XIII CONFERENZA

STATO O MERCATO?
Intervento pubblico e architettura dei mercati
Pavia, Università, 5 - 6 ottobre 2001

FRANCESCA BARIGOZZI - BERTRAND VILLENEUVE

INFLUENCING THE MISINFORMED MISBEHAVER: AN ANALYSIS OF PUBLIC POLICY TOWARDS UNCERTAINTY AND EXTERNAL EFFECTS

pubblicazione internet realizzata con contributo della COMPAGNIA di San Paolo

Società italiana di economia pubblica

Dipartimento di economia pubblica e territoriale – Università di Pavia

In‡uencing the Misinformed Misbehaver: An Analysis of Public Policy towards Uncertainty and External E¤ects[¤]

Francesca Barigozzi[y] Bertrand Villeneuve[z]

May 2001

Abstract

We study a situation in which government in‡uences consumers' behaviors by providing both information and incentives. More generally, we propose a methodology for solving models of signal cum cheap talk.

We develop the case of consumption choice in the presence of uncertainty and external e¤ects. The instruments used by the government are information campaigns and taxes. A di¢culty arises because the government would like to bolster these less than perfectly e¤ective tools by delivering biased information to the misbehaver. We study the equilibrium trade-o¤ between informing and o¤ering incentives. Environmental tax policy, anti-smoking campaigns and policy against over-consumption of antibiotics serve as illustrations.

Keywords: information campaigns, tax policy, cheap talk, signaling, skeleton.

JEL: I18, H30, D82.

1 Introduction

As for the "mad cow" disease and other hotly debated issues concerning public health, food safety and the environment, risk controversies have mushroomed.

[¤]We have received valuable comments, suggestions or corrections from Giacomo Calzolari, Helmuth Cremer, Thomas de Garidel, Bruno Jullien, Alessandro Lizzeri, Maurice Marchand and seminar participants in Bologna, Bolzano, Madrid, Paris, Seattle. Hélène Bories, Patrick Courvalin and Anna Mingardi helped us with the account of antibiotics in the introduction. Errors, in any case, are our own exclusive responsibility.

[y]University of Bologna.

[z]CEA and IDEI, University of Toulouse. Corresponding author: Bertrand Villeneuve, IDEI, Université de Toulouse 1, 21 allée de Brienne, 31000 Toulouse, FRANCE. E-mail: bertrand.villeneuve@cict.fr.

Since policy makers must often assess and communicate such risks, the individuals' confidence in government or other authorities is a decisive component of policy-making. Our work focuses on communication as disclosing the conflicts between a benevolent authority and consumers. Two ingredients are indispensable: the public's ignorance of the risk to be regulated and the impossibility for individuals alone to rightly internalize certain negative consequences of their actions. In isolation, each is relatively easy to solve: the former needs information provided to the public; the latter, optimized incentives. But in combination these remedies interfere with one another and result in political confusion when incentives and coercive instruments are defective. The maximization of social welfare does not necessarily imply truthful policies, and consumers know it. Indeed, they are increasingly skeptical of promotional strategies that use the outcome of the scientific literature or cite expert advisory committees.

In the present paper, the policy-making process is analyzed as a game in which government wants to influence consumers' behaviors through both tax policy and information campaigns, and where rational consumers react in a Bayesian manner. Instruments being imperfect, the government is often tempted to "improve" behavior by providing biased information. Confidence, we show, is not easily controlled. Depending on the coordination between government and consumers, the same background data can produce various policies and real effects. We determine the structure of practicable policies and discuss the trade-off between vagueness in communication and distortion of incentives.

Influence Games We introduce a general methodology for tackling influence games, i.e. games in which the principal is an informed party who combines different instruments to transmit information and provide incentives to the agent. The relevant literature was initiated by Crawford and Sobel (1982) and Milgrom and Roberts (1986). Crawford and Sobel (1982) show that the precision of the information depends on the intensity of the conflict between the two parties' objectives. Quite recently, these ideas have been applied to political games in which a lobby tries to influence policy makers.[1] In a major contribution to the rational foundations of advertising, Milgrom and Roberts (1986) model a firm that signals its product quality through price and dissipative advertising (burned money) to enhance consumers' willingness to pay for the product. In line with these articles, and contrary to Maskin and Tirole (1992), we are not mainly interested in characterizing optimal mechanisms; rather, we study the combination of imperfect mechanisms, giving priority to their practical structure and consequences.

In a recent development along these lines, starting from the classical model of Crawford and Sobel, Austen-Smith and Banks (2000) show how burning money

[1] See Helpman (2000), Presidential Address at the Econometric Society World Congress, Seattle and Grossman and Helpman (2001), for surveys.

can improve cheap talk. In particular, they clearly show why the information transmitted can be perfect, and why the most informative equilibrium need not be the most e¢cient.

We retrieve these results in our context but using a distinctively di¤erent methodology, that is conceived to be applicable to a variety of models thus enabling us to make other useful …ndings. First we change the perspective: we show that cheap talk is almost useless when costly signals are available[2] and explain why more precise equilibria are typically more distorted. Second, rather than searching for the equilibria of a given economy, we de…ne the minimal amount of information (a skeleton) to describe an equilibrium and determine the full set of economies that admit a given skeleton as equilibrium. We prove the uniqueness of the fully revealing equilibrium and characterize it in detail. More generally, the skeleton approach helps understanding of the interaction between costly and free signals and opens the way to interesting comparative statics.

The Analysis of Health and Environmental Policy Our model deals with policies that a¤ect the consumption of commodities that are detrimental to consumers' welfare both individually and collectively. Typically, side e¤ects are due to individual consumption; external e¤ects are due to overall consumption in the economy. Broad-spectrum antibiotics display (apart from the obvious bene…ts) this double negative impact. At the individual level they clear the way to opportunistic infection by more resistant germs;[3] and at the societal level, they enhance the resistance of the germs involved in contagious diseases. Analogously, in the case of tobacco and alcohol, one can readily distinguish between disease related to individual consumption and the social damage from passive smoking or the cost to the health care system (not to mention psychosocial issues like drunk driving and addiction).

These two types of negative e¤ects explain why, without government intervention in the form of information or incentives, consumers may not consume e¢ciently. First, side e¤ects are not necessarily recognized by consumers. For example, the real magnitude of side e¤ects of antibiotics is no more than a vague notion for most people; likewise, the risk smokers perceive can be under- or over-estimated (Viscusi 1990). Second, external e¤ects (e.g. the rise of resistant strains, passive smoke) are very largely ignored by consumers in the absence of incentives such as taxes, norms, controls.

[2] Our proof is direct; an indirect proof is given in Manelli (1996).

[3] Some broad-spectrum antibiotics decrease the individual's immunological response, and as a consequence new diseases can arise. For example, many antibiotics based on penicillin are used to treat diseases like bronchitis, otitis and tonsillitis caused by di¤erent bacteria (staphilococcus aureus, haemo…lus in‡uenzae, streptococcus pneumoniae). Possible side e¤ects of penicillin consumption are candida albicans and herpes. See Levy (1992) for the medical viewpoint and Brown and Layton (1996) for an excellent economic analysis of the external e¤ects.

Political attitudes towards tobacco are typical of the schizophrenia we discuss in the present work. Efficient taxation is, generally, difficult to establish, but compared to others levies, tobacco taxes are an easy source of funds. Government might try to optimize health and budgetary objectives by manipulating consumers' beliefs on the individual consequences of smoking. Obviously, rational consumers form their opinion with this danger in mind, so the success of such attempts is uncertain. In the same vein, let us remark that the pervasive opinion that consuming a lot of antibiotics may cause individual resistance to the treatment is unfounded. (Actually the problem is the resistance acquired by the germs, which concerns the society rather than the individual strictly speaking). Authorities (which may not be directly responsible for this belief) are clearly tempted not to bother correcting it, since it serves (at a low cost) the practical goal of curbing consumption.

We assume that the government is better informed and benevolent and that it maximizes the utility of the representative consumer.

To support our assumption that the government is better informed, note that the government may appeal to experts (civil servants, professionals, academics) who are able to transform dispersed data and results into operational knowledge. We do not need to assume that this operation is perfect, only that it is performed better by the experts than by the general public. Moreover, informing the public is a never-ending task. We know that discouraging teenagers from smoking require renovated strategies year after year. Though one may think that the "society" is already saturated with information on the relationship between tobacco and cancer, each new cohort of consumers still has to be educated.[4]

Nevertheless, the government confronts the following dilemma: taxes are imperfect instruments, and it is tempting to make them work better by disseminating biased information. This engenders a sort of paternalism: the government wants consumers to consume efficiently, but, being unable to commit to neutral and truthful information transmission, it may send interested messages. In our view, in all circumstances where there is some imperfection that impedes economic efficiency, taxes must be understood as a metaphor for the entire comprehensive policy package, also comprising contracts, restrictions, standards and norms.[5]

One crucial aspect of the model is the analysis of the tax as a signal that transmits information on the value of side effects. The tax has two consequences: first, by modifying the price it provides an incentive to internalize the external

[4] Publication of rough scientific findings in the mass media, though in principle a contribution to the formation of public opinion, may generate in confusion. Sensationalism and caricatures are common. The most telling examples are extreme dietary recommendations.

[5] Causes of imperfection are often related to asymmetric information. In moral hazard models, for example, outcomes are observable, but the contribution of effort cannot be perfectly separated from random effects. As a consequence, contracts are rarely able to implement first-best efficiency.

e¤ects; second, it acts as a signal of the value of side e¤ects.

Information campaigns are analyzed as messages, i.e. short statements aiming at informing individuals about the e¤ects of certain goods. The best example are warning labels on cigarette packs or on the hazards of alcoholic beverage. A fundamental characteristic of these campaigns is that they have no direct impact on government's or consumers' utility: they are inexpensive, and to simplify, we can class them as cheap talk messages.[6,7] This implies that the literal meaning of the information conveyed is vague enough not to be falsi…able. Take the warning label "Seriously Harmful to Health" on cigarette packs. This is not false, but the exact nuance it carries is a matter of social convention (speci…cally, how cheap talk is interpreted). Like taxes, information campaigns are imperfect instruments of policy.

Another crucial element is the analysis of tax distortions. In the model, the sign of the marginal cost of public funds is not restricted a priori. For this reason, the paper draws some practical conclusion from the literature on the "double dividend". In this literature, a tax on a polluting good is welfare-improving for two reasons: it reduces pollution and it also reduces the distortions caused by preexisting taxes. Now it has been shown that the double dividend exists only under certain conditions; in particular, this means that the sign of the marginal cost of public funds is not determined a priori (see Goulder 1995 for a survey).[8]

Finally we show, given benevolent authorities and rational consumers, that the major cause of trouble is lack of government credibility. By de…nition, then, government able to commit ex ante to inform truthfully would not encounter the di¢culties we discuss. If such a government is in place, …ne. Yet a distrustful attitude on the part of the public towards informed authorities is frequent: people often feel that the government's actions are motivated by economic interest more that by the public interest (think of the di¤usion of information about the HIV-contaminated blood in the eighties in France and Germany, or "mad cow" disease in Europe).

To solve the social game sketched here, we take an approach based on Bayesian equilibrium: people are not systematically fooled and the government tries to make the best of the instruments available. We analyze one possible cause of consumers' distrust, establishing the trade-o¤ between the precision of the information transmitted and the optimality of the policy implemented: precision is

[6] See Crawford and Sobel (1982).

[7] Note that our approach remain valid if information campaigns are costly, as long as the cost is independent of the message the authority decides to send. Since the di¤usion cost of "smoke is detrimental to one's health" is the same as that of "smoke is very detrimental to one's health", we can, without generality loss, normalize this cost to zero, if no-message is not a choice.

[8] Actually the marginal cost of public funds can be either positive or negative, depending on the relationship between preexisting taxes and the tax on the harmful good. Our model covers all the cases.

higher with less efficient political programs, and conversely. We prove that the equilibrium is never efficient ex ante, and that there exists a unique fully-revealing equilibrium that is almost surely inefficient ex post.

Plan Section 2 presents the terms of the policy dilemma in the case of commodities affecting health and the environment. Section 3 defines the equilibrium. The main body of the paper develops the methodology. After a few results on the structure of the government's preferences (Section 4), the analysis is developed in three steps. First we show that an equilibrium can be summed up by its "skeleton", i.e. a relatively small set of policies satisfying incentive compatibility for the sender (Section 5). Second, we show under what circumstances a given "skeleton" can be implemented in an equilibrium. This is crucial to getting insights into the structure of partially revealing equilibria (Section 6). Finally we characterize the game's unique fully-revealing equilibrium (Section 7). Some implications concerning tax policy for fuels, SO_2 emissions and drugs are discussed in the conclusion.

2 The Model

2.1 The Consumers

Consumers live two periods and the value of their period-2 consumption x_2 is negatively affected by period-1 consumption x_1. Preferences can be written as:

$$U[x_1] + x_2 - \theta x_1 - \eta \bar{x}_1 \tag{1}$$

where U is the logarithmic utility function.[9] The consequences of x_1 on period-2 utility pass through two distinct channels:

- The term $-\theta x_1$ measures **side effects** due to the consumer's own consumption in period 1. The intensity θ is not precisely known to consumers. The cumulative distribution function $F(\theta)$ and its density $f(\theta)$, both supported in $[\underline{\theta}, \overline{\theta}]$, represent consumers' priors on θ. In general, f is continuous and non-negative on the support.

- The term $-\eta \bar{x}_1$ indicates the negative **externality** that depends on $\bar{x}_1$, i.e. average period-1 consumption in the economy. The intensity η is supposed to be known to all the agents.[10] The consumer does not internalize the

[9] Most of the propositions in Section 4 (characterization of the equilibria) do not rely on this restriction on utility, as can be seen in the proofs, but it does improve the legibility of the explicit calculations of the first- and second-best.

[10] An alternative model could put uncertainty on η. In general, though, this uncertainty alone would not exhibit the sort of conflict we are pointing at, since the consumer's behavior is not

social consequences of x_1. This is because there are a large number of atomistic consumers in the economy: each knows that he would affect the externality only marginally.

Let t be the tax rate set by the government. The representative consumer, not internalizing the externality ´; solves:

$$
(2) \qquad \begin{cases} \max_{x_1;x_2} E\left[U[x_1] + x_2 - \mu x_1\right] \\ \text{s.t.}: (p_1 + t)x_1 + p_2 x_2 = W \end{cases}
$$

where the expected value of utility is conditional on the consumer's information; p_1 and p_2 are the prices for, respectively, period-1 and -2 consumption, and W is the consumer's endowment.

To further simplify the program we normalize p_2 to 1 and p_1 to 0 without loss of generality since the support of μ can be translated to account for the price, which is exogenous: Then we substitute the budget constraint into the objective function and drop the subscripts to write the first period consumption as x: The simplified consumer's program is:[11]

$$
(3) \qquad \max_x E\left[U[x] - (\mu + t)x\right]
$$

As a consequence, consumption choice x^* depends on the consumer's information and on the tax rate t:

$$
(4) \qquad x^*[t; E\mu] \text{ solves } U'[x] = E\mu + t
$$

that is

$$
(5) \qquad x^*[t; E\mu] = \frac{1}{E\mu + t}
$$

2.2 Social Welfare and the Marginal Cost of Public Funds

The social welfare function that is maximized by the government corresponds to the consumers' utility once the externality and the exact value of side effects are taken into account. The revenue from income-, capital-tax, or other levies is exogenous and taxation is distortionary. Within this public finance perspective, we calculate how x should be taxed. In addition to the fact that taxes are imperfect, the government is hampered by its inability to commit to a policy that informs truthfully on the value of μ. In other words, the government is benevolent in that it evaluates consumption in consumers' best interest, but opportunistic because

affected by the intensity of the externality. In our specification, consumption does not even depend on ´:

[11] Notice that the linear part in the preferences also represent the utility from goods other than x:

it does not value truthful information *per se*, and would deceive consumers provided this induces "better" behavior (and reduces the distortions caused by the tax): in short it practices a variety of paternalism.

All consumers being identical, in equilibrium $\hat{x} = x$, and the government's objective function can be represented as:

$$U[x] - (\theta + \delta)x + S - (1 + \lambda)R \tag{6}$$

where S is the consumer's surplus from public expenditures R: The government raises R with general taxation at the welfare cost $(1 + \lambda)R$; with $\lambda > -1$: Most partial equilibrium models call parameter λ the "shadow cost" of public funds; it represents the distortion due to the raising of fiscal revenue.[12]

In our model, R and S remain constant, while a new tax on the good x is added to existing taxes. As modern public finance theory has shown, no general conclusion can be drawn about the sign of the shadow cost of taxation when a revenue-neutral substitution between different taxes is implemented. The sign of λ is not restricted *a priori* and depends on the structure of preexisting taxes (in particular their efficiency) and on how they interact with the new tax.[13]

When the government introduces a tax t on good x and revenue tx is devoted to reducing preexisting taxes, (6) becomes, after simplification:

$$U[x] - (\theta + \delta - \lambda t)x + S - (1 + \lambda)R \tag{7}$$

Comparing (3) and (7), we see that the government's program differs from the consumer's in three ways: superior information on θ; internalization of δ; and the presence of λ in the government's objective function. As for the externality, the consumer does not internalize the effect of his contribution tx on the total distortion caused by taxation.

The case $\lambda > 0$ (preexisting taxes inflict a welfare cost larger than R) relates to the recently debated "double dividend" effect. According to this literature, a revenue-neutral substitution of environmental taxes for ordinary income taxes might offer a double dividend: not only does it (1) improve the environment but it also (2) reduces the costs of the tax system through cuts in distortionary taxes (see Goulder 1995). To intuit this result, assume that x and the other taxed goods (labor included) are gross substitutes. In that case, typically, the tax t reduces the consumption of x and increases the consumption of the other taxed goods. Thus total fiscal revenue increases and taxes on the other goods can be reduced,

[12] See, for example, the shadow cost of public funds used in the theory of regulation (Laffont and Tirole 1993). In general equilibrium models of taxation (e.g. Ramsey), λ would be the (endogenous) Lagrange multiplier associated with the government's budget constraint. Under some regularity conditions, the Lagrange multiplier is equivalent to what the theory of cost-benefit analysis calls the shadow cost of a marginal change in a public project. See Drèze and Stern (1987).

[13] For a concise discussion, see Ballard and Fullerton (1992) and Goulder (1995).

which attenuates distortions.[14,15] Notice that we could also reason in terms of relative efficiency: when $\lambda > 0$ a tax on good x is relatively less distortionary than preexisting taxes; when $\lambda < 0$, it is more distortionary.

Dropping constant terms that are not relevant to policy decisions, we get a reduced form of the government's objective function:

$$SW[x;t;\theta] \equiv U[x] - (\theta + \eta - \lambda t)x \tag{8}$$

2.3 Constrained Efficient Allocations

The first-best allocation is defined as the allocation that maximizes consumers' utility when θ is known, η internalized, and the economy is not distorted ($\lambda = 0$). This gives $x_{FB}(\theta) = \frac{1}{\theta+\eta}$ (see equation (3) with η instead of t). This allocation can be implemented even without full control over x with the standard Pigovian tax $t = \eta$.

The second-best allocation is defined as the best the government can attain if consumers are perfectly informed on θ when (1) it is constrained to a linear tax on x; and (2) the marginal cost of public funds is not zero. This can be written as follows:

$$\begin{cases} \max\limits_{t} U[x] - (\eta + \theta - \lambda t)x \\ \text{s.t.} : x = \frac{1}{\theta+t} \end{cases} \tag{9}$$

where the constraint on x corresponds to consumers' reaction function (5) when θ is known. Hence the second-best consumption and tax rate are:

$$\begin{cases} x_{SB}(\theta) = \frac{1}{\eta+(1+\lambda)\theta} \\ t_{SB}(\theta) = \eta + \lambda\theta \end{cases} \tag{10}$$

Because of the "double dividend", the second-best tax is higher than the first-best when $\lambda > 0$. The opposite holds for $\lambda < 0$. For a given λ, the tax rate is strictly increasing (decreasing) with respect to θ when $\lambda > 0$ ($\lambda < 0$). In any case, it is important to note at this step that the tax rate is potentially informative on the value of side effects.

Straightforward calculations lead to a sort of Ramsey-Boiteux pricing rule:[16]

$$\frac{t_{SB}(\theta) - \frac{\eta}{1+\lambda}}{t_{SB}(\theta)} = \frac{\lambda}{1+\lambda}\frac{1}{\varepsilon} \tag{11}$$

---

[14] Complementarity between x and the other taxed goods allows the same reasoning to hold when $\lambda < 0$.

[15] The same reasoning in terms of substitutability and complementarity between x, the other goods, and the public project (financed by R) applies. In other words, the cost of public funds also depends on the interaction between the public expenditures and the taxed activities.

[16] A similar expression can be found in Sandmo (1975). See also Bovenberg and van der Ploeg (1994).

where $\varepsilon \equiv -\frac{\partial x^*}{\partial t}\frac{t}{x^*} = \frac{t}{\theta+t}$ is the tax elasticity of demand. As tax elasticity is decreasing in θ for all t, (11) shows that for positive λ, the stronger the side effects, the higher the tax. For negative λ, the opposite holds.

From (8), and for a given tax t, the marginal net external effect of consumption x is $\delta - \lambda t$, where δ is for external effects sensu stricto, and $-\lambda t$ for the effects of the tax on public finance that the consumers do not internalize. Assume that $\lambda > 0$. First, the social welfare function (8) shows that, others things (in particular x) equal, the government would prefer to impose a "high" (in fact infinite) tax, in order to engender large positive externalities. Second, given such positive externalities, the government must decide how to tax x in order to make consumers internalize them; according to the Pigou rule $|t - (\delta - \lambda t)|$ should be minimized, and this is attained at $t = \bar{t} \equiv \frac{\delta}{1+\lambda}$. Unfortunately, these two arguments draw in opposite directions and the two goals rely on the same instrument, t. This explains why the optimal trade-off is $t_{SB}(\theta)$, a value between $\bar{t}$ and $+\infty$.[17]

3 The Influence Game

The timing of the model is as follows: first the government observes θ; then it chooses its policy, and finally the consumer, observing the policy, updates his beliefs on θ and chooses his consumption level.

A policy $P = (t; m) \in \mathbb{R} \times M$ is composed of the tax rate t and a (cheap talk) "message" m selected from a certain large set, M. Through the choice of a policy P, the government wants to induce the consumer to approach efficient consumption. The tax has the two-fold role of incentive and information; cheap talk can only transmit information. We can think of m as composed of a "sentence". We assume that M is broad enough to say what needs to be said; it might be composed, say, of all reasonably short utterances (see, e.g., Farrell and Rabin 1996 on what cheap talk is and is not). It is useful, at this point, to make a distinction between the message the government sends and the consumers' interpretation of it at equilibrium. What really matters is not the message itself but the way the consumer understands the policy. To be clearer, whatever the phrasing of the communication, we concentrate on the meaning (the revised $E\theta$) that the consumer assigns to every policy.[18]

After observing the policy, the consumer updates his priors which are then denoted by $\mu(P)$ (with $\mu(P) \in \Delta([\underline{\theta}; \overline{\theta}])$, the set of probability distributions over

[17] When $\lambda < 0$, with fixed x, the government would like large subsidies ($t \to -\infty$); $t_{SB}(\theta)$ is between $-\infty$ and $\bar{t}$.

[18] As an example, let m_1 and m_2 denote two messages sent in a fully-revealing equilibrium. Assume that m_1 corresponds to the word "dog" and m_2 corresponds to the word "cat". This is an equilibrium as long as the receiver understands this language and assigns to the message "dog" the meaning, say "$\theta = \theta_1$", and to the message "cat" the meaning, say "$\theta = \theta_2$", where θ_1 and $\theta_2 \in [\underline{\theta}; \overline{\theta}]$.

$[\underline{\theta}; \overline{\theta}])$. We denote $E(\theta|P)$ by $\hat{\theta}(P)$.

Definition 1 A perfect Bayesian equilibrium (PBE) of the game is a pure strategy P mapping $[\underline{\theta}; \overline{\theta}]$ into $\mathbb{R}_+ \times M$ and a belief μ mapping $\mathbb{R} \times M$ into $\Delta([\underline{\theta}; \overline{\theta}])$ such that:

1. Policies are optimal given beliefs: for each $\theta \in [\underline{\theta}; \overline{\theta}]$, $P(\theta)$ solves

$$\max_{P} SW[x^*[t; \hat{\theta}(P)]; t; \theta]. \tag{12}$$

2. Beliefs are rational given equilibrium policy: for each P, $x^*[t; \hat{\theta}(P)]$ solves

$$\max_{x} \int_{\underline{\theta}}^{\overline{\theta}} [U[x] - (\theta + t)x]\mu(\theta|P)d\theta, \tag{13}$$

where $\mu(\theta|P) \equiv \frac{\mathbb{1}_{\{P(\theta)=P\}} \Delta f(\theta)}{\int_{\underline{\theta}}^{\overline{\theta}} \mathbb{1}_{\{P(\theta)=P\}} \Delta f(s)ds}$; $\mathbb{1}$ being the indicator function.

The concern that revised beliefs may not always be well-defined is dealt with in Proposition 3, below.

4 Government Policy Preferences

Bad news for communication gurus: the consumer's rationality prevents the government from transforming lead into gold by clever communication strategies. In other words, we preclude any perverse mechanism whereby propaganda can make less desirable states of the world (larger side effects) preferable.

Proposition 1 In any equilibrium, the larger the side effects, the lower the social welfare.

Proof. Let θ_1 and θ_2 be two possible states of the world, $P_1 = (t_1; m_1)$ and $P_2 = (t_2; m_2)$ two equilibrium policies, and x_1 and x_2 the consumption levels induced. If $\theta_1 < \theta_2$, then $U[x_2] - (\eta + \theta_1 - \lambda t_2)x_2 \geq U[x_2] - (\eta + \theta_2 - \lambda t_2)x_2$. On the other hand, the incentive constraint of the type-θ_1 social planner reads: $U[x_1] - (\eta + \theta_1 - \lambda t_1)x_1 \geq U[x_2] - (\eta + \theta_1 - \lambda t_2)x_2$. By transitivity, we get: $U[x_1] - (\eta + \theta_1 - \lambda t_1)x_1 \geq U[x_2] - (\eta + \theta_2 - \lambda t_2)x_2$. Thus the social planner's pay-off decreases with respect to the side effects. ∎

Now let us analyze the government's incentive to manipulate information: that is, the reasons why consumers are likely to be suspicious of the government's actions and claims, and how suspicious.

Remark 1 In equilibrium, any policy P can be analyzed without loss of insight as a pair $(t; \hat{\theta})$, where t is the tax rate, and $\hat{\theta}$ the belief associated to the policy.

We define $\overline{SW}[t, \hat\theta, \theta] \equiv SW[x^*[t, \hat\theta], t, \theta]$ as the value of a policy characterized by the tax-belief pair $(t, \hat\theta)$ for a government of type θ. Reasoning directly on tax-belief pairs allows a simpler analysis of incentive constraints, independently of the cheap talk message chosen. Indeed, incentive compatibility for $P(\theta) = (t, \hat\theta)$ and $P(\theta') = (t', \hat\theta')$ can clearly be checked by comparing $\overline{SW}[t, \hat\theta, \theta]$ with $\overline{SW}[t', \hat\theta', \theta]$, and $\overline{SW}[t, \hat\theta, \theta']$ with $\overline{SW}[t', \hat\theta', \theta']$.

The consumer solves $U'[x] = \hat\theta(P) + t$. Therefore from (8), we can see that the consumer's choice is equal to the socially optimal consumption when:

$$\hat\theta(P) + (1 + \lambda)t = \theta + \delta \tag{14}$$

Suppose that the consumer is naive and believes whatever the government announces. The government would set the tax rate and induce beliefs so that (14) is verified. Note that the right-hand side of (14) is a constant. When $\lambda > 0$ (t increases social welfare), the government prefers to set high taxes and to induce low beliefs. In other words, it prefers to make the consumer optimistic about side effects, and relies mostly on taxation. The opposite is true when $\lambda < 0$ (t yields deadweight losses): the government prefers to make the consumer pessimistic about side effects and to drive taxation to its lowest level.

Vis-à-vis a rational consumer, such a policy is obviously never consistent; nevertheless, it provides useful indications on the incentives that the government perceives. For instance, when only cheap talk is available, equation (14) becomes $\hat\theta(P) = \theta + \delta$. In this case, the government always has incentives to overstate the value of θ so as to make the consumer internalize the externality. The setting then resembles Crawford and Sobel (1982), which as we saw earlier, explains why the health authorities are better off when consumers have an exaggerated perception of the side effects of antibiotics.

Now we become more formal. Policies are restricted to induce finite consumption. Thus feasible policies are such that $t + \hat\theta > 0$. The difficulty here is that indifference curves are not monotonic: there is an optimal policy (unfortunately inconsistent with Bayesian consumers, as we shall show), and utility decreases as the tax and the belief get farther from the optimum. Nevertheless, the following proposition gives us some useful properties to proceed with the analysis of incentive compatibility.

Proposition 2

1. For all θ, the upper contours of $\overline{SW}$ with respect to t and $\hat\theta$ are convex.
2. For all θ, tangents to indifference curves are horizontal along the straight line $(1+\lambda)t+\hat\theta = \delta+\theta$, and vertical along the straight line $t+(1-\lambda)\hat\theta = \delta+\theta$. The overall optimum is the intersection of these lines ($t = \frac{\delta+\theta}{\lambda}$, $\hat\theta = -\frac{\delta+\theta}{\lambda}$); the optimum with $t = 0$ is $\hat\theta = \theta + \delta$.

3. Let $V(\theta)$ be an indifference curve for type θ passing through $(t, \text{þ})$. $V(\theta)$ turns continuously clockwise if $\lambda > 0$ (anti-clockwise if $\lambda < 0$) locally at $(\text{þ}, t)$ as θ increases and indifference curves related to two different types cross once at most.

Proof. 1. It suffices to verify that the utility is quasi-concave. To do this, we check that the successive principal minors of the bordered Hessian matrix have alternate signs (odd principal minors must be positive). The bordered Hessian matrix is:

$$(15)\quad \begin{bmatrix} 0 & \frac{\eta+\theta-t-(1-\lambda)\text{þ}}{(t+\text{þ})^2} & \frac{\eta+\theta-(1+\lambda)t-\text{þ}}{(t+\text{þ})^2} \\ \frac{\eta+\theta-t-(1-\lambda)\text{þ}}{(t+\text{þ})^2} & -\frac{2\eta+2\theta-t-(1-2\lambda)\text{þ}}{(t+\text{þ})^3} & -\frac{2\eta+2\theta-(1+\lambda)t-(1-\lambda)\text{þ}}{(t+\text{þ})^3} \\ \frac{\eta+\theta-(1+\lambda)t-\text{þ}}{(t+\text{þ})^2} & -\frac{2\eta+2\theta-(1+\lambda)t-(1-\lambda)\text{þ}}{(t+\text{þ})^3} & -\frac{2\eta+2\theta-(1+2\lambda)t-\text{þ}}{(t+\text{þ})^3} \end{bmatrix}$$

The first principal minor is equal to zero, the second is negative, and we find $\frac{\lambda^2}{(t+\text{þ})^4}$ for the third, which gives result required.

2. and 3. The MRS between t and þ is

$$(16)\quad \left.\frac{dt}{d\text{þ}}\right|_{SW=\text{constant}} = -\frac{\eta+\theta-(1+\lambda)t-\text{þ}}{\eta+\theta-t-(1-\lambda)\text{þ}}$$

Its derivative with respect to θ is

$$(17)\quad -\frac{\lambda(t+\text{þ})}{(\eta+\theta-t-(1-\lambda)\text{þ})^2}$$

which is negative (positive) for $\lambda > 0$ ($\lambda < 0$) for $t+\text{þ} > 0$. Tangents to indifference curves being vertical whenever $\eta+\theta-t-(1-\lambda)\text{þ} = 0$, the claim is correct. The optimum is the singular point where both the numerator and the denominator of (16) are equal to zero.

Notice that if one sets aside domain restrictions, upper contours are closed, meaning that two indifference curves related to two different types, if ever they cross, cross twice at least. Since we have proved here that in the relevant range ($t+\text{þ} > 0$) crossing occurs once at most, the standard argument based on single crossing can be retrieved. ■

Figure 1 shows the government's indifference curves for $\eta = 1$; $\underline{\theta} = 0$; $\bar{\theta} = 1$; $\lambda = .7$ and $\theta = 1$.

Insert Figure 1 here.

We now can prove that the second-best policy is not implementable in a Bayesian equilibrium. Suppose the consumer thinks that the government is playing the second-best strategy. The tax schedule t_{SB} being invertible; if $t_{SB}(\theta)$ is imposed, the individual can infer θ unambiguously. Nevertheless, it is not possible to implement this allocation in a Bayesian equilibrium. In fact, the fiscal revenue $t_{SB}(\theta)x_{SB}(\theta)$ increases as θ decreases. When $\lambda > 0$ and θ is high, the government may have an interest in reporting a lower θ; in other words in making the consumer optimistic. On the contrary, when $\lambda < 0$; the government may have interest in making the consumer pessimistic. Proposition 1 confirms that the government faces strong incentives to provide biased information.

Corollary 1 The second-best allocation is never an equilibrium if $\lambda \neq 0$:

In fact, at $(t_{SB}(\theta); \theta) = (\delta + \lambda\theta; \theta)$; for all θ; the tangent of the indifference curve of the government is vertical (see point 2 in the Proposition): small changes in the tax have first-order effects, whereas small changes in the beliefs have only second-order effects on the government's objective. Consequently, if $\lambda > 0$; any policy close to the second-best $(\delta + \lambda\theta; \theta)$ but with $t < \delta + \lambda\theta$ is preferred; this holds for a second-best policy associated with a slightly type. If $\lambda < 0$; second-best policies associated with slightly higher types are preferred. In any case, the second-best allocation is not incentive-compatible, which confirms that the government has a strong incentive to bias its information campaign.

Note in contrast that when $\lambda = 0$; the first-best allocation is implementable in a PBE. Indeed, given that the government has no incentive to lie (see (14)), the tax is specifically used to internalize the externality ($t = \delta$), but since the tax rate is uninformative on θ, cheap talk must be used to eliminate asymmetric information. With a slight abuse, equilibrium policies can be written as $P = (t = \delta; \hat{\theta} = \theta)$; where information is fully transmitted. This is, of course, a very particular case.

5 Skeletons

Description of all the equilibria given the prior type distribution is difficult. So rather than look for equilibria in the traditional way for signalling games, we introduce a different technique. That is we solve the inverse problem: we find the set of types and the distributions of types that are consistent with a certain equilibrium allocation. This new approach has some relationship with mathematical tools mostly used in imagery (namely Voronoi diagrams, and their dual, Delaunay triangulation) from which we borrow our terminology (the skeleton).[19] The analogy is the following: given a partition of the types, types in each subset applying the same (unknown) policy, and two different subsets applying different

[19] The idea is basically the following: the Voronoi diagram of a point set P is a subdivision of the plane with the property that the Voronoi cell of point p contains all locations that are closer to p than to every other point of P. The points of P are also called Voronoi generators. Each edge of a Voronoi cell is the bisector of the connection of p to the corresponding neighbour cell. See http://www.voronoi.com/ for theory, algorithms, and examples of applications.

policies, one may want to inquire into the underlying policies. Conversely, given a certain set of policies, and given that the government responds to its incentives, one may be interested in the types that have to be associated with each policy. In all these problems, preferences can be seen as a measure of distance.

The following proposition generalizes the well-known result (Crawford and Sobel 1982, henceforth CS) that all equilibria are "partition equilibria". See also Austen-Smith and Banks (2000).

Proposition 3 Any equilibrium allocation can be implemented in an equilibrium in which there exists a partition of $[\underline{\theta}, \overline{\theta}]$ into a set of intervals $\{I_k\}_{i \in K}$ (K is a minimal set of indices) and a set of policies $\{P_k\}_{k \in K}$ such that (i) the policy chosen in I_k is P_k; and (ii) $k \neq k'$ implies $I_k \neq I_{k'}$ and $P_k \neq P_{k'}$. Moreover, the effects of policy P_k are entirely characterized by the pair $(t_k, \hat{\theta}_k)$, where $\hat{\theta}_k \equiv E(\theta|P_k) = E(\theta|I_k)), \forall k.$

Proof. In this proof, optimal is used in the weak sense. In any PBE, for all P that are equilibrium actions, the set of types for which P is optimal is a convex subset of $[\underline{\theta}, \overline{\theta}]$. To see this, consider the sender's incentive constraint in a given equilibrium. Type θ will prefer policy $P_1 = (t_1, m_1)$ to any $P_2 = (t_2, m_2)$, implying, respectively, consumptions x_1 and x_2, if and only if:

$$(18) \quad U[x_1] - (\eta + \theta - \lambda t_1) x_1 \geq U[x_2] - (\eta + \theta - \lambda t_2) x_2 ,$$
$$\theta(x_2 - x_1) \geq U[x_2] - U[x_1] + \eta(x_1 - x_2) + \lambda t_2 x_2 - \lambda t_1 x_1$$

This equation defines either a half straight-line in the space of types ($x_1 - x_2 \neq 0$) or the whole real line ($x_1 = x_2$). From this, it follows that if policy P is optimal for two values of θ, then it is optimal for any type that lies between these two values.

Let us denote by (θ_1, θ_2), with $\theta_1 \neq \theta_2$, an interval in which P_1 is optimal. We now check that there is only one optimal policy in the interval. Suppose, for purpose of argument, that this is not the case, e.g. $\exists \theta \in (\theta_1, \theta_2)$ for which both P_1 and $P_2 (\neq P_1)$ are optimal. Equation (18) thus becomes

$$(19) \qquad \theta(x_2 - x_1) = U[x_2] - U[x_1] + \eta(x_1 - x_2) + \lambda t_2 x_2 - \lambda t_1 x_1$$

Now either $x_1 \neq x_2$, and, according to (18), P_1 is strictly preferred to P_2 on one side of θ, and P_2 is strictly preferred to P_1 on the other side, which is in contradiction with our assumption that P_1 is optimal on (θ_1, θ_2), or else $x_1 = x_2$, which implies in turn that $t_1 = t_2$, and (given that consumptions are only a function of the tax and the beliefs) that P_1 and P_2 imply the same beliefs. In this case, P_1 and P_2 are the same in terms of tax and beliefs. Though they may differ in their cheap talk dimension, they can be seen as identical, and Remark 1 shows why.

If an equilibrium allocation were not implementable by a strategy based on a partition into intervals, then the latter result would be false. Hence, our claim is proven. ■

One substantial implication of this result is that Condition 1 in the definition of the equilibrium (Section 3) implies the non-evident property that beliefs in Condition 2 are well-defined (indeed, strategies inherit the measurability of the space of types).

Proposition 3 suggests that only a "few" tax-belief pairs are interesting. We can go further and show that only a "few" incentive-compatibility constraints need to be checked to ensure that an allocation is an equilibrium.

Definition 2 (Skeleton) Let $\{\theta_k\}_{k\in K}$ be a closed subset of $[\underline{\theta};\bar{\theta}]$ in which $k \neq k'$ implies $\theta_k \neq \theta_{k'}$ (K is a minimal set of indices); and let $\{t_k\}_{k\in K}$ be a set of real numbers. $\{(t_k;\theta_k)\}_{k\in K}$ is said to be a skeleton if and only if $\forall k;k' \in K; \overline{SW}[t_k;\theta_k;\theta_k] \geq \overline{SW}[t_{k'};\theta_{k'};\theta_k]$ (incentive compatibility).

One particularity of the skeleton is that any equilibrium to which it is connected is revealing for the types of the skeleton, and for these types only (this is represented by the incentive constraints in the definition of skeleton).

Proposition 4 below is the reciprocal of Proposition 3. We exploit the idea that the skeleton represents the essential data that characterize an equilibrium. The type support can be divided into intervals in which the strategy is pooling, and we specify the restrictions on the "flesh" (the distribution F) that may be put on the "bones" (the skeleton) to obtain an equilibrium.

Proposition 4 Let $\mathcal{F}$ be the set of type distribution F such that the skeleton $\{(t_k;\theta_k)\}_{k\in K}$ is an equilibrium set of policies. There exists a partition of $[\underline{\theta};\bar{\theta}]$ into a set of intervals $\{I_k\}_{k\in K}$ with $\theta_k \in I_k$ such that: $\forall F \in \mathcal{F}; \forall k; t(\cdot) = t_k$ over I_k and $E(\theta|I_k) = \theta_k$.

Proof. By convention, we denote the lowest element of $\{\theta_k\}_{k\in K}$ as θ_1; and the highest as $\theta_{\bar{1}}$. Given θ_k; we define its successor in $\{\theta_k\}_{k\in K}$ as $\theta_{k+1} \equiv \min_{k'\in K}\{\theta_{k'} > \theta_k\}$ (this "+1" is just a convention, inspired by the fact that when K is finite, it can be reformulated as a set of successive integers). The type θ_{k+1} is well defined since a skeleton is closed.[20] We reason on incentive compatibility.

If $\theta_{k+1} \neq \theta_k$; we denote by τ_k a type which is indifferent between $P(\theta_k)$ and $P(\theta_{k+1})$; i.e. $\overline{SW}[t_k;\theta_k;\tau_k] = \overline{SW}[t_{k+1};\theta_{k+1};\tau_k]$. Given the single crossing property, and given the continuity of the government's welfare function with respect to the true type, τ_k is unique and belongs to $[\theta_k;\theta_{k+1}]$. We define $I_k =$

[20] Notice that we assume that $\{\theta_k\}_{k\in K}$ is closed only to simplify our reasoning. This assumption is in facti without loss of generality: if an accumulation point of $\{\theta_k\}_{k\in K}$ were missing (i.e. if $\{\theta_k\}_{k\in K}$ were not complete), we could add it to $\{\theta_k\}_{k\in K}$, with a corresponding accumulation point in $\{t_k\}_{k\in K}$. Due to the continuity of the incentive constraints, incentives are not reversed, and the skeleton is completed.

$(\tau_{k-1}, \tau_k]$: If the successor of θ_k is θ_k itself (this happens if θ_k is, on the right, an accumulation point in $\{\theta_k\}_{k\in K}$), then $I_k = \{\theta_k\}$: The lower bound of the lowest interval (i.e. containing θ_1) is $\underline{\theta}$, and the upper bound of the upper interval (containing θ_1) is $\bar{\theta}$: Given Proposition 3, $t(\cdot) = t_k$ over I_k for all k is incentive-compatible. Finally, to ensure that the equilibrium beliefs of the consumer are Bayesian, it is necessary and sufficient that $F(\cdot)$ be such that $E(\theta|I_k) = \theta_k$: ■

Conditional expectations (with respect to the policy, or to the interval) are independent of one another. The probability associated with the interval I_k not being constrained, $F \in \mathcal{F}$ can be chosen to be as smooth as wanted.

Figure 2 summarizes the notation and the main properties of the skeleton.

Insert Figure 2 here.

Corollary 2 If two different intervals are associated with two different tax rates, then the tax rate is sufficiently informative for the consumer, and the message can be ignored. If there exists $k \neq k'$ such that $t_k = t_{k'}$; then messages are indispensable to signal the right interval and ensure the right beliefs.

When the tax rate is the same for two or more intervals, cheap talk serves to transmit some information. In the terminology of Austen-Smith and Banks (2000), cost-free signalling is influential if two different cheap talk messages associated with the same tax rate have to be used to distinguish two different intervals.

Similarities with CS are obvious: Propositions 3 and 4 show that the government can use meaningful yet imprecise policies to communicate the side effects to consumers. Since the government has interest in lying on the value of θ to make consumers internalize the externality, powerful communication campaigns would give the government the means of manipulating consumers' beliefs. As a consequence, the government is restricted in equilibrium to vague statements that only specify broad ranges within which θ may lie.

This trade-off is classical for readers accustomed to cheap talk: the partition $\{I_k\}_{k\in K}$ entails a loss in precision, but now, if the government wants to lie, it has to pretend that the side effects are in a different subinterval, which changes consumers' consumption by a discrete amount. Such "big lies" are less attractive than telling the truth.

A less evident conclusion is that there are also considerable differences from CS. The skeleton approach enables us to show that partitions need not be finite, meaning that the precision of the message may be arbitrarily high locally. In this sense a new trade-off arises: as precision increases, tax policies are more severely constrained by incentive compatibility, and distortions away from the second-best become large.

6 Partially Revealing Equilibria

We can start to build an equilibrium by choosing a skeleton and filling the distribution while preserving conditional expectations. Proceeding in this way, we readily provide examples in which the tax rate is not monotonic, where it is revealing on certain subsets of $[\underline{\mu}; \overline{\mu}]$ with bundles elsewhere, etc. Hence a multiplicity of partially revealing and pooling equilibria are conceivable.

Proposition 4 does not claim that some distribution F always exists. Indeed, even off-equilibrium beliefs are constrained to be in $[\underline{\mu}; \overline{\mu}]$, and we may be short of sufficiently dissuasive off-equilibrium beliefs to support a skeleton. We are nevertheless able to furnish a simple way of extending a skeleton to make F non empty, in other words, to implement the skeleton in a PBE.

Proposition 5 Any skeleton is either directly implementable or can be made implementable by adding one policy (one belief and its associated tax).

Proof. Let us take an un-implementable skeleton. If $\theta > 0$; the simplest way to complete it is to add a sufficiently low type, say $\mu_{min} < \mu_1$; coupled with $t_{SB}(\mu_{min})$. This may mean enlarging $[\underline{\mu}; \overline{\mu}]$ by replacing $\underline{\mu}$ with μ_{min}: If we associate belief μ_{min} with any tax outside $\{t_k\}_{k \in K}$; we still have a skeleton. To check incentive compatibility, note that if belief μ_{min} is sufficiently small compared to μ_0; it is necessarily too small compared to any type of government drawn in $\{\mu_k\}_{k \in K}$. Moreover $t = t_{SB}(\mu_{min})$ is better than any other value of the tax for a government of type μ_{min}: The new skeleton is now implementable. If $\theta < 0$; the same reasoning holds for a large $\mu_{max} > \mu_1$ (coupled with $t_{SB}(\mu_{max})$). ■

This suggests that, as long as priors are defined over a sufficiently large set, and even if extreme types are extremely unlikely, one may exploit the presence of "scarecrow" types to build equilibria.

Our model differs significantly from pure cheap talk models. Indeed, with finite skeletons, cheap talk doesn't really need to be influential (i.e. useful), since either all taxes are different or some are identical and we can use the continuity of the incentive constraints to modify the skeleton slightly and make all tax rates different, in which case cheap talk is useless. In other words, suppose that $(t_k; m_k)$ and $(t_{k'}; m_{k'})$ are two equilibrium policies with $t_k = t_{k'} = t$ and $m_k \neq m_{k'}$ (cheap talk is influential). If we change one of the two taxes, the partition in the skeleton has to be modified, but changes remain small because there are only a finite number of bones (hence a finite number of continuous incentive constraints) and the welfare cost of changing is arbitrarily low. The extension of this intuition to a large set of signals is not developed here. See Manelli (1996) for another approach to the same sort of result (i.e. cheap talk closes but does not substantially extend the set of equilibrium allocations). The previous reasoning shows that the role of cheap talk as stated in Proposition 4 is sharply diminished, since in a very strong sense cheap talk is almost useless when a costly message (here the tax) is available.

7 Fully Revealing Equilibria

By de…nition, fully informative equilibria have exhaustive skeletons in which all types are represented. Moreover, Proposition 4 implies that the corresponding allocation is a universal skeleton, that is an equilibrium for any distribution F in $[\underline{\mu}; \bar{\mu}]$: The following proposition establishes that for any given $[\underline{\mu}; \bar{\mu}]$ there is a unique fully revealing equilibrium (or a unique universal skeleton), which we characterize in detail.

Proposition 6 There exists a unique fully revealing equilibrium. The tax rate is the unique solution to the ordinary di¤erential equation $\frac{t'}{1+¸} = ¡\frac{t ¡ \frac{´}{1+¸}}{t ¡ t_{SB}(\mu)}$ with the boundary condition $t(\bar{\mu}) = t_{SB}(\bar{\mu})$ if $¸ > 0$; and $t(\underline{\mu}) = t_{SB}(\underline{\mu})$ if $¸ < 0$: In particular:

1. Cheap talk is ine¤ective, and the strategy $t(¢)$ is strictly increasing and differentiable.
2. Consumption decreases with respect to μ:
3. If $¸ > 0$; the tax rate exhibits no distortion at $\bar{\mu}$: For other values, the tax rate is lower than the second-best tax rate and higher than $\frac{´}{1+¸}$:
4. If $¸ < 0$; the tax rate exhibits no distortion at $\underline{\mu}$: For other values the tax rate is higher than the second-best tax rate and lower than $\frac{´}{1+¸}$

Proof. See the Appendix. ■

On the role of cheap talk, it is clear from point 1 that in the fully informative equilibrium all the information is transmitted through the tax rate. When $¸ > 0$ ($¸ < 0$); the fully informative equilibrium allocation, compared to the second-best one, is characterized by taxes that are too low (or too high). Moreover, when $¸ > 0$ ($¸ > 0$) taxes are decreasing (increasing) with respect to the type.

The di¤erential equation assigns essential roles to the second-best tax and to $\bar{t} = \frac{´}{1+¸}$: The government (though under incentive constraints) tries to maximize welfare, hence to approach $t_{SB}(\mu)$ as nearly as possible. On the one hand, the closer the tax is to $t_{SB}(\mu)$, the higher social welfare, but the stronger the incentives for lying and the steeper the slope of the revealing tax schedule. On the other hand, for tax rates approaching the suboptimal $\bar{t}$; incentives to manipulate beliefs vanish, and the revealing tax schedule ‡attens.[21]

Here lies the origin of the distortion: credibility is gained by moving away from the optimal schedule. Note that if the government could commit *ex ante* to tax $\bar{t}$ whatever the state of nature μ; then telling the truth by means of cheap talk would be sequentially optimal since no credibility problem would arise. Unfortunately, this easy credibility would come at the price of a severe lack of e¢ciency!

[21] On the properties of $\bar{t}$; see Subsection 2.3.

Figure 3 shows the fully revealing tax rate for $´ = 1; \underline{µ} = 0; \bar{µ} = 1$ and $¸ = .3.$ Notice the indifference curves passing through the equilibrium value for $µ = .8$ and $µ = 1$. Figure 4 corresponds to $¸ = -.3$ (other parameters are equal to those in Figure 3). This illustrates the non-negligible size of the distortion.

Insert Figure 3 here.
Insert Figure 4 here.

The limited role of cheap talk can also be viewed another way. As $¸ \to 0;$ one can find a sequence of fully-revealing equilibrium allocations converging on the first-best where $t = ´$ for all $µ$: However, if we keep restricting the signal to be supported only by the tax, the first-best is not an equilibrium since no precise information on $µ$ can be transmitted: at the limit, cheap talk is indispensable, but very close approximations in which it is not used are available.

In CS, the most informative equilibrium Pareto-dominates, the others ex ante.[22] Austen-Smith and Banks (2000) find that this is not true when cheap talk and burning money to signal the type are used together. With our skeleton approach, it is relatively plain that the unique-fully informative equilibrium allocation need not be efficient with our approach in terms of skeletons. To see this, take an equilibrium and take its skeleton. The substance of Proposition 4 is that any economy that satisfies the restrictions on the conditional expected type in the intervals associated with the skeleton can implement it in equilibrium. If the mass of an interval where the distortion is substantial is sufficiently large, then the equilibrium is necessarily inefficient ex ante. More generally, given two skeletons, one more informative than the other (a finer partition in intervals), the less informative one can be made more efficient by choosing the distribution appropriately.

8 Conclusion

We have examined the conflict between providing incentives and transmitting information that arises when an informed and benevolent government combines linear taxes and information campaigns. Our model suggests that the government may have trouble gaining credibility for its actions and messages, even though its objective is aligned with the consumers'. The problem is that the government cannot commit to reveal information truthfully. Its instruments being imperfect, it has strong incentives to enhance their impact with providing biased information.

Depending on the kind of distortion that prevails in the fiscal system (i.e. whether the tax generates a "double dividend" or not), the government would like

[22] See Theorems 3 and 5 in CS, which say that both the sender and the receiver strictly prefer equilibrium partition with more steps.

to make consumers either pessimistic or optimistic about the e¤ect of consumption on individuals. In the likely case of positive marginal cost of public funds, if consumers were more optimistic, the government could set higher taxes without curbing consumption too severely, and the distortions created by preexisting taxes could be easily alleviated. Fuel taxes are an example: the government may not wish to stress the dangers of motor vehicles in order to preserve this ready source of revenue. The paradoxical consequence is that, at the fully revealing equilibrium, there is a bias towards excessively low taxes.

Our example of negative costs of public funds is quite informal. In France, SO_2 emissions are subjected to a "para…scal" tax: an independent agency is in charge of collecting the tax and also redistributes the proceeds in the form of subsidies for abatement e¤orts. Thus even if the Treasury has positive marginal costs of public funds, the agency may have negative marginal costs. It may be tempting to exaggerate local e¤ects (represented by our μ in the agents'—here the …rms'—programs) to economize resources wasted in the costly collection/redistribution process. As a result, at the fully revealing equilibrium, paradoxically the agency should be biased towards exaggeratedly high taxes.[23]

Another policy implication of the model is that information campaigns à la Crawford and Sobel are practically super‡uous when the social planner can also rely on costly signals. Quite simply, a costly tax is taken more seriously than mere propaganda, and information campaigns characterized by short phrases whose content is too vague to be veri…able ("smoking is harmful to health") are often of very limited e¢cacy. This result is in line with the empirical evidence of Bardsley and Olekalns (1999) on the impact of health warnings on cigarette packs. To be sure, we must distinguish here between information campaigns and "hard information". The former take the form of "free" advertising while the latter implies that the government collects and presents detailed scienti…c evidence in support of its views and that it employs other relays (academics, teachers, social workers, newspapers, etc.), in hopes that credibility will cease to be an issue. This process is slower, but presumably more e¤ective.

References

[1] Austen-Smith, David (1990): "Information Transmission in Debate," American Journal of Political Science, 124–152.

[2] Austen-Smith, David and Je¤rey S. Banks (2000): "Cheap Talk and Burned Money," Journal of Economic Theory, 91, 1, 1–16.

[23] As far as health policy is concerned, one can imagine that health autorities are driven to deliver cautious messages on the side e¤ects of antibiotics. In equilibrium, however, they must severely limit reimbursements (which is similar to imposing high taxex) to signal toxic drugs.

[3] Ballard, Charles L. and Don Fullerton (1992): "Distortionary Taxes and the Provision of Public Goods," Journal of Economic Perspectives, 6, 117–131.

[4] Bardsley, P. and N. Olekalns (1999): Economic Record, 75, 230, September, 225–40.

[5] Bovenberg, A. Lans and Frederick van der Ploeg (1994): "Environmental Policy, Public Goods and the Marginal Cost of Public Funds," Economic Journal, 104, 444–554.

[6] Brown, Gardner and David F. Layton (1996): "Resistance Economics: Social Cost and The Evolution of Antibiotic Resistance," Environment and Development Economics, 1, 3, 349–355.

[7] Crawford, Vincent P. and Joel Sobel (1982): "Strategic Information Transmission", Econometrica, 50, 6, 1431–1451.

[8] Drèze, Jean and Nicholas Stern (1987): "The Theory of Cost-Bene…t Analysis", in Handbook of Public Economics, II, edited by A.J. Auerbach and M. Feldstein, Elsevier Science Publishers.

[9] Farrell, Joseph and Matthew Rabin (1996): "Cheap Talk," Journal of Economic Perspectives, 10, 3, 103–118.

[10] Goulder, Lawrence H. (1995): "Environmental Taxation and the Double Dividend: A Reader's Guide," International Tax and Public Finance, 2, 157–83.

[11] Grossman, Gene M. and Elhanan Helpman (2001): Special Interest Politics, forthcoming.

[12] La¤ont, Jean-Jacques and Jean Tirole (1993): A Theory of Incentives in Procurement and Regulation, MIT Press, Cambridge and London.

[13] Levy, Stuart B. (1992): "The Antibiotics Paradox", Plenum Press, New York.

[14] Manelli, Alejandro M. (1996): "Cheap Talk and Sequential Equilibria in Signaling Games," Econometrica, 64, 917–942.

[15] Maskin, Eric and Jean Tirole (1992): "The Principal-Agent Relationship with an Informed Principal, II : Common Values," Econometrica, 60 , 1–42.

[16] Milgrom, Paul and John Roberts (1986): "Prices and Advertising Signals of Product Quality," Journal of Political Economy, 94, 796–822.

[17] Sandmo, Agnar (1975): "Optimal Taxation in the Presence of Externalities," Swedish-Journal-of-Economics, 77, 1, 86–98.

[18] Okeke, Iruka N., Laminkanra, Adebayo, Edelman, Robert (1999): "Socioeconomic and Behavioral Factors Leading to Acquired Bacterial Resistance to Antibiotics in Developing Countries," Emerging Infectious Disease, 5, 1, 18–27.

[19] Viscusi, W. Kip (1990): "Do Smokers Underestimate Risks?" Journal of Political Economy, 98, 6, 1253-1269.

A Appendix

A.1 Proof of Proposition 6

We establish the result in two steps. The first analyzes differentiable fully-revealing equilibria; uniqueness in this category is proved. The second step shows that any fully-revealing equilibrium is essentially identical to the differentiable one.

A.1.1 Differentiable Equilibrium

The analysis follows this plan: reasoning on local incentive compatibility, we find the ordinary differential equation satisfied by any fully-revealing equilibrium tax policy and we eliminate solutions with tax rates that do not fall between $\frac{\eta}{1+\lambda}$ and the second-best schedule (whichever is the higher); we check global incentive compatibility along the equilibrium policy; we search for off-equilibrium beliefs (i.e. associated with off-equilibrium tax rates) that discourage deviations. This gives a unique equilibrium.

Local Incentive Compatibility The government prefers $t(\theta)$ (and the implied $x(\theta)$) to $t(\theta + d\theta)$ and to $t(\theta - d\theta)$; taking limits we get

(20) $$x'U' + \lambda t'x - (\eta + \theta - \lambda t)x' = 0.$$

Given that the consumer's first-order condition is

(21) $$U' = \theta + t,$$

we can eliminate U' to get (after simplification)

(22) $$tx' = -\lambda t'x + (\eta - \lambda t)x'.$$

t and x being separable, (22) is easily integrated to give

(23) $$\forall \theta, \theta_0 : \frac{\frac{\eta}{1+\lambda} - t(\theta)}{\frac{\eta}{1+\lambda} - t(\theta_0)} = \left(\frac{x(\theta)}{x(\theta_0)}\right)^{\frac{1+\lambda}{\lambda}}$$

where θ_0 and $t_0 = t(\theta_0)$ are initial conditions. Equations (23) and (5) implicitly but entirely determine the solutions $t(\theta)$ and $x(\theta)$. In particular, x solves

$$(24)\qquad \left((\theta + \frac{\delta}{1+\lambda})x - 1\right) x^{\frac{1}{\lambda}} = \left((\theta_0 + \frac{\delta}{1+\lambda})x_0 - 1\right) x_0^{\frac{1}{\lambda}} = \text{Constant}$$

By differentiation, we get

$$(25)\qquad x' = \frac{\lambda x^2}{1 - (\delta + (1+\lambda)\theta)x}$$

The second-order condition is:

$$(26)\qquad 0 \geq x'^2 U'' + x'' U' + \lambda t'' x + 2\lambda t' x' - (\delta + \theta - \lambda t)x'',$$

while the derivative of the first-order condition is:

$$(27)\qquad 0 = x'^2 U'' + x'' U' + \lambda t'' x + 2\lambda t' x' - x' - (\delta + \theta - \lambda t)x'',$$

Simplifying (26) with (27) we get:

$$(28)\qquad x' \leq 0$$

Applying (28) to (25) and using (5), we can see that, when $\lambda > 0$, $x' \leq 0$ if and only if $t < \delta + \lambda\theta$ and when $\lambda < 0$, $x' \leq 0$ if and only if $t > \delta + \lambda\theta$.

Starting from (25) and (22), straightforward calculations prove that the differential equation satisfied by t is

$$(29)\qquad t' = \frac{\delta - (1+\lambda)t}{t - \delta - \lambda\theta}.$$

Global Incentive Compatibility The first- and second-order conditions exclude infinitesimal deviations. We check that discrete deviations are also precluded.

Let θ be the true value of the side-effects parameter. Using (5) we calculate the derivative of the government's utility with respect to $\hat\theta$ assuming that the government offers $t(\hat\theta)$, thereby inducing $x(\hat\theta)$:

$$(30)\qquad \begin{aligned} & x'(\hat\theta)U'[x(\hat\theta)] + (\lambda t'(\hat\theta)x(\hat\theta) - (\delta + \theta - \lambda t(\hat\theta))x'(\hat\theta)) \\ = & \frac{x'(\hat\theta)}{x(\hat\theta)} + \lambda t'(\hat\theta)x(\hat\theta) - (\delta + \theta - \lambda t(\hat\theta))x'(\hat\theta) \end{aligned}$$

From (22), we find that the following expression has the same sign as (30)

$$(31)\qquad x(\hat\theta)(t(\hat\theta) + \theta) - 1$$

Using (5), it follows that (30) is positive for $\hat\theta < \theta$ and negative for $\hat\theta > \theta$. This means that incentive compatibility is satisfied everywhere for equilibrium actions.

Uniqueness of the Differentiable Equilibrium We give the full reasoning for $\lambda > 0$. A symmetrical argument proves the proposition for $\lambda < 0$. We show now that the boundary condition $t(\bar{\theta}) = \alpha + \lambda\bar{\theta}$ is necessary.

Reasoning by contradiction, we show that there exist beliefs compatible with the equilibrium for off-equilibrium actions if and only if the condition is satisfied. Let $t(\cdot)$ be a solution to (29) such that $t(\bar{\theta}) < \alpha + \lambda\bar{\theta}$.[24] Given (29), either $t(\theta)$ is systematically below $\frac{\alpha}{1+\lambda}$ or $t(\theta)$ is strictly increasing. In any case $\max_\theta t(\theta) < \alpha + \lambda\bar{\theta}$; we choose an arbitrary (off-equilibrium) t in the interval $(\max_\theta t(\theta); \alpha + \lambda\bar{\theta})$ and we denote by $\hat{\theta}$ the associated belief. Now we prove that there always exists a type θ such that the government prefers policy $(t, \hat{\theta})$ to policy $(t(\theta); \theta)$.

From Proposition 2, we know that for each θ, the absolute best policy is $\frac{\alpha+\theta}{\lambda}$ for the tax rate and $-\frac{\alpha+\theta}{\lambda}$ for the belief; moreover, the second-best $(\alpha + \lambda\theta; \theta)$ is preferred to $(t(\theta); \theta)$. The convexity of the upper contours of the government's objective function implies that any policy in the triangle $\Delta(\theta) = ((t(\theta); \theta); (\alpha + \lambda\theta; \theta); (\frac{\alpha+\theta}{\lambda}; -\frac{\alpha+\theta}{\lambda}))$, except $(t(\theta); \theta)$, is strictly better than $(t(\theta); \theta)$ when θ is the type. It suffices now to check that $(t, \hat{\theta})$ is necessarily in $\Delta(\theta)$ for a certain $\theta \in [\underline{\theta}; \bar{\theta}]$. Indeed, $\cup_{\theta \in [\underline{\theta}; \bar{\theta}]} \Delta(\theta)$ contains (a) the triangle $((\alpha + \lambda\underline{\theta}; \underline{\theta}); (\alpha + \lambda\bar{\theta}; \bar{\theta}); (\frac{\alpha+\bar{\theta}}{\lambda}; -\frac{\alpha+\bar{\theta}}{\lambda}))$; and (b) the policies between $(t(\theta); \theta)$ and $(t_{SB}(\theta); \theta)$ for $\theta \in [\underline{\theta}; \bar{\theta}]$. Provided $\frac{\alpha+\bar{\theta}}{\lambda}$ is larger than $\alpha + \lambda\bar{\theta}$, then $(t, \hat{\theta})$ is either in (a) or (b) in the latter union, hence the existence of a θ for which the deviation is desirable. Given that $\frac{\alpha+\bar{\theta}}{\lambda} > \alpha + \lambda\bar{\theta}$, the proof is complete. The consequence is that $t(\bar{\theta}) = \alpha + \lambda\bar{\theta}$ (no distortion at the top).

Now we prove that associating belief $\bar{\theta}$ to any tax rate above $t(\bar{\theta})$ does not induce deviations. The value to the government of type θ of imposing $t > t(\bar{\theta})$, thereby inducing belief $\bar{\theta}$, is: $-\log(\bar{\theta} + t) - \frac{\alpha+\theta-\lambda t}{\bar{\theta}+t}$. The root of the derivative with respect to t is $\alpha + \theta - (1 - \lambda)\bar{\theta}$ which is lower than $\alpha + \lambda\bar{\theta} = t(\bar{\theta})$. The value being decreasing with respect to t over $[t(\bar{\theta}); +\infty[$, $t(\bar{\theta})$ is a better move than any $t > t(\bar{\theta})$. Given that equilibrium actions are incentive-compatible, neither $t(\bar{\theta})$ nor t is desirable, compared to $t(\theta)$. By the same reasoning, we can check that, if for $t < t(\underline{\theta})$, beliefs are $\underline{\theta}$, then t is not attractive: for a government of type θ, the value of imposing $t < t(\underline{\theta})$ is smaller than that of imposing $(t(\underline{\theta}); \underline{\theta})$.

We conclude that the unique revealing allocation found is an equilibrium.

A.1.2 Uniqueness in General

Let us take a fully-revealing equilibrium. Given that the government's preferences, for constant beliefs, are single-peaked with respect to t (a direct consequence of the convexity in Proposition 2), and given the value of its equilibrium strategy, there exist a maximum of two tax rates per θ, $t_L(\theta)$ and $t_U(\theta)$, both

[24] For $\lambda > 0$, we already excluded that the tax rate could be higher than the second-best tax rate in the preceding subsection.

being suboptimal (as compared to the second-best) when they are different. More precisely, $t_L(\theta) \le \eta + \lambda\theta \le t_U(\theta)$). The theorem of the maximum ensures that the value of the government's equilibrium strategy is continuous with respect to θ, therefore functions $t_L(\cdot)$ and $t_U(\cdot)$ are continuous with respect to θ. We denote by Θ_L and Θ_U the subsets of $[\underline{\theta}, \overline{\theta}]$ leading to a move in the lower and the upper selection respectively. Notice that $\Theta_L \cup \Theta_U = [\underline{\theta}, \overline{\theta}]$ but $\Theta_L \cap \Theta_U \neq \emptyset$ if mixed strategies are used. For fixing ideas, the following reasoning assumes that $\lambda > 0$.

The first step is to prove that Θ_U is not dense in any interval of $[\underline{\theta}, \overline{\theta}]$. We reason by contradiction: take J an interval in $[\underline{\theta}, \overline{\theta}]$ in which Θ_U is dense. Take $\theta_0 \in J$, and a strictly monotonic sequence $(\theta_n)_{n\ge1}$ in Θ_U converging to θ_0. We prove that for all sequences $(\theta_n)_{n\ge1}$, $\lim_{n\to\infty} \frac{t_n - t_0}{\theta_n - \theta_0} = \frac{\eta - (1+\lambda)t_0}{t_0 - \eta + \lambda\theta_0}$, where t_n denotes $t_U(\theta_n)$. Indeed, incentive constraints (θ_n should not mimic θ_0, and vice-versa) imply that:

$$(32) \qquad -\log(\theta_n + t_n) - \frac{\eta + \theta_n - \lambda t_n}{\theta_n + t_n} \ge -\log(\theta_0 + t_0) - \frac{\eta + \theta_n - \lambda t_0}{\theta_0 + t_0}$$

$$(33) \qquad -\log(\theta_n + t_n) - \frac{\eta + \theta_0 - \lambda t_n}{\theta_n + t_n} \le -\log(\theta_0 + t_0) - \frac{\eta + \theta_0 - \lambda t_0}{\theta_0 + t_0}$$

Therefore, taking a first-order approximation, and multiplying by $(\theta_0 + t_0)^2$ yields

(34)

$$0 \ge ((1+\lambda)t - \eta)(\theta_n - \theta_0) + (t - \eta - \lambda\theta)(t_n - t_0) + o(\theta_n - \theta_0) + o(t_n - t_0)$$

(35)

$$0 \le ((1+\lambda)t - \eta)(\theta_n - \theta_0) + (t - \eta - \lambda\theta)(t_n - t_0) + o(\theta_n - \theta_0) + o(t_n - t_0)$$

The limit of the rate of variations is the same for all sequences, which implies that t_U is differentiable at θ_0, hence differentiable on interval J.

A solution of the differential equation (29) situated above the second-best taxes is not incentive compatible at any point, because the second-order condition is never satisfied. We can conclude that strategy t_U is not-incentive compatible, and that interval J does not exist.

It is now easy to conclude that Θ_L is dense in $[\underline{\theta}, \overline{\theta}]$: being the complementary set (in an interval) of a set Θ_U which is not dense anywhere, Θ_L is dense. Consequently, t_L satisfies the differential equation (29) in a dense subset of $[\underline{\theta}, \overline{\theta}]$, which implies that it does so everywhere. The lower selection is necessarily equal to the unique differentiable equilibrium strategy, since we can apply to $t_L(\cdot)$ the reasoning suited for differentiable equilibria.

It remains now to prove that Θ_U contains a finite number of points. We take θ_1 and $\theta_2 \in \Theta_U$ (where $\theta_1 \neq \theta_2$) with corresponding tax rates t_1 and t_2. Let us denote by $t_i(\cdot)$ $(i = 1, 2)$ the solution to (29) with maximal definition domain passing through t_i at θ_i. Note that either $t_1(\cdot)$ and $t_2(\cdot)$ are the same, or one is systematically above the other, because, according to the Cauchy-Lipschitz Theorem, two different solutions to differential equation (29) never cross.

Assume for …xing ideas that $t_2(¢)$ is above $t_1(¢)$: (a) If the two curves are close enough to each other, $t_2(µ_1)$ is de…ned and is larger than t_1: Notice that $t_2(µ_1)$ is closer to the second-best than t_1: Our study of the incentives when taxes are above the second-best shows that solutions to the di¤erential equations minimize welfare (the second-order condition is violated everywhere): when the type is $µ_1$; t_2 with belief $µ_2$ is preferred to $t_2(µ_1)$ with belief $µ_1$: By transitivity, t_2 is preferable to t_1 when the true type is $µ_2$: This is in contradiction with incentives. (b) If there is an in…nite number of types in $£_U$; we can always exhibit $µ_1$ and $µ_2$ which are close enough to each other to apply reasoning (a). We conclude that $£_U$ contains a …nite number of points.

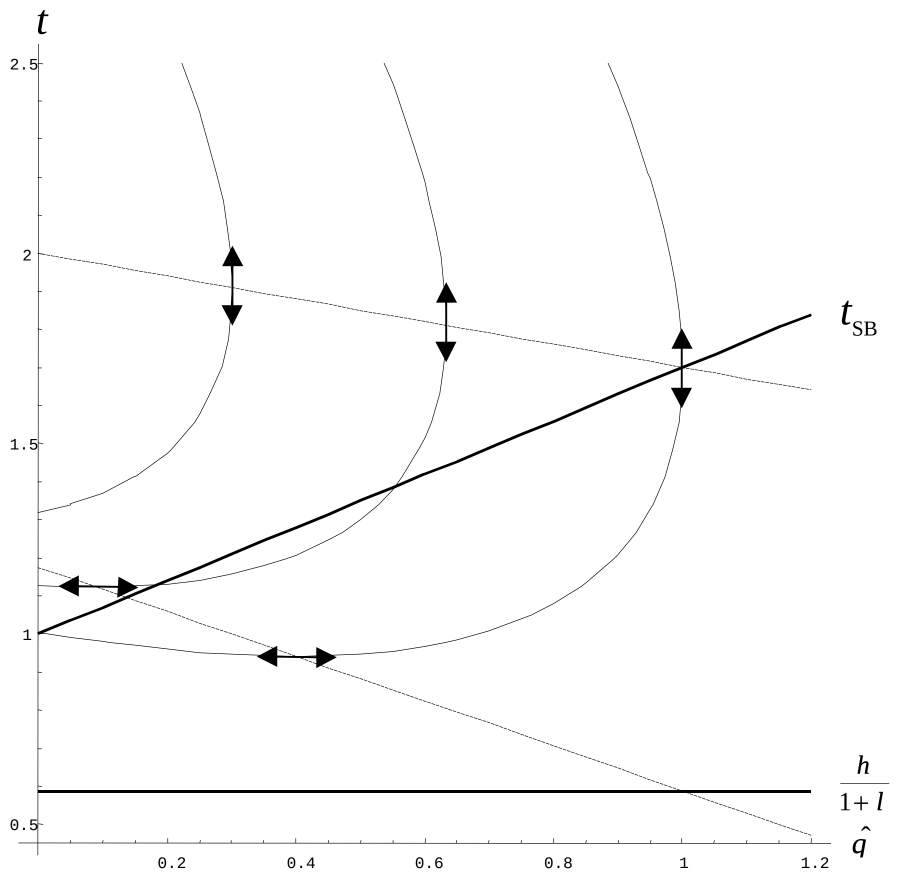

Figure 1: government indifference curves with λ>0.

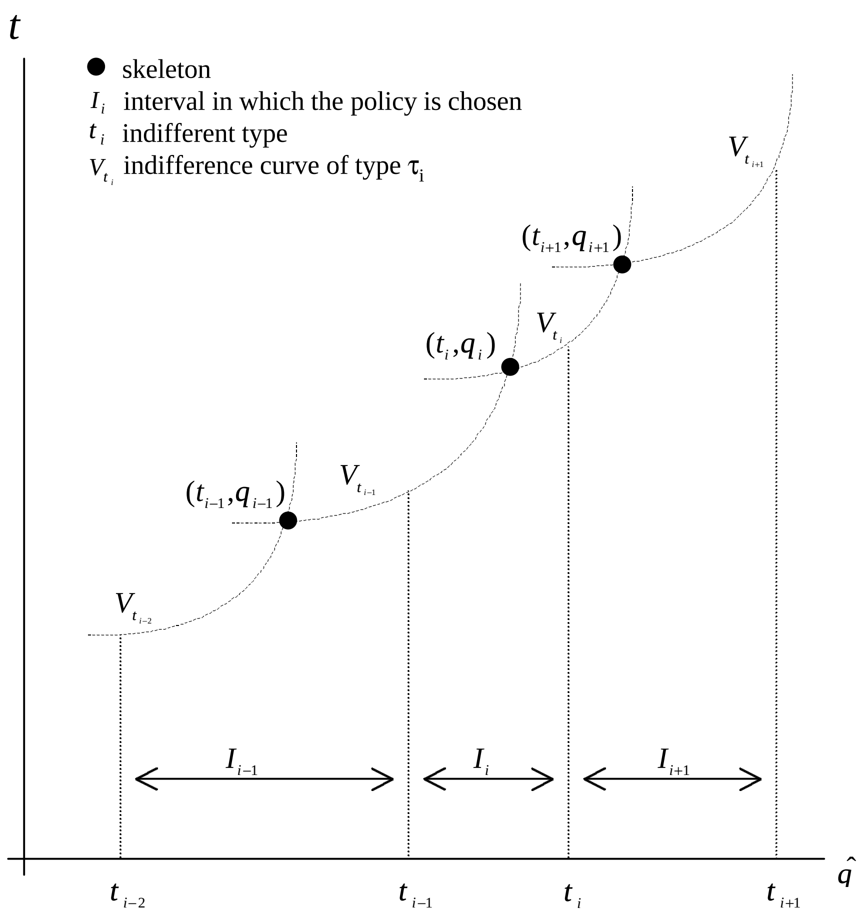

Figure 2: the skeleton.

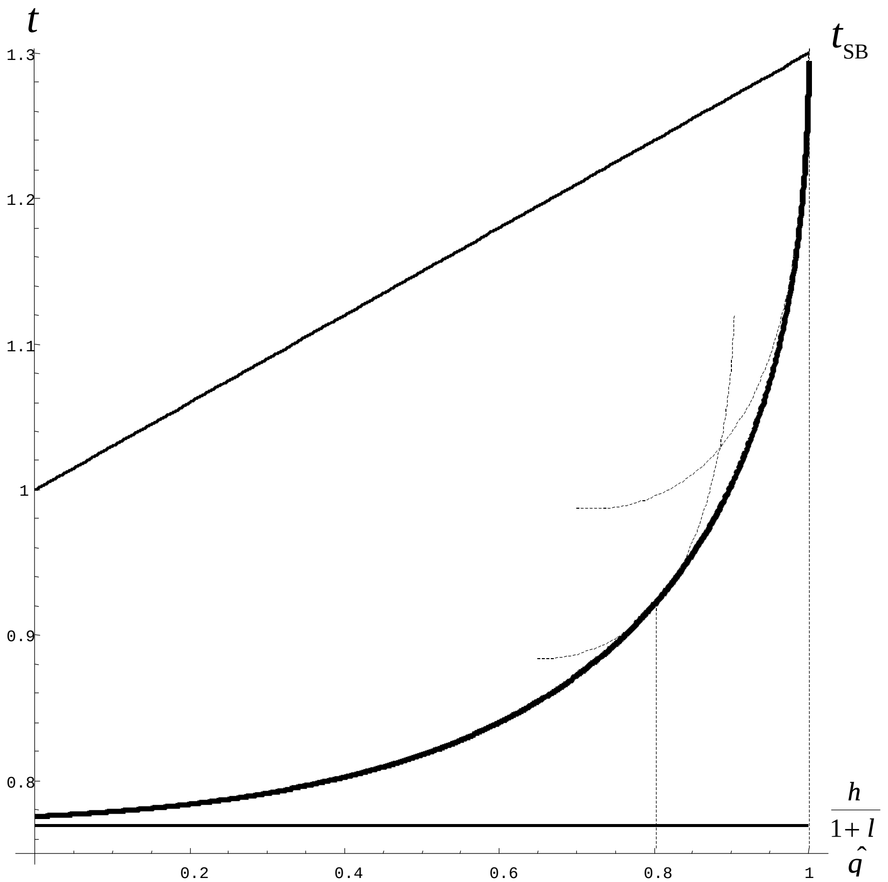

Figure 3: the fully-revealing equilibrium with λ>0.

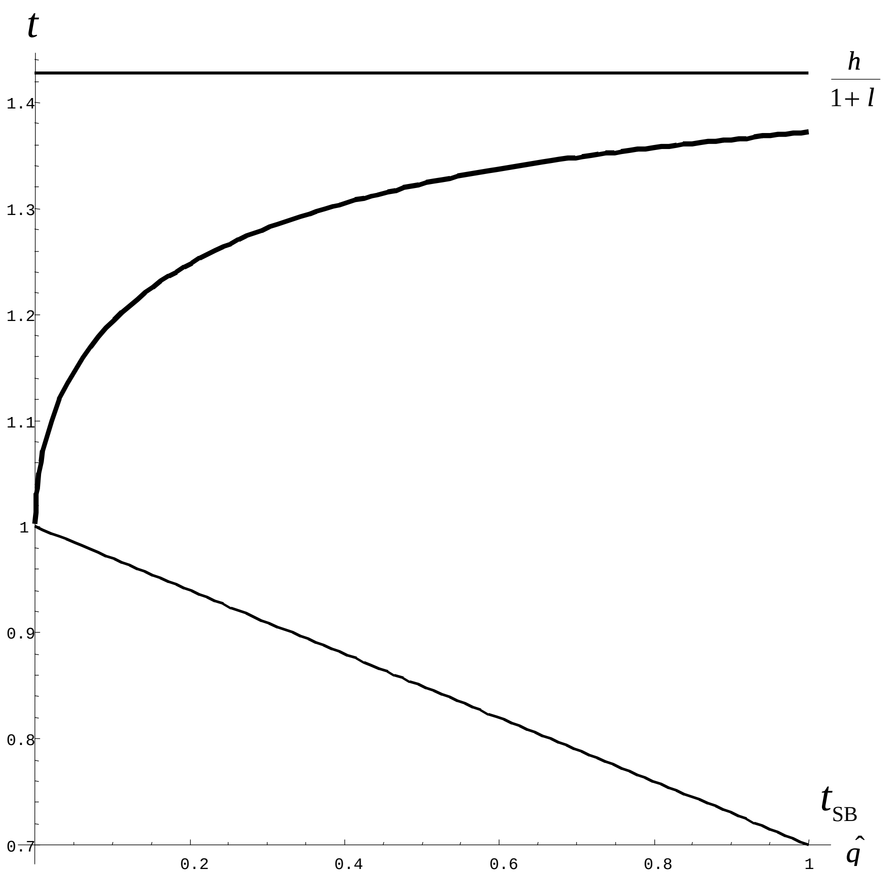

Figure 4: the fully-revealing equilibrium with λ<0.