 400,000$ as a reasonable value. Note that one could consider instead the total estimated number of internet users throughout the world - which is of the order of 2,900 million - to set the value of c. A look at the graphs in the next section shows however that both bounds quickly reach an asymptotic value well before $c = 400,000$.

2. n: while it is hard to set a definite limit on the numbers of cultural goods of a given type (music, movies, series, ...) that generate significant traffic, a look at `http://www.cegtek.com/afm/` reveals that 500 is probably a sensible figure.

3. a, p, τ: we estimate jointly these parameters by taking advantage of the fact that the measures μ of the set of non-consumers, buyers, and pirates are all functions of the triplet (a, p, τ) and of no other parameters. More precisely, writing $\mathcal{NC}$ for the set of non-consumers, easy computations reveal that, in Configuration 1 with $c = 1$ (recall Figure 1),

$$\begin{cases} \mu(\mathcal{NC}) & = (1-\tau)((1-a+p) + \tau(1-a) + \frac{p\tau}{2}, \\ \mu(\mathcal{B}) & = (1-\tau)(a-p), \\ \mu(\mathcal{P}) & = \tau(a-p) + \frac{p\tau}{2}. \end{cases} \tag{22}$$

This is a set of three non-linear equations in the unknowns (a, p, τ), which have multiple solutions. Relevant ones may then be searched for once $\mu(\mathcal{NC}), \mu(\mathcal{B})$ and $\mu(\mathcal{P})$ are known[5]. The figures on p. 26 of [21] are useful for this purpose. They indicate that sensible (normalized) values, e.g. in the case of music, are $\mu(\mathcal{NC}) = 0.48, \mu(\mathcal{B}) = 0.66 \times 0.52 = 0.3432$, and $\mu(\mathcal{P}) = 0.34 \times 0.52 = 0.1768$. Slightly different figures would be obtained if one considered other types of cultural good, or all cultural goods. For instance, $\mu(\mathcal{NC})$ drops down to 0.32 if one considers all types of cultural goods. We neglect these variations as our aim here is to obtain a general behaviour, and also because precision at this stage is not meaningful since several parameters are unreachable at this time. We note also that our estimates are comparable with the ones obtained for various countries in [23].

To decide which solutions of (22) are most relevant, we use two more facts from our sources: first, p. 106 of [12] hints that the value of $a - \xi$ averaged over all consumers should be close to $0.6p$. Since, in Configuration 1, this average is simply $a - \frac{1}{2}$, we expect that a should not be too far from $0.5 + 0.6p$. This estimate is also consistent with figures given p. 78,79 and 82 of [12] and roughly consistent with the ones on p. 59 of [23] and with findings in [8]. Second, all our first four sources indicate that τ should be somewhere between 0.2 and 0.7, depending for instance on the type of cultural good - music, films, series- (see for instance p. 26 to 29 of [21] or p.19 of [23]). Solving (22) with the constraints that $a \in [0.5, 1], p \in [0, 0.5]$ and $\tau \in [0.2, 0.7]$ yields the solution $(a = 0.72, p = 0.23, \tau = 0.29)$.

A value of 0.29 for τ might seem underestimated. However, one must recall that, in our model, individuals are divided into three non-overlapping classes: non-consumers, buyers, and pirates. Thus, the value of τ represents the percentage of individuals that will typically both consume digitally a cultural good and do it in an illegal way.

Note that the values of (a, p, τ) above have been derived for Configuration 1. Clearly, Formulas (22) are different in other configurations. We will however stick to these values as sensible default ones in all cases.

4. $\hat{\tau}$: the figures on p.23 of [22] suggest that setting $\hat{\tau} = 2\tau = 0.58$ is a relevant choice.

5. σ: this parameter is difficult to estimate. However, since only the lower bound in (18) depends on σ, and since it does so simply in an affine way, its impact is not crucial. Setting $\sigma = 0.1$ does not seem to be unreasonable.

[5]Note that knowledge of these three quantities is not enough to determine the bounds in (18) since moments of μ are also required.

6. $\tilde{p}$: we have found no reliable source to set the value of this parameter. We have settled on $\tilde{p} = 0.1p = 0.023$.

7. $\hat{p}$: this is even more difficult to estimate, as it refers to a hypothetical situation. We believe that taking $\hat{p} = 0.2p = 0.046$ may be appropriate. See Section 4.3 for a discussion on this aspect.

8. γ: the decay rate may be roughly estimated for instance from the figures given at `http://torrentfreak.com/game-of-thrones-most-pirated-tv-show-of-2013-131225/`, which exhibit a rather good linear fit between the logarithm of the rank and the one of the number of BitTorrent downloads for series, with slope close to $\gamma = 0.41$ (see Figure 4). As our experiments show, the bounds are only slightly sensitive to a change of γ.

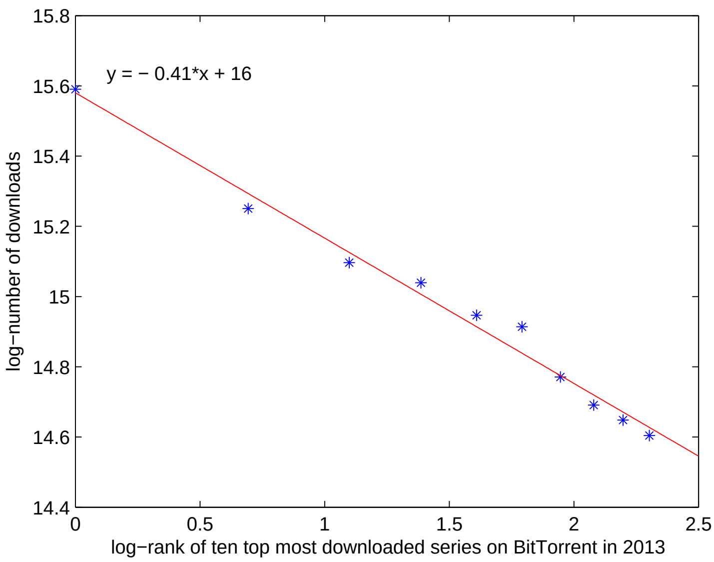

Figure 4: Log-log plot of downloads versus rank for the ten top most downloaded series on BitTorrent in 2013.

To sum up, we will use the parameters displayed in Table 1 (not all parameters are relevant in each case). Obviously, these values are only approximations, as only indirect sources are available at this time to estimate them. In addition, we have several times taken average between values that would fit either for music, films, or series, rather that performing distinct estimations for each type of cultural good: while the latter procedure would clearly be more correct, we lack at this time sufficient data to go to this level of detail. We believe nevertheless that our default values provide sensible starting points. This statement is supported by the fact that the values we have derived roughly fit the situation in many countries including France, Germany, the Netherlands, the UK and the USA. Finally, we note that inspection of (18) shows that both our bounds are smooth functions of the parameters: as a consequence, small inaccuracies in their estimation do not alter qualitatively our results.

c	n	σ	a	p	$\tilde{p}$	$\hat{p}$	τ	$\hat{\tau}$	γ
400,000	500	0.1	0.72	0.23	0.023	0.046	0.29	0.58	0.41

Table 1: Default values used for the parameters in numerical experiments.

4 Numerical computations

Equipped with the estimates of the previous section, we now proceed to compute the bounds of (18) on the three configurations described earlier. As explained above, we let the parameters vary one at a time, letting all the other ones taking the default values indicated in Table 1. This procedure yields curves, or, rather, strips, characterizing the possible win-win situations when they exist, that is, the cases where both industry and forums benefit from legalization, along with the corresponding admissible values for the contribution r.

4.1 Configuration 1

Recall that, for Configuration 1, the bounds in (21) do not depend on c nor on n. Obviously, γ is also irrelevant here, which leaves us with seven parameters to be varied. Figures 5 to 11 display graphs showing the evolution of the bounds as functions of the remaining parameters, that is, $\sigma, a, p, \tilde{p}, \hat{p}, \tau, \hat{\tau}$.

The figures show that, when μ is uniform, the estimated values for the parameters are such that the proposition of the Hadopi is not economically feasible. A retribution r would exist only if:

- σ was larger than 0.55, which means that Industry would gather more than half of what consumers spend for cultural products they have bought or pirated (aside from p). As mentioned above, estimating σ is a difficult task, and our default value 0.1 may well be underestimated. However, a value of 0.55 seems unrealistic.
- a was smaller that 0.3, a value that is quite far from our estimate of 0.72 and rather close to the default one of 0.23 of p. Symmetrically, a value of r would exist if p was larger that 0.67 and close to the default value of a. In other words, Industry price has to be large enough for the proposition to be applicable.
- The earning of legalized forums $\hat{p}$ was larger that 0.1, which means that it would have to be multiplied by four after legalization assuming that $\tilde{p} = 0.023$.
- the proportion of pirates τ was larger than 0.5. This might be not totally unrealistic.
- The proportion of users downloading from legalized forum was smaller than 0.33: at least in a first phase, such a figure would be possible. However, in time, one would expect that $\hat{\tau}$ would become significantly larger than $\tau = 0.29$.

Note also that no reasonable values of $\tilde{p}$ leads to a feasible scenario.

Summing up, we see that, for a retribution r to exist in Configuration 1, the values of the parameters would have to be significantly different from the default ones. Since some of our estimates are rather imprecise, it may be the case that the proposition of Hadopi is feasible under a uniform distribution, but it is more probable that it is not so.

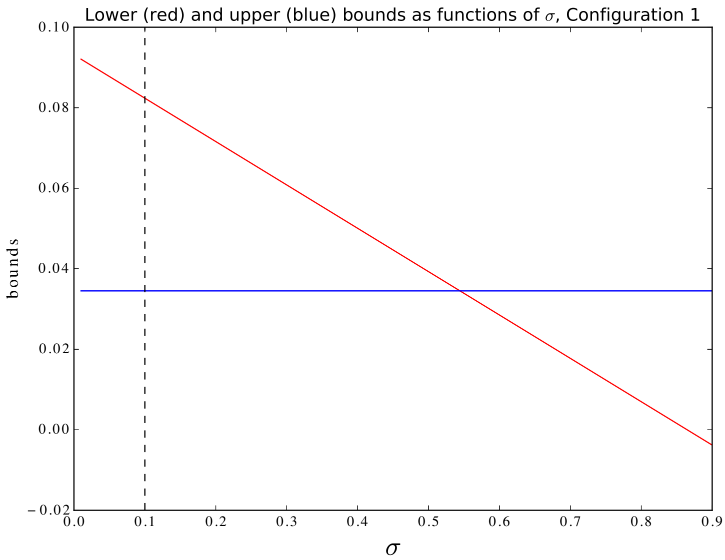

Figure 5: Lower bound (red) and upper bound (blue) in (21) as functions of σ in Configuration 1. The vertical black dashed line indicates the default value.

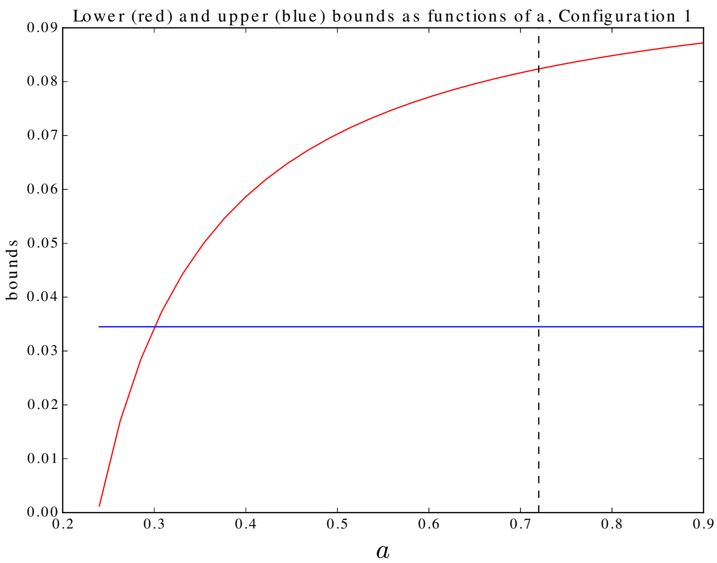

Figure 6: Lower bound (red) and upper bound (blue) in (21) as functions of a in Configuration 1. The vertical black dashed line indicates the default value.

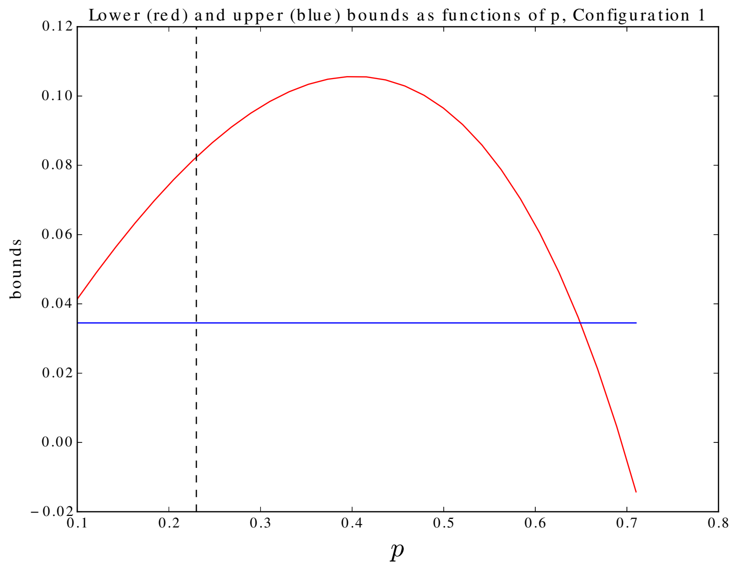

Figure 7: Lower bound (red) and upper bound (blue) in (21) as functions of p in Configuration 1. The vertical black dashed line indicates the default value.

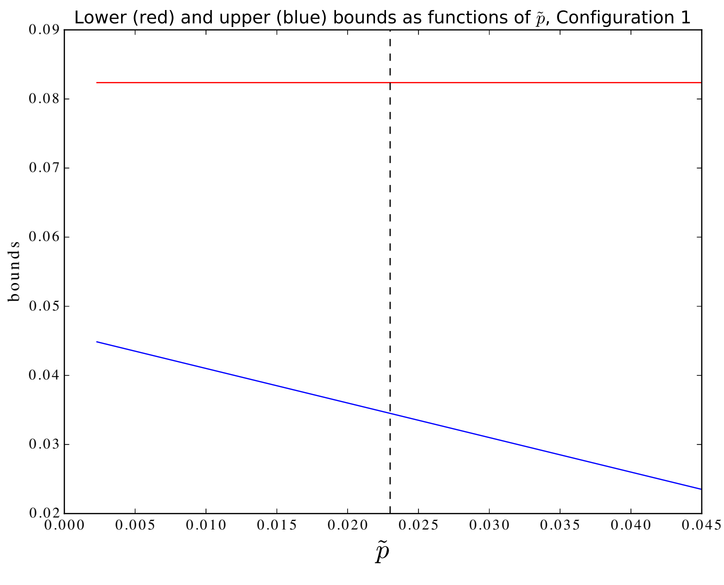

Figure 8: Lower bound (red) and upper bound (blue) in (21) as functions of $\tilde{p}$ in Configuration 1. The vertical black dashed line indicates the default value.

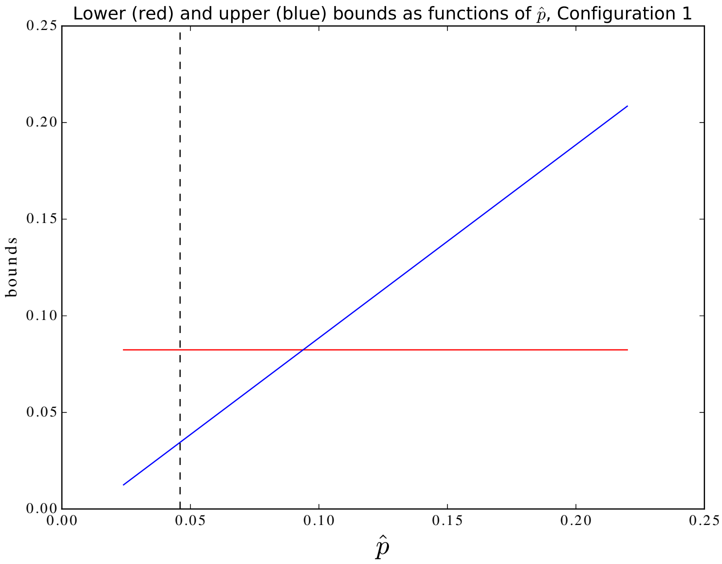

Figure 9: Lower bound (red) and upper bound (blue) in (21) as functions of $\hat{p}$ in Configuration 1. The vertical black dashed line indicates the default value.

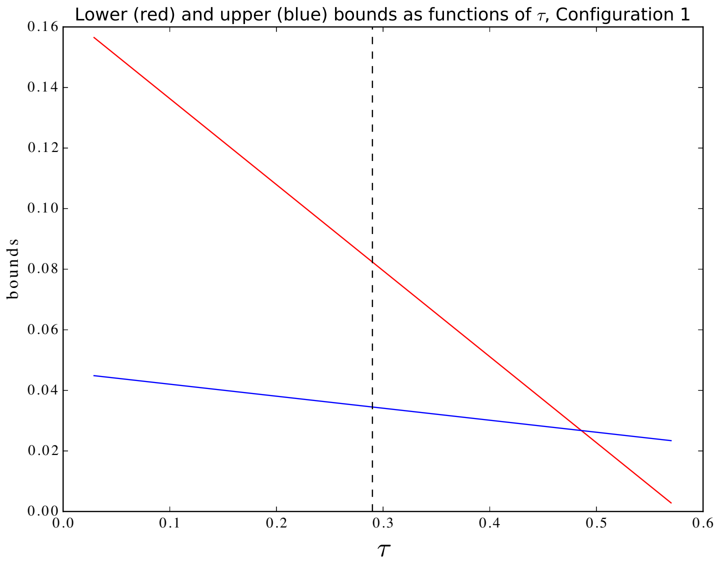

Figure 10: Lower bound (red) and upper bound (blue) in (21) as functions of τ in Configuration 1. The vertical black dashed line indicates the default value.

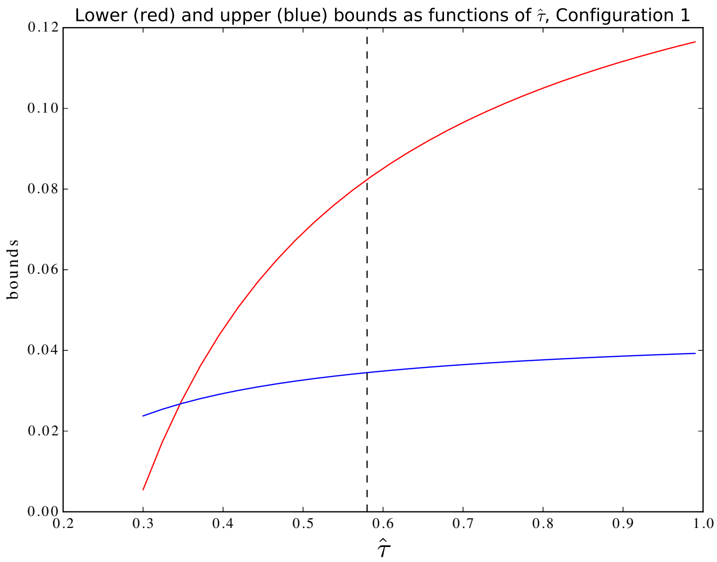

Figure 11: Lower bound (red) and upper bound (blue) in (21) as functions of $\hat{\tau}$ in Configuration 1. The vertical black dashed line indicates the default value.

4.2 Configuration 2

As already noted, (21) shows that the upper bound is the same for Configurations 1 and 2. Our numerical experiments further indicate that the lower bounds in Configuration 2 are very close to the ones in Configuration 1 when all the parameters $\sigma, a, p, \tilde{p}, \hat{p}, \tau, \hat{\tau}$ are varied in their admissible ranges: the difference is of the order of 10^{-5}. For this reason, we do not present the graphs corresponding to these parameters.

Maybe more surprisingly, the lower bound in Configuration 2, which depends on γ, has negligible discrepancy with the lower bound in Configuration 1 with default parameters when $\gamma < 1$. The difference increases with γ, but it only reaches 0.001 when $\gamma = 2$. Figure 12 displays the bounds as functions of γ in Configuration 2 (of course, only the lower bound varies with γ).

These facts may be an indication that our model is robust with respect to a change of distribution of consumers in the class of separable measures. It also reveals that, if the behaviour of internet users as regards pirating does not depend on the popularity of goods, the proposition of the Hadopi is likely to be economically infeasible.

4.3 Configuration 3

Assuming the same form for the distribution in x as in Configuration 2, we set in this section

$$\mu(d\theta, dx) = \left(1 + \frac{c}{1 + x^\gamma}\left(\frac{1}{2} - \theta\right)\right) d\theta dx.$$

We draw on Figure 13 an example of the ruled surface which is the graph of the density of the measure. Note that, in this configuration, the bounds in (18) do depend on all our ten parameters.

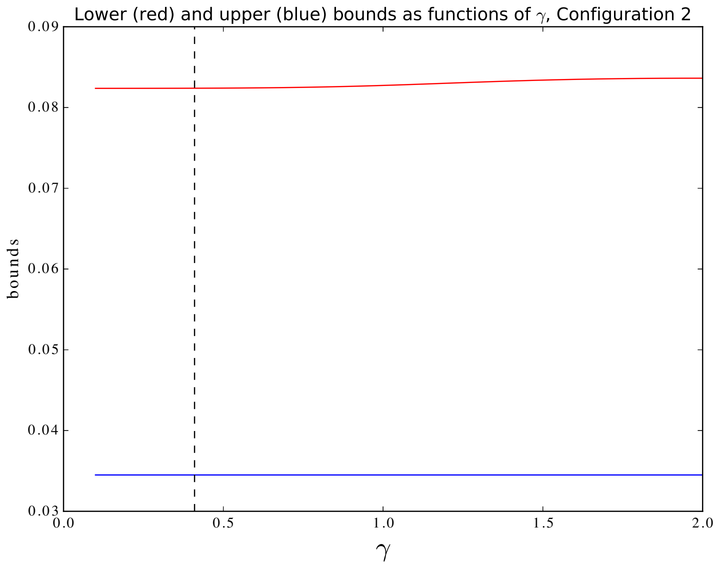

Figure 12: Lower bound (red) and upper bound (blue) in (21) as functions of γ in Configuration 2. The vertical black dashed line indicates the default value.

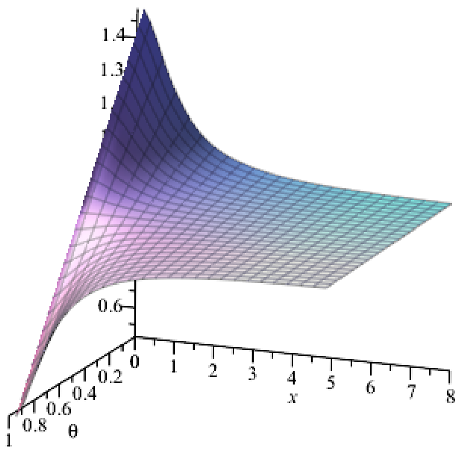

Figure 13: The density of the measure in Configuration 3 when $\gamma = 2$.

We display our results on Figures 14 to 23. One sees that, in this configuration, the default values lead to a non-empty interval for the retribution r, so that the proposition of the Hadopi is indeed economically feasible: r would be of the order of 0.025, that is, roughly comparable to the remuneration of pirate forums. At first sight, this result might seems counter-intuitive: pirate forums earn $\tilde{p}$ per download, while legalized ones earn $\hat{p} = 2\tilde{p}$ with our default values. With a value of $r \approx \tilde{p}$, one has that $\hat{p} - r \approx \tilde{p}$, so that there is nothing to gain for forums through legalization: the increase in profit serves as the retribution to Industry. However, one must recall that, after legalization, these forums will attract (significantly) more users since $\hat{\tau} = 2\tau$, so that their total earning will become much larger. In other words, the net profit per download would not change, but the volume will increase. The fact that the per-download increase in profit of legalized forums will essentially go to Industry is noteworthy: it reveals the mechanism through which the proposition of the Hadopi will function.

Our graphs indicate that these results are rather robust with respect to changes in the parameters: in most cases, the lower bound remains smaller than the upper bound throughout the whole admissible range of the parameter. The exceptions are τ and the couple $(\tilde{p}, \hat{p})$: too small a proportion of pirates prior to legalization (that is, in the current situation) would not allow for r to exist. Given that our estimate for τ is probably a lower bound, this is not likely to be the case. The other problematic situation is when the increase in the earnings of forums before and after legalization is too small: legalization should allow the profit per download to be multiplied by at least 1.5. While this does not seem unreasonable, verifying whether this assumption is realistic should be the topic of further studies.

To sum up, even if our estimates are strongly in error (except for an overestimation of $\frac{\hat{p}}{\tilde{p}}$, which, admittedly, is hard to evaluate), a range of possible values for r does exist. As a consequence, the proposition of the Hadopi seems to make economical sense if one believes that more popular goods are more prone to pirating.

5 Discussion and extensions

The numerical results of the preceding section show that proportional retribution may be economically feasible provided it is true that the tendency to use pirate download depends on the popularity of goods. More precisely, we need to assume that popular goods are more prone to pirating. This does not seem unreasonable, and is probably closer to reality that the separable scenarios of Configurations 1 and 2 where internet users do not adjust their behaviour based on what they wish to consume.

In this frame, our findings show that a retribution in exchange of legalization, that would be of the order of the current profit of pirate forums, would allow to augment the earnings of both Industry and legalized forums, under the two conditions that the proportion of pirates is large enough (which seems largely true) and that the increase in the amount of money that forums will make from advertisement when becoming legal is sufficient. We note the interesting fact that this increase will essentially be used for the retribution to Industry, while the gains for forums will come from the fact that they will attract a larger number of consumers.

Economic feasibility is of course but one aspect of the proposition put forward by Hadopi. Legal and technical matters are of equal importance. Finding an adequate legal frame is a challenge, specially since many rights, including fundamental ones such as intellectual property, are involved. As mentioned above, the proposition can be seen as an instance of global licence analogous to [15, 16] with the twist that this fee is financed by advertisement. Another way to see it is to realize that it proposes a sort of low-

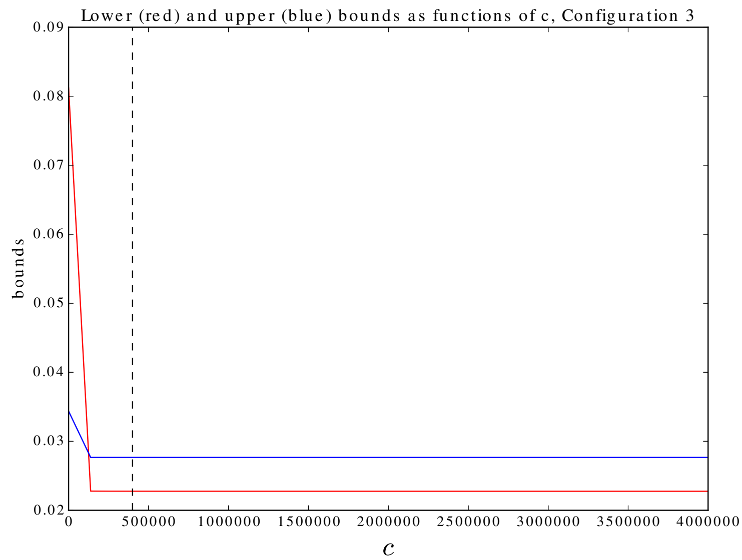

Figure 14: Lower bound (red) and upper bound (blue) in (18) as functions of c in Configuration 3. The vertical black dashed line indicates the default value.

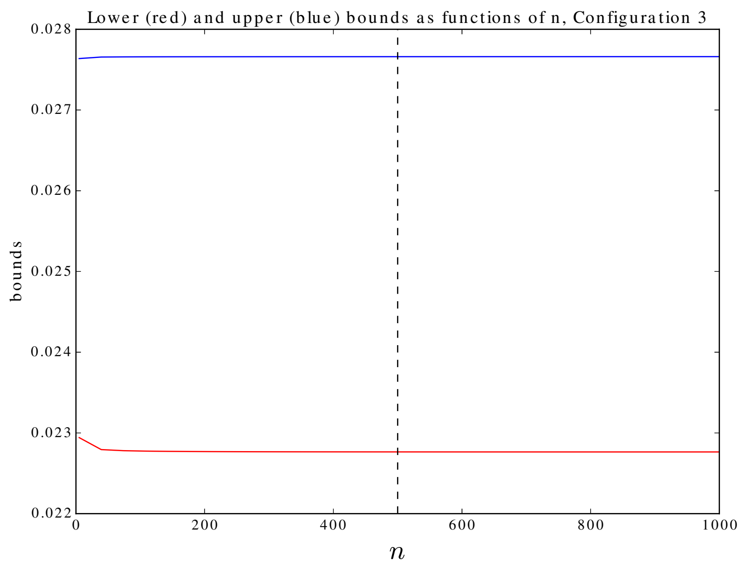

Figure 15: Lower bound (red) and upper bound (blue) in (18) as functions of n in Configuration 3. The vertical black dashed line indicates the default value.

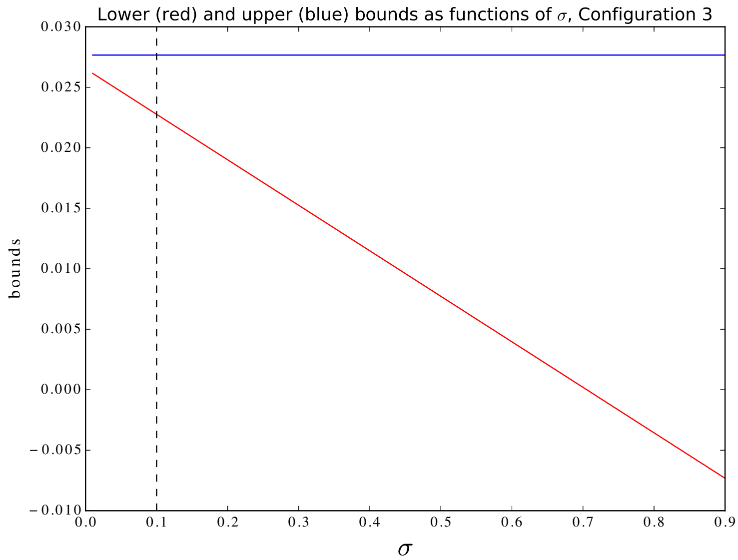

Figure 16: Lower bound (red) and upper bound (blue) in (18) as functions of σ in Configuration 3. The vertical black dashed line indicates the default value.

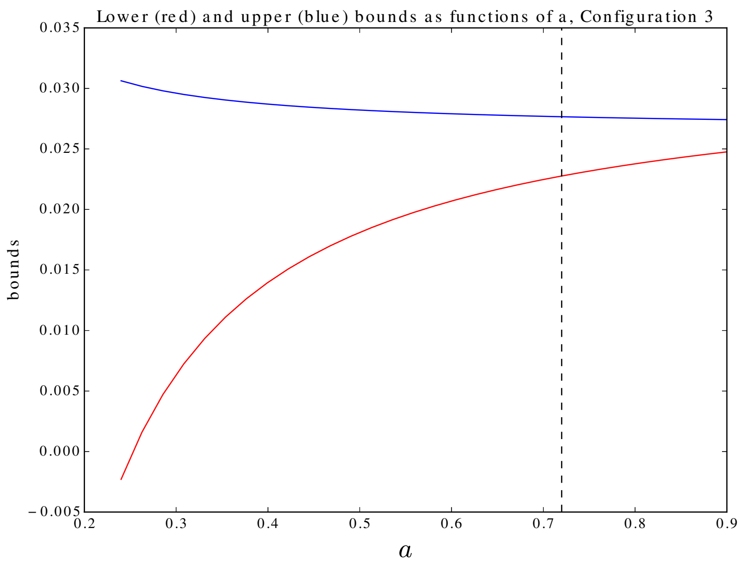

Figure 17: Lower bound (red) and upper bound (blue) in (18) as functions of a in Configuration 3. The vertical black dashed line indicates the default value.

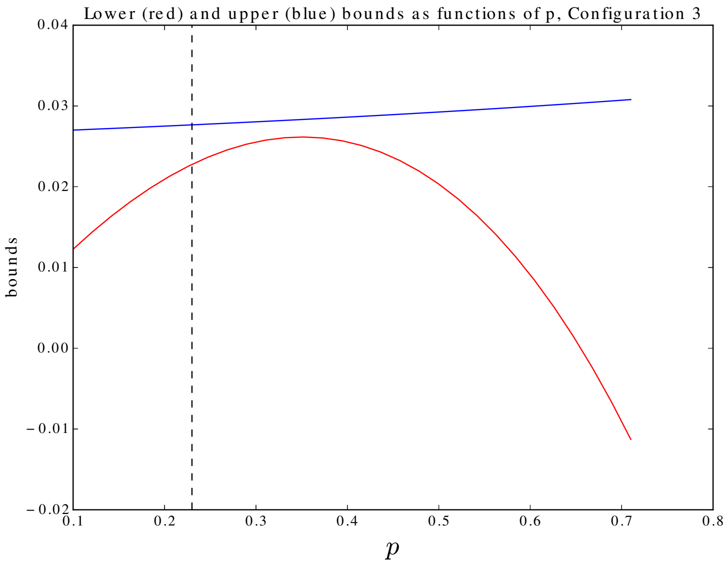

Figure 18: Lower bound (red) and upper bound (blue) in (18) as functions of p in Configuration 3. The vertical black dashed line indicates the default value.

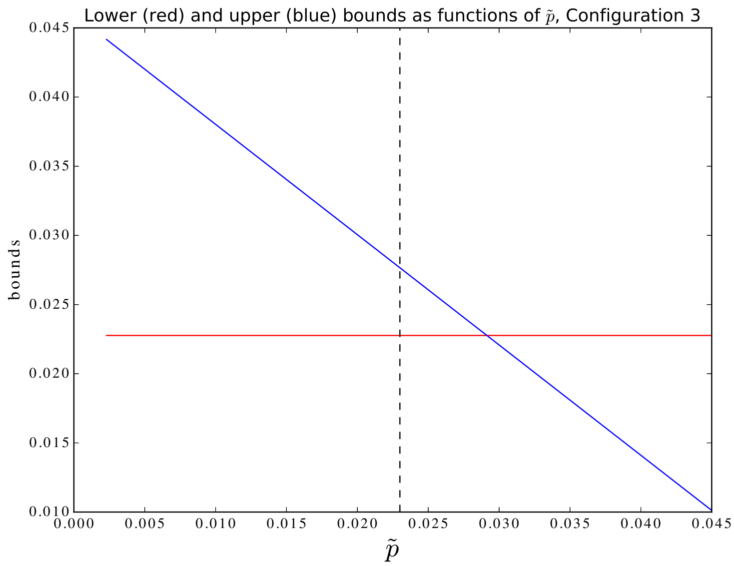

Figure 19: Lower bound (red) and upper bound (blue) in (18) as functions of $\tilde{p}$ in Configuration 3. The vertical black dashed line indicates the default value.

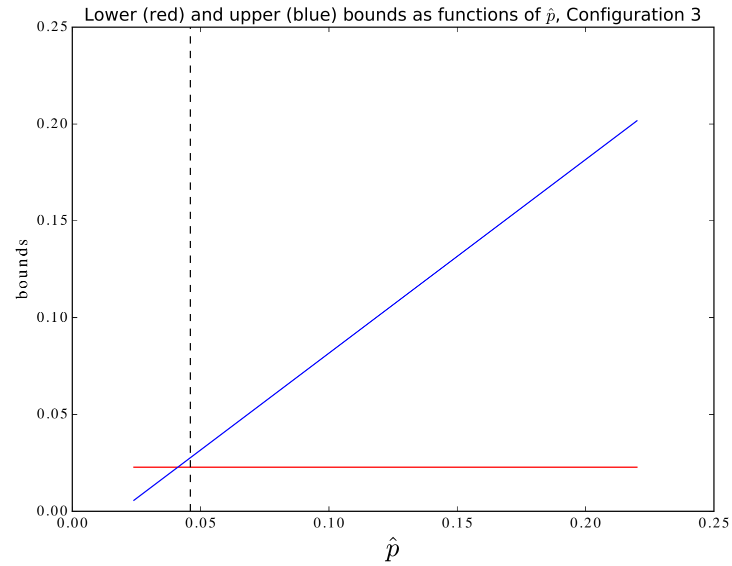

Figure 20: Lower bound (red) and upper bound (blue) in (18) as functions of $\hat{p}$ in Configuration 3. The vertical black dashed line indicates the default value.

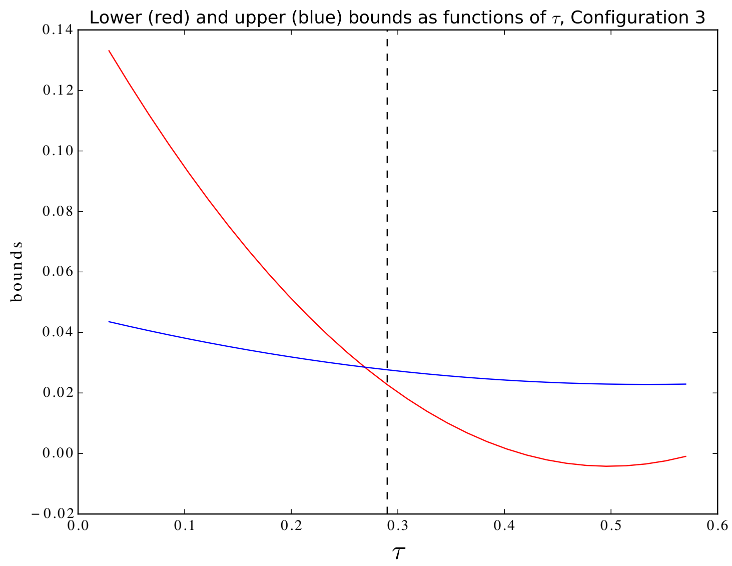

Figure 21: Lower bound (red) and upper bound (blue) in (18) as functions of τ in Configuration 3. The vertical black dashed line indicates the default value.

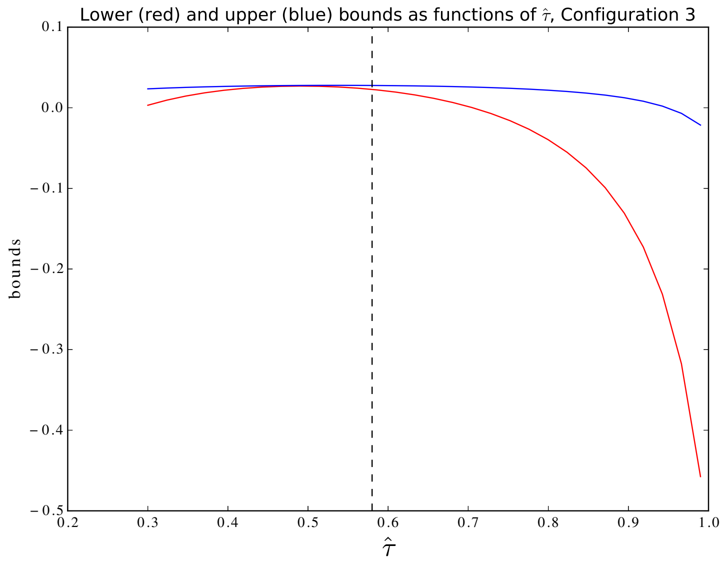

Figure 22: Lower bound (red) and upper bound (blue) in (18) as functions of $\hat{\tau}$ in Configuration 3. The vertical black dashed line indicates the default value.

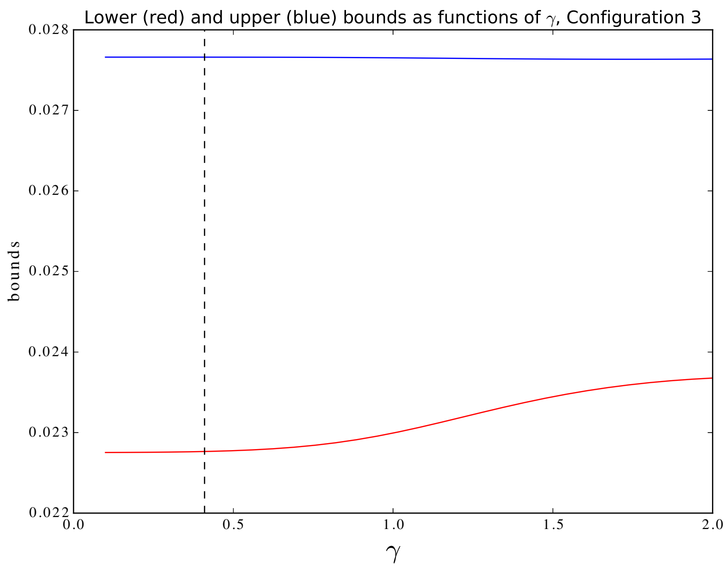

Figure 23: Lower bound (red) and upper bound (blue) in (18) as functions of γ in Configuration 3. The vertical black dashed line indicates the default value.

cost legal alternative to acquiring cultural goods: consumers that are not prepared to pay the monetary price p for a good g may still legally enjoy a lesser quality version of g at the non-monetary price of some technical skills through legalized forums. Such a paradigm raises a number of important legal issues [25].

Technical aspects are also difficult to deal with. Establishing communication and dealing with pirate sites, and then controlling the behaviour of the ones entering in the legalisation process will not be easily achieved. A point must be taken into account that may threaten the whole paradigm: that pirate sites do not conform with the legal system does not mean that they do not obey any rules. To the contrary, as stated in [24], *the emergence of voluntary IP regimes in piratical communities is an important signal that even if they have little respect for statutory copyright or are unwilling to honor every claim made by rights-holders, those millions who engage in piratical practices are well aware of their responsibilities as members of a cultural community. Their decisions, however, are less based on laws, and more defined by the ethics negotiated within the community.* In other words, even though an offer for legalization may be advantageous to most pirate sites, at least some of them will not accept it because it does not fit in their vision of the sharing of cultural goods on the internet.

The version of our model studied above is rather minimal. One needs to extend it in a variety of directions, including the following ones:

1. introduce actors 3, 5 and 6 to obtain a more complete and realistic view of the implications of legalizing some pirate forums;
2. model the increase of the probabilities $(z_k)_{k \leqslant L}$ and deduce the corresponding variations of profits for all actors;
3. it seems natural to assume that all goods are not available at all times from forums. Instead, more popular ones will be more likely to be available. Thus, a more realistic model will attach to each good g_i a probability of being available which is directly proportional to its popularity;
4. consider more general forms for the contribution paid to industry by forums willing to become legal;
5. consider in greater details the specific cases of the music and film industries.

Acknowledgements

JLV wishes to thank Loïc Baud and Eric Walter from the Hadopi for introducing him to this topic and for many fruitful discussions, as well as Soufiane Khoudmi also from Hadopi for several remarks on earlier drafts of this article.

References

[1] CURIEN, N. AND MOREAU, F. (2005). The Music Industry in the Digital Era: Towards New Business Frontiers? *Working paper, Laboratoire d'Econométrie, Conservatoire National des Arts et Métiers Paris, February 9.*

[2] HAUTE AUTORITÉ POUR LA DIFFUSION DES OEUVRES ET LA PROTECTION DES DROITS SUR INTERNET, www.hadopi.fr

[3] CAMMAERTS, B., MENG, B. AND MANSELL, R. (2013) Copyright and creation: a case for promoting inclusive online sharing *LSE Media Policy Project Series, Broughton Micova, Sally and Tambini, Damian, eds, Media Policy Brief 9. Department of Media and Communications , London, UK.*

[4] MCKENZIE, J. (2009). Illegal music downloading and its impact on legitimate sales, *Australian Economic Papers, 48 (4), 296-307, December*.

[5] LIEBOWITZ, S.J. (2008). Research Note: Testing File-Sharing's Impact on Music Album Sales in Cities, *Management Science, 54, 852-859.*

[6] OBERHOLZER F. AND STRUMPF, K. (2007). The Effect of File Sharing on Record Sales: An Empirical Analysis, *Journal of Political Economy, 115, 1-42.*

[7] AGUIAR L. AND MARTENS, B. (2013). Digital Music Consumption on the Internet: Evidence from Clickstream Data, *JRC Technical Reports, Institute for Prospective Technological Studies Digital Economy Working Paper 2013/04*.

[8] ROB, R. AND WALDFOGEL, J. (2006). Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students, *Journal of Law and Economics, 49, 29-62*.

[9] WALDFOGEL, J. (2010). Music file sharing and sales displacement in the iTunes era, *Information Economics and Policy, 22, 306-314*.

[10] PEITZ, M. AND WAELBROECK, P. (2004). The Effect of Internet Piracy on Music Sales: Cross-Section Evidence, *Review of Economic Research on Copyright Issues, 1, 71-79.*

[11] PEITZ, M. AND WAELBROECK, P. (2006). Why the music industry may gain from free downloading - The role of sampling, *International Journal of Industrial Organization, 24 (5), 907-913*.

[12] HUYGEN, A., HELBERGER, N., POORT, J., RUTTEN, P., AND VAN EIJK, N. (2009). Ups and Downs; Economic and Cultural Effects of File Sharing on Music, Film and Games, *TNO Information and Communication Technology Series, (Feb. 18, 2009).* Available at SSRN. http://dx.doi.org/10.2139/ssrn.1350451

[13] http://www.adami.fr/accueil.html

[14] http://www.spedidam.fr/

[15] Enquête SPEDIDAM sur la consommation et échanges d'oeuvres numériques par les internautes français (2007). http://www.spedidam.fr/medias/Mediametrie_rapport_enquete.pdf

[16] ZELNICK, P., TOUBON J. AND CERUTTI G. (2010). Creation et Internet. http://www.adami.fr/fileadmin/user_upload/pdf___docs/02_Defendre/Les_droits_sur_internet/rapport_zelnik.pdf

[17] BAE, S.H. AND CHOI, J.P. (2006). A model of piracy, *Information Economics and Policy 18, 303-320.*

[18] KIEMA, I. (2008). Commercial piracy and intellectual property policy, *Journal of Economic Behavior & Organization, 68, 304-318*

[19] YOON, K. (2002). The optimal level of copyright protection, *Information Economics and Policy 14, 327-348*.

[20] CREMER, H. AND PESTIEAU, P. (2009). Piracy prevention and the pricing of information goods, *Information Economics and Policy, 21 (1), 34-42*.

[21] HADOPI (2013). Etude du volume de consommation des biens et services culturels dématérialisés. www.hadopi.fr/sites/default/files/page/pdf/HADOPI-OPINION-WAY-26062013-Carnet_de_consommation%20%282%29.pdf

[22] HADOPI (2012). Hadopi, biens culturels et usages d'Internet : pratiques et perceptions des internautes français.
http://hadopi.fr/sites/default/files/page/pdf/HADOPI-160113-BU2-Complet.pdf

[23] KARAGANIS, J. AND RENKEMA, L. (2013). Copy Culture in the US and Germany *The American Assembly, Columbia University*
http://piracy.americanassembly.org/copy-culture-report/

[24] BODÓ, B. (2013). Set the fox to watch the geese: voluntary IP regimes in piratical file-sharing communities. *In M. Fredriksson & J. Arvanitakis (Eds.), Piracy: Leakages from Modernity. Sacramento, CA: Litwin Books.*

[25] TBA