The signaling effect of tax policy[*]

Francesca Barigozzi[†]   Bertrand Villeneuve[‡]

February 2004

Abstract

The paper focuses on the signaling value of a tax when agents are less informed on the effect of their consumption than the policy-maker. When he chooses the tax, the policy-maker optimizes both the incentive effect and the effect on beliefs. We show that optimal taxes under symmetric information are not implementable under asymmetric information. The tax distortion required for credibility depends on the marginal cost of public funds and brings to undertaxation.

Keywords: tax policy, marginal cost of public funds, information bias, signaling.

JEL: I18, H30, D82.

1 Introduction

Environmental and health policy issues often involve commodities combining *external* effects (pollution, costs to the healthcare system) and *individual* effects (future costs or benefits related to own consumption). Once reasonable effort to inform the public has been undertaken by the authorities, the perception of individual effects may remain fuzzy. This paper investigates the triple task of tax policy (raising public funds, correcting external effects, and signaling missing information to individuals) and particularly the policy-maker's conflict between budgetary and informational objectives.

Policy design is particularly delicate for commodities that are detrimental to consumers' welfare through multiple channels. Broad-spectrum antibiotics display (apart from the obvious benefits) a double negative impact. At the individual level they clear the way to opportunistic infection by more resistant germs; at the societal level, they enhance the resistance of the germs involved in contagious diseases.[1] In the case of tobacco, we have on the one hand, diseases related to individual consumption, and on the other hand the social damage from passive smoking and the cost to the healthcare system. High levels of alcohol consumption exhibit the same double negative effect. Risky driving exposes drivers themselves and also others to accidents, whose costs are partly private and partly social.[2] For a discussion on environmental policies, the reader can see the conclusion of the paper.

The appropriate tax level depends on the constraints faced by the policy-maker. The Pigouvian tax, set at the marginal damage, is efficient only if the economy is otherwise undistorted. When the policy-maker needs revenue to finance public activities, the Ramsey rule stipulates that inelastic demands must be taxed more. Both arguments are combined in Sandmo (1975): for commodities causing externalities, a simplified lesson is that the optimal tax is the sum of the Pigouvian term (the marginal external cost) and the Ramsey term (the inverse of the elasticity of demand), both corrected by factors accounting for the level of distortion in the economy. If the budget constraint of the policy-maker is sufficiently tight and demand sufficiently inelastic, optimal taxes are well above the Pigouvian level, without contradicting orthodox theory. Increasing need for public expenditures may explain increasing taxes over the years.

This view applies to excise taxes on tobacco. A vast empirical literature has tried to establish the exact magnitude of smoking externalities. Results

---

[1]Some broad-spectrum antibiotics decrease the individual's immunological response; as a consequence, he may get new diseases. For many antibiotics based on penicillin (used to treat, e.g., bronchitis, otitis and tonsillitis), possible side effects 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 effects.

[2]For an empirical analysis of driving externalities and Pigouvian taxes, see Edlin and Karaca-Mandic (2003).

are subject to considerable controversy, as Chaloupka and Warner (2000) review; nevertheless departure from the Pigouvian rule is well documented and there is an agreement that tobacco taxes are typically higher than marginal external effects. Theoretically, this simply means than the Pigouvian term could be well behind the Ramsey one, a possibility that is seldom mentioned in empirical studies.

The classical approaches assume that economic agents have perfect perception of all individual effects of their consumption. We relax this assumption to consider the policy-maker as the informed party: he can collect and process detailed evidence and data and rely on expert advice about all the aspects of health and environmental issues relevant to economic agents. Moreover, given the public good nature of information, one of the policy-maker's objectives is obviously to transmit it to consumers. However, once all direct instruments have been employed (education, information campaigns, etc.), consumers may remain, by and large, underinformed. By exploring the possibility of simultaneously providing signals and incentives with taxes, we offer an interpretation of tax policy for which the Ramsey-Sandmo theory is insufficient.[3]

Our baseline is that the policy-maker wishes to promote the public interest. The marginal cost of public funds is a crucial ingredient of the analysis; like the externality, it enters in the social welfare function but is neglected by agents at the moment of choice. The policy-making process is analyzed as a game in which the government wants to influence consumers' behaviors through tax policy, consumers being rational and Bayesian. In our view, whenever some imperfection impedes economic efficiency, taxes must be understood as a metaphor of the entire comprehensive policy package, also comprising contracts, bans, standards and norms.[4,5]

The policy-maker is aware of the signaling property of the chosen policy and will take any opportunity to distort the tax provided the total effect on consumption (through incentives and beliefs) is favorable to social welfare.[6] Consumers anticipate this threat and, at the equilibrium, the tax is distorted away from the Ramsey-Sandmo level.[7] Indeed, the Ramsey-Sandmo tax "reveals" the parameter that the consumer does not precisely perceive: the parameter affects the

---

[3]In the case of tobacco, which is a mere example, our explanation is based on the necessity to improve consumers' information, and is complementary to that in Gruber and Köszegi (2003) which relies on a complete information model of addiction.

[4]Anti-smoking measures indeed comprise a large span of instruments, like advertising bans, prohibition of sales to minors, smoke-free areas, etc.

[5]Cheap talk has been recently used to study warning labels by Hanson (2003). In his model an informed regulator is empowered to require one or several warning labels on a drug, and may also be empowered to ban the drug. This selection between good (not banned) and bad (banned) drugs is based on relatively coarse policy instruments (cheap talk and bans).

[6]A similar form of paternalism is found in Benabou and Tirole (2004).

[7]Although the impossibility to set taxes at their second-best levels also arises when no external effects exist, situations with externalities present the most interesting examples and provide a primary justification for the utilization of the tax.

elasticity of demand, which in turn affects the impact of the tax. By choosing a tax slightly below the second-best, the policy-maker signals smaller individual effects, which enhances consumption and consequently the tax revenue; the effect on consumers' welfare is negative but of second order whereas the effect on the policy-maker's budget is positive and of first order.

The analysis of the fully revealing equilibrium shows that the above conflict results in taxes lying between the Pigouvian and the Ramsey-Sandmo taxes. The informative role of the tax is not, per se, a cause of overtaxation, quite the contrary. However, we show that a small difference in the severity of individual effects causes a large difference of the equilibrium tax. In a sense, an amplification of minor differences can be seen as an important characteristic of informative taxation that the theory of the second-best cannot justify.

A technical remark is necessary at this stage. The level of welfare losses is not monotonic with the level of the tax, therefore signaling costs have an unusual structure that classical models do not exhibit. In the proof of the uniqueness of the fully revealing equilibrium, Mailath (1987) is inapplicable.

Plan   Section 2 presents the terms of the policy dilemma. Section 3 defines the equilibrium. After a few results on the structure of the government's preferences (Section 4), we characterize the game's unique fully-revealing equilibrium (Section 5). All the proofs are relegated to the Appendix.

2   The Model

2.1   The Consumers

Consumers live two periods and their period-2 utility is negatively affected by period-1 consumption x_1. Preferences can be written as

$$(1) \qquad\qquad U[x_1] + x_2 - \theta x_1 - \eta \overline{x}_1,$$

where U is increasing, concave and twice continuously differentiable over $\mathbb{R}^+$. 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 $-\eta \overline{x}_1$ indicates the negative **externality** that depends on $\overline{x}_1$, i.e. *average* period-1 consumption in the economy. The intensity η is supposed

to be known to all the agents,[8] nevertheless consumers do not internalize the social consequences of x_1.

Individual consumption is assumed to be unobservable. This implies that the government is constrained to impose a linear tax on commodity x, denoted by t. The representative consumer solves

$$(2) \qquad \begin{cases} \max\limits_{x_1,x_2} \quad E\left[U[x_1] + x_2 - \theta x_1\right], \\ \text{s.t.} : (p_1 + t)x_1 + p_2 x_2 = W. \end{cases}$$

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. We assume that W is large enough for x_1 and x_2 to be strictly positive.

We normalize p_2 to 1 and p_1 to 0 without loss of generality since the support of θ can be translated to account for the (exogenous) price. 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[9]

$$(3) \qquad \max_x \quad E\left[U[x] + W - (\theta + t)x\right].$$

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

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

With $U = \log$, the previous reads

$$(5) \qquad x^*[E\theta, t] = \frac{1}{E\theta + t}.$$

2.2 Social Welfare and the Marginal Cost of Public Funds

The policy-maker maximizes social welfare, i.e. the consumers' utility once the externality, the fiscal impact of the tax policy and the exact value of side effects are taken into account. The policy-maker is constrained by the fact that the tax t must be linear, by preexisting distortions (the marginal cost of public funds) and by his supposed inability to commit to a policy that informs truthfully on the value of side effects θ.

The policy-maker is benevolent in that he evaluates consumption in consumers' best interest, but he practices a variety of paternalism since he does

---

[8] This is the simplest assumption since with our specification, consumption does not depend on η.

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

not value truthful information *per se*, and would deceive consumers provided this induces "better" behavior.

All consumers being identical, in equilibrium $\bar{x} = x$, and the policy-maker's objective function can be represented, for $t = 0$, as

$$(6) \qquad U[x] + W - (\theta + \eta)x + S - (1 + \lambda) R,$$

where S is the consumer's surplus from public expenditures R (R and S are fixed). The government raises the exogenous amount R with general taxation (income-, capital-tax, or other levies) at the welfare cost $(1 + \lambda) R$, with $\lambda > -1$. The parameter λ is the "shadow cost" of public funds; it represents the distortion due to the raising of fiscal revenue.[10]

When the government introduces a tax t on good x and revenue tx is devoted to reducing preexisting taxes (revenue accruing from other fiscal sources is $R - tx$), (6) becomes, after simplification,

$$(7) \qquad U[x] + W - (\theta + \eta - \lambda t)x + S - (1 + \lambda)R.$$

Comparing (3) and (7), we see that the policy-maker's objective function differs from the consumer's in three ways: the government has superior information on θ and he takes into account both η and λ. In fact the consumer does not internalize that his contribution tx will replace revenue from other distortionary taxes in the economy.

Dropping constant terms, we get a reduced form of the policy-maker's objective function

$$(8) \qquad SW[x, t, \theta] \equiv U[x] - (\theta + \eta - \lambda t)x.$$

Fiscal revenue tx being small compared to R, we assume for simplicity that λ remains constant whatever t. Though 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,[11] we assume throughout the paper that $\lambda \geq 0$.[12]

The assumption $\lambda \geq 0$ implies that preexisting taxes inflict a welfare cost larger than R and relates to the debated "double dividend" effect (Pearce, 1991). According to this literature, a revenue-neutral substitution of environmental taxes for other taxes might offer a double dividend: not only does it improve the environment but it also reduces the costs of the tax system through cuts in distortionary taxes (see Goulder, 1995). To intuit this result, assume that x

---

[10]In the Ramsey model of taxation, λ would be the (endogenous) Lagrange multiplier associated with the government's budget constraint. For partial equilibrium models (where λ is exogenous), see, e.g., the theory of regulation (Laffont and Tirole, 1993).

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

[12]Results in the case $\lambda < 0$ are available in a previous version of the paper (see Barigozzi and Villeneuve, 2003).

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, which attenuates distortions. 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.

2.3 Constrained Efficient Allocations

The *first-best allocation* is defined as the allocation that maximizes consumers' utility when θ is known, η is internalized, and rising funds does not generate distortion ($\lambda = 0$). This gives $x_{\mathrm{FB}}(\theta) = U'^{-1}[\theta + \eta]$. This allocation can be decentralized even for unobservable x with the Pigouvian linear tax $t = \eta$.

We define the *second-best* allocation as the best the policy-maker can achieve, with consumers perfectly informed on θ, when he is constrained to a linear tax on x and the marginal cost of public funds is not zero. This can be written as follows

(9)
$$\begin{cases} \max_t \quad U[x] - (\eta + \theta - \lambda t)x, \\ \text{s.t.} : x = U'^{-1}[\theta + t], \end{cases}$$

where the constraint on x corresponds to consumers' reaction function (4) when θ is known. Straightforward rearrangement of the first-order condition leads to the Ramsey-Sandmo rule[13]

(10)
$$\frac{t_{\mathrm{SB}}(\theta) - \overline{t}}{t_{\mathrm{SB}}(\theta)} = \frac{\lambda}{1 + \lambda} \frac{1}{\varepsilon},$$

where $\varepsilon \equiv -\frac{\partial x^*}{\partial t}\frac{t}{x^*}$ is the tax elasticity of demand, and $\overline{t} = \frac{\eta}{1+\lambda}$. Note that, in the second-best tax, the Pigouvian term $\overline{t}$ (the marginal external cost) and the Ramsey term (the inverse of the elasticity of demand) are both corrected for the tax distortion λ.

To pursue the analysis, we need two additional restrictions on U that we maintain throughout the paper. Let $\pi(x) = \frac{-xU'''[x]}{U''[x]}$ be the index of relative prudence of U. The first condition implies that program (9) is concave:

Concavity. π is not restricted for $\lambda = 0$. If $\lambda > 0$, then $\pi \leqslant 2 + 1/\lambda$.[14]

The second one facilitates the characterization of the second-best policy and requires that $\frac{-x'[s]}{x[s]}$ (relative variation of demand) decreases as the price increases (the proof is direct):

---

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

[14] Denote U'^{-1} by X. The first- and second-order conditions of (9), $t - \overline{t} = -\frac{\lambda}{1+\lambda}\frac{X}{X'}$ and $(t-\overline{t})X'' + \frac{1+2\lambda}{1+\lambda}X' \leqslant 0$, yield $\frac{-\lambda X X''}{X'} \leqslant -(1+2\lambda)X'$. Conditions on concavity directly follow.

Monotone relative variation (MRV). $\pi \geqslant 1$.

Both conditions are satisfied by a fairly large range of function, including log, for which $\pi = 2$, and CRRA utility functions $\frac{x^{1-\alpha}}{1-\alpha}$ with $0 \leqslant \alpha = \pi - 1 \leqslant 1 + 1/\lambda$ if $\lambda > 0$. Moreover, for small degrees of distortion (as measured by λ), the condition on concavity becomes extremely loose.

Proposition 1 (Second-best Tax) *Under Concavity and MRV, $t_{SB}(\theta)$ increases for $\lambda > 0$. If $\lambda = 0, t_{SB}(\theta) = \eta$.*

The intuition is that, larger side effects discouraging consumption more, and small consumption being relatively inelastic, fiscal revenue is partly preserved by larger taxes: the higher the *individual* risk, the higher the tax.

With $U = \log$, we have

$$(11) \qquad \begin{cases} t_{\mathrm{SB}}(\theta) = \eta + \lambda\theta, \\ x_{\mathrm{SB}}(\theta) = \frac{1}{\eta+(1+\lambda)\theta}. \end{cases}$$

Proposition 1 proves that, when the objective of the policy-maker is twofold (correcting the external effect and rising funds), optimal taxes are above the Pigouvian level. A simple and often overlooked explanation of the observed gap between cigarette taxes and tobacco externality is the existing distortions in the fiscal system. Moreover, an increasing need of revenue on the part of the policy-maker can explain increasing taxes over the years.

3 The Influence Game

The timing of the model is as follows: first the policy-maker observes θ, then he chooses his tax policy, and finally the consumer, observing the tax, updates his beliefs on θ and chooses his consumption level.

The tax has the two-fold role of providing incentive and information. After observing the policy, the consumer updates his priors which are then denoted by $\mu(t)$, with $\mu(t) \in \Delta([\underline{\theta}, \overline{\theta}])$, the set of probability distributions over $[\underline{\theta}, \overline{\theta}]$. We denote $E(\theta|t)$ by $\widehat{\theta}(t)$.

Definition 1 *A perfect Bayesian equilibrium (PBE) in pure strategies of the game is a strategy $\mathcal{P}$ mapping $[\underline{\theta}, \overline{\theta}]$ into $\mathbb{R}$, a belief μ mapping $\mathbb{R}$ into $\Delta([\underline{\theta}, \overline{\theta}])$, and a consumption rule x^* mapping $\Delta([\underline{\theta}, \overline{\theta}]) \times \mathbb{R}$ into $\mathbb{R}_+$ such that:*

1. Policies are optimal given beliefs and consumption function:

$$(12) \qquad \text{for all } \theta \in [\underline{\theta}, \overline{\theta}], \mathcal{P}(\theta) = \arg\max_t SW[x^*[\widehat{\theta}(t), t], t, \theta].$$

2. *Beliefs are rational given equilibrium policy:*

$$(13) \qquad \mu(\theta|t) \equiv \frac{\mathbb{I}_{\{\mathcal{P}(\theta)=t\}} \cdot f(\theta)}{\int_{\underline{\theta}}^{\overline{\theta}} \mathbb{I}_{\{\mathcal{P}(s)=t\}} \cdot f(s) \, ds},$$

$\mathbb{I}$ *being the indicator function. Beliefs for off-equilibrium actions are not restricted.*

3. *Consumption is optimal given the observed policy and the beliefs:*

$$(14) \qquad x^*[\widehat{\theta}(t), t] = \arg\max_x \int_{\underline{\theta}}^{\overline{\theta}} [U[x] - (\theta + t)x]\mu(\theta|t)d\theta.$$

4   Policy-maker's Preferences

In any equilibrium, the larger the side effects, the lower the social welfare: the consumer's rationality prevents the policy-maker from transforming lead into gold by clever communication strategies. To see that, let θ_1 and θ_2 be two possible states of the world, t_1 and t_2 two equilibrium policies, and x_1 and x_2 the consumption levels induced. If $\theta_1 < \theta_2$, then $U[x_2] - (\eta + \theta_1 - \lambda t_2)x_2 \geqslant U[x_2] - (\eta + \theta_2 - \lambda t_2)x_2$. On the other hand, the incentive constraint of the type-θ_1 policy-maker reads: $U[x_1] - (\eta + \theta_1 - \lambda t_1)x_1 \geqslant U[x_2] - (\eta + \theta_1 - \lambda t_2)x_2$. We get: $U[x_1] - (\eta + \theta_1 - \lambda t_1)x_1 \geqslant U[x_2] - (\eta + \theta_2 - \lambda t_2)x_2$. Thus the social planner's payoff decreases with respect to the side effects. In other words, we preclude any perverse mechanism whereby propaganda can make less desirable states of the world (larger side effects) preferable.

We analyze now the policy-maker's incentive to distort taxes, i.e. the reasons why consumers are likely to be suspicious of the policy-maker's policy and how suspicious.

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

Policies are restricted to induce finite consumption, thus feasible policies are such that $t + \widehat{\theta} > 0$. Indifference curves are not monotonic with respect to the tax or the belief. The following proposition gives us the properties required to proceed with the analysis of incentive compatibility.

Proposition 2 (Single-crossing) *For all $(\widehat{\theta}, t)$, the tangent of the indifference curve for type θ passing through $(\widehat{\theta}, t)$ turns continuously clockwise if $\lambda > 0$ as θ increases. This has two related implications: 1. Indifference curves related to two different types cross once at most; 2. Upper contours are never interior (the branches of indifference curves go to infinity or cross the boundary $\widehat{\theta} + t > 0$).*

Figure 1 shows the policy-maker's indifference curves for $U = \log$, $\eta = 1$, $\underline{\theta} = 0$, $\bar{\theta} = 1$, $\lambda = .7$ and $\theta = 1$. For all θ, tangents to indifference curves are horizontal along the straight line $(1 + \lambda)t + \widehat{\theta} = \eta + \theta$, and vertical along the straight line $t + (1 - \lambda)\widehat{\theta} = \eta + \theta$.

<div style="text-align:center; border:1px solid; display:inline-block;">Insert Figure 1 Here</div>

We shall verify whether second-best tax policy is implementable in a Bayesian equilibrium. Suppose the consumer *thinks* that the policy-maker is playing the second-best strategy. The tax schedule t_{SB} being invertible, if $t_{\mathrm{SB}}(\theta)$ is imposed, the individual can infer θ unambiguously. In fact, the policy-maker faces strong incentives to provide biased information.

Corollary 1 *The second-best allocation is never an equilibrium if $\lambda > 0$.*

In fact, at $(\theta, t_{\mathrm{SB}}(\theta))$, for all θ, the tangent of the indifference curve of the government is vertical (see proof of Proposition 2): small changes in the tax have only second-order effects, whereas small changes in the beliefs have first-order effects on the policy-maker's objective. Consequently, any policy close to the second-best but with $\widehat{\theta} < \theta$ is preferred: the second-best allocation is not incentive-compatible, which confirms that the policy-maker has an incentive to bias tax policy. Next section characterizes the fully informative equilibrium.

Note in contrast that when $\lambda = 0$, the second-best allocation is not informative on side effects ($t_{\mathrm{SB}} = \eta$). The tax perfectly internalizes the externality, but consumers' priors are not updated and $x = U'^{-1}[E\theta + \eta]$. The final allocation will imply under- or over-consumption with respect to the first-best depending on the prior distribution and on the value taken in effect by θ.

Remark that the externality η is not essential for incentives to bias information. It serves mainly for the realism of the application to health and environment issues.

5   Fully Revealing Equilibria

Fully informative equilibria are useful for exposing the tradeoff between informing consumers and setting incentives. Rational consumers anticipate potential manipulations and we show that the policy-maker is obliged, to gain credibility, to distort taxes towards lower levels.

We derive in this section the structure of the fully revealing equilibrium and its uniqueness. The result is not an application of Mailath (1987) since signals have non-monotonic costs (the minimal cost being reached at t_{SB}) and an appropriate proof has to be written in full.

Proposition 3 *There is a unique fully revealing equilibrium. The tax rate is the unique solution to the ordinary differential equation defining $t(\theta)$:*

$$(15) \qquad t' = -\frac{t - \bar{t}}{t - \frac{\eta}{1+\lambda} + \frac{\lambda}{1+\lambda}\frac{U'^{-1}[\theta+t]}{[U'^{-1}]'[\theta+t]}}$$

with the boundary condition $t(\bar{\theta}) = t_{SB}(\bar{\theta})$. The tax policy $t(\cdot)$ is strictly increasing and differentiable; consumption decreases with respect to θ. For values of θ lower than $\bar{\theta}$, $\bar{t} < t(\theta) < t_{SB}(\theta)$.

Figure 2 shows the fully revealing tax rate for $U = \log, \eta = 1, \underline{\theta} = 0, \overline{\theta} = 1$ and $\lambda = .3$. The differential equation simplifies to

$$(16) \qquad \frac{t'}{1+\lambda} = -\frac{t - \bar{t}}{t - t_{\mathrm{SB}}(\theta)}.$$

> Insert Figure 2 Here

The solution to the differential equation lies between the optimal tax $t_{\mathrm{SB}}(\theta)$ and $\bar{t} = \frac{\eta}{1+\lambda}$ (which appeared in the Ramsey-Sandmo equation (10)). On the one hand, when the tax is close to the optimal t_{SB}, incentives to "improve" beliefs are strong and the slope of the revealing tax schedule is very steep. On the other hand, as the tax approaches the suboptimal $\bar{t}$, incentives to manipulate beliefs vanish, and the revealing tax schedule flattens. As the policy-maker would like consumers to under-estimate side effects, but not too much, at the fully revealing equilibrium there is a bias towards excessively low taxes. In the case of risky driving, the tax on x (excess speed) is interpretable as a fine proportional to x times the probability of being caught. At the fully revealing equilibrium, the policy-maker is constrained to impose fines that are too low with respect to the second-best.

To understand the role of $\bar{t}$, compare (3) and (8) and remark that, when $t = \bar{t}$, the consumer's and the policy-maker's objective functions are perfectly aligned. This means that, from the policy-maker's viewpoint, the optimal consumer's belief is exactly the true θ. However, if $\bar{t}$ were always chosen, it would not be informative on side effects since it does not depend on θ. The agent's priors would not be updated and x would always be $U'^{-1}[E\theta + \frac{\eta}{1+\lambda}]$. Still, the most important problem with $\bar{t}$ is not informational: $\bar{t}$ is simply systematically too low because, contrary to the Ramsey tax, it neglects the role of the elasticity of demand.

Some comparative statics on λ is worth noting. As λ decreases, the slope of the second-best tax flattens while the tax $\bar{t}$ shifts towards the value of the externality η. Both the range of variation of the equilibrium tax and its maximal value shrink. The intuition is that the conflict between raising fiscal revenues and

11

providing information damps and the signaling costs are substantially reduced. When $\lambda = 0$, the second-best tax and $\bar{t}$ degenerate into the Pigouvian tax η and the equilibrium cannot be informative.

We mentioned in the introduction the debate on cigarettes taxes. According to the empirical literature, excise taxes are too high with respect to the negative externality induced by tobacco consumption. However, since the Ramsey-Sandmo tax is increasing in individual effects and in the marginal cost of public funds, the observed level of cigarettes taxes can be explained by second-best theory. On the other hand, the incentive compatible fully revealing equilibrium tax is always below the second-best one. Figure 2 illustrates a general argument: the fully revealing tax is higher than the Pigouvian tax η for the upper tail of the distribution. In other words, the signaling role of tax policy preserves the expected gap between the observed tax and η. Moreover, the fully revealing tax schedule is increasing and convex: the higher θ, the higher the variation of the tax induced by a small change in side effects.

In conclusion, the fully revealing tax schedule significantly amplifies, at the upper tail of the distribution, minor differences in side effects. In this sense, and in this sense only, we can say that more severe side effects are signalled through drastic taxation.

6   Conclusion

In a previous version of our model (Barigozzi and Villeneuve, 2003), we considered policies combining informative taxes and information campaigns ("cheap talk" or costless messages, e.g. warning labels on cigarette packs or on the hazards of alcoholic beverage). The paper analyzed the relative impact and credibility of cheap talk with respect to tax policies. However, the expositional and technical costs are considerable, and the results basically show that the costly message (tax) is almost sufficient to convey information. From a theoretical point of view, this confirms Manelli's (1996) finding that, in a signaling game, cheap talk closes but does not substantially extend the set of equilibrium allocations. If we see taxes as a proxy for the various imperfect and costly instruments that a policy-maker can impose to affect cigarette consumption, a practical conclusion is that they are typically more efficient, in terms of the precision of information conveyed, than words. From an empirical point of view, there is evidence that short phrases whose content is too vague to be verifiable ("smoking is harmful to health") are often of very limited efficacy, and that a costly tax is taken more seriously than mere propaganda. Bardsley and Olekalns (1999), working on the impact of health warnings on cigarette packs, showed that over the past 35 years, price (including tobacco taxes), real income, and demographic effects explain most of the variation in tobacco consumption, whereas health warnings on cigarette packs have had a relatively minor impact.

In the text, we focused on health policy issues, that is on the taxation of commodities which induce a negative externality and side effects: smoking, alcoholic beverages, fast driving and antibiotics. Applications of the results in the field of environmental policies and multinational firms taxation are possible.

When a resource-saving technology emerges, the diffusion process could be blocked by an imprecise perception of potential economies on operating costs. The policy-maker may decide to encourage the new technology through a favorable tax policy (large subsidies signal large savings). This decision reduces adoption cost directly incurred by consumers, which in the long run may be favorable to the environment to an extent that overpasses the cost in public funds. When a multinational firm evaluates investment in a host country, the rate of return of the investment may be partially unknown.[15] The government in the host country may impose a tax on the invested capital, the level of taxation signals to the firm some credible information about the profitability of its investment.

References

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

[2] Bardsley, P. and N. Olekalns (1999): "Cigarette and Tobacco Consumption: Have Anti-smoking Policies Made a Difference?," Economic Record, 75 (230), 225–240.

[3] Barigozzi, Francesca and Bertrand Villeneuve (2003): "Influencing the Misinformed Misbehaver: An Analysis of Public Policy towards Uncertainty and External Effects," downloadable at http://www.idei.asso.fr/Commun/Articles/Villeneuve/Misinformed%20Misbehaver.pdf

[4] Benabou, Roland and Jean Tirole (2003): "Intrinsic and Extrinsic Motivation," Review of Economic Studies, 70(3), 489-520.

[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.

---

[15]For a model on regulation of multinational firms through tax on capital investment, see Olsen and Osmundsen (2001).

[7] Chaloupka, Frank J. and Kenneth E. Warner (2000): "The Economics of Smoking," The Handbook of Health Economics, edited by Joseph P. Newhouse and Anthony J. Cuyler, New York: North-Holland, Elsevier Science B.V., 1539-1627.

[8] Edlin, Aaron S. and Pinar Karaca-Mandic (2003): "The Accident Externality from Driving," UC Berkeley WP.

[9] Goulder, Lawrence H. (1995): "Environmental Taxation and the Double Dividend: A Reader's Guide," International Tax and Public Finance, 2, 157–183.

[10] Gruber, Jonathan and Botond Közsegi (2003): "Tax Incidence when Individuals are Time-inconsistent: The Case of Cigarette Excise Taxes," Journal of Public Economics, forthcoming.

[11] Hanson, Robin (2003): "Warning Labels as Cheap-talk: Why Regulators Ban Drugs," Journal of Public Economics, forthcoming.

[12] Laffont, 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] Mailath, George J. (1987): "Incentive Compatibility in Signaling Games with a Continuum of Types," Econometrica, 55, 1349–1365.

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

[16] Olsen, Trond Egil and Petter Osmundsen (2001): "Strategic Tax Competition: Implications of National Ownership," Journal of Public Economics 81(2), 253-277.

[17] Pearce, David W. (1991): "The Role of Carbon Taxes in Adjusting to Global Warming," Economic Journal, 101, 938–948.

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

A   Appendix

In the following, we denote U'^{-1} by X. x_1 and x_2 denote values (consumption) taken by the function.

A.1 Proof of Proposition 1

Implicit derivation of the FOC of (9) yields

$$(17) \qquad t'_{\text{SB}}(\theta) = -\frac{H - (1 + \lambda)X'}{H},$$

where $H = X''((1 + \lambda)t - \eta) + (1 + 2\lambda)X' < 0$ (second-order condition). This implies that $t'_{\text{SB}}(\theta) > -1$, i.e. $\theta + t_{\text{SB}}(\theta)$ increases w.r.t. θ.

The FOC of (9) can be written as

$$(18) \qquad t_{\text{SB}}(\theta) = \frac{\eta}{1 + \lambda} - \frac{\lambda}{1 + \lambda}\frac{X[\theta + t_{\text{SB}}(\theta)]}{X'[\theta + t_{\text{SB}}(\theta)]}.$$

To prove the claim, we reason by contradiction. If $t'_{\text{SB}} < 0$, then the RHS decreases w.r.t. θ, which, given the MRV hypothesis, implies that $\theta + t_{\text{SB}}(\theta)$ decreases w.r.t. θ, a contradiction with the previous paragraph.

A.2 Proof of Proposition 2

Along an indifference curve $(\overline{SW} = \text{constant})$, given that $X' = 1/U''$, we have

$$(19) \quad \left[X'((1 + \lambda)t - \eta + \widehat{\theta} - \theta) + \lambda X \right] dt + X' \left[(1 + \lambda)t - \eta + \widehat{\theta} - \theta \right] d\widehat{\theta} = 0.$$

The derivative with respect to $\widehat{\theta}$ of the slope of the normal vector (i.e. orthogonal to the tangent) is

$$(20) \qquad \frac{\lambda U'' X}{(\eta + \theta - t - (1 - \lambda)\widehat{\theta})^2},$$

which is negative. Single-crossing implies that the two branches of an indifference curve go to infinity or cross (at the limit) the boundary $\widehat{\theta} + t = 0$ (this precludes closed curves).

A.3 Proof of Proposition 3

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.3.1 Differentiable Equilibrium

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 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.

In the absence of ambiguity, t denotes $t(\theta)$ and X_{E} denotes $X_{\mathrm{E}}(\theta) = X(\theta + t(\theta)) = U'^{-1}[\theta + t(\theta)]$.

Differential Equation. The government prefers $t(\theta)$ (and the implied consumption) to $t(\theta + d\theta)$ and to $t(\theta - d\theta)$; taking limits we get

$$(21) \qquad X_{\mathrm{E}}'U' + \lambda t'X_{\mathrm{E}} - (\eta + \theta - \lambda t)X_{\mathrm{E}}' = 0.$$

Given that the consumer's first-order condition is $U' = \theta + t$, we can eliminate U' to get the ordinary differential equation

$$(22) \qquad \lambda t' = -(1 + \lambda)\frac{X_{\mathrm{E}}'}{X_{\mathrm{E}}}\left(t - \frac{\eta}{1 + \lambda}\right),$$

which, with $X_{\mathrm{E}}'(\theta) = (1 + t'(\theta))X'(\theta + t(\theta))$ and after rearrangement, is the ODE of the text.

Monotonicity. The second-order condition is:

$$(23) \qquad 0 \geqslant X_{\mathrm{E}}'^2 U'' + X_{\mathrm{E}}''U' + \lambda t''X_{\mathrm{E}} + 2\lambda t'X_{\mathrm{E}}' - (\eta + \theta - \lambda t)X_{\mathrm{E}}'',$$

while the derivative of the first-order condition is:

$$(24) \qquad 0 = X_{\mathrm{E}}'^2 U'' + X_{\mathrm{E}}''U' + \lambda t''X_{\mathrm{E}} + 2\lambda t'X_{\mathrm{E}}' - X_{\mathrm{E}}' - (\eta + \theta - \lambda t)X_{\mathrm{E}}'',$$

Simplifying (23) with (24) we get:

$$(25) \qquad X_{\mathrm{E}}' \leqslant 0.$$

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. Assuming that the government offers $t(\widehat{\theta})$, thereby inducing $X_{\mathrm{E}}(\widehat{\theta})$, we calculate the derivative of the government's utility with respect to $\widehat{\theta}$:

$$(26) \qquad X_{\mathrm{E}}'(\widehat{\theta})U'[X_{\mathrm{E}}(\widehat{\theta})] + (\lambda t'(\widehat{\theta})X_{\mathrm{E}}(\widehat{\theta}) - (\eta + \theta - \lambda t(\widehat{\theta}))X_{\mathrm{E}}'(\widehat{\theta}))$$
$$= X_{\mathrm{E}}'(\widehat{\theta})((1 + \lambda)t - \eta + \widehat{\theta} - \theta) + \lambda t'(\widehat{\theta})X_{\mathrm{E}}(\widehat{\theta}).$$

Simplifying with (22), we find that (26) has the same sign as $X_{\mathrm{E}}'(\widehat{\theta})(\widehat{\theta} - \theta)$. It follows that (26) is positive for $\widehat{\theta} < \theta$ and negative for $\widehat{\theta} > \theta$. This means that incentive compatibility is satisfied everywhere for equilibrium actions.

16

Position of the Solution. In equation (15), the denominator is null (vertical tangent) if only if $t(\theta) = t_{\mathrm{SB}}(\theta)$. We know from the analysis of (18) that this value $t_{\mathrm{SB}}(\theta)$ is unique for each θ.

Assume that $t(\cdot)$ is located below $\bar{t}$. Consider an off-equilibrium tax t_1 slightly above $\bar{t}$ and denote the associated belief by $\widehat{\theta}$. Clearly, when the government's type is $\widehat{\theta}$, $(\widehat{\theta}, t_1)$ is preferred to $(\widehat{\theta}, t(\widehat{\theta}))$. The candidate solution is not incentive compatible. Assume now that $t(\cdot)$ is located above the second-best schedule. $t(\cdot)$ is necessarily decreasing (while t_{SB} is increasing). As θ increases, $t(\cdot)$ approaches t_{SB} and t' goes to $-\infty$. This implies that consumption decreases, which is impossible.

We conclude that $t(\cdot)$ is located between $\bar{t}$ and the second-best schedule, and that it is increasing.

Uniqueness (Boundary Condition). We show that there are incentive compatible beliefs associated with off-equilibrium actions only if the boundary condition $t(\overline{\theta}) = t_{\mathrm{SB}}(\overline{\theta})$ is satisfied.

We reason by contradiction. Let $t(\cdot)$ be a solution to (15) such that $t(\overline{\theta}) < t_{\mathrm{SB}}(\overline{\theta})$. We already excluded that $t(\theta) > t_{\mathrm{SB}}(\theta)$ and we know that $t(\theta)$ is above $\frac{\eta}{1+\lambda}$ and that $t(\theta)$ is strictly increasing. Thus we have $\max_\theta t(\theta) = t(\overline{\theta}) < t_{\mathrm{SB}}(\overline{\theta})$. We choose an arbitrary (off-equilibrium) $\widehat{t}$ in the interval $(t(\overline{\theta}), t_{\mathrm{SB}}(\overline{\theta}))$ and we denote by $\widehat{\theta}$ the associated belief. We prove that there always exists a type $\widetilde{\theta}$ of government that prefers $(\widehat{\theta}, \widehat{t})$ to $(\widehat{\theta}, t(\widehat{\theta}))$, which contradicts incentive compatibility.

Clearly, $t(\widehat{\theta}) < \widehat{t}$, therefore if $\widehat{t} \leqslant t_{\mathrm{SB}}(\widehat{\theta})$, then $\widehat{t}$ is preferable to $t(\widehat{\theta})$ for type $\widehat{\theta}$. This contradiction implies that necessarily, $\widehat{t} > t_{\mathrm{SB}}(\widehat{\theta})$. Consider now the solution $T(\cdot)$ to (15) passing through $(\widehat{\theta}, \widehat{t})$. (1) The previous analysis (Position of the Solution) proves that this solution is decreasing, therefore T hits the second-best curve between $\widehat{\theta}$ and $\overline{\theta}$, for a belief denoted $\widetilde{\theta}$. (2) Moreover, for type $\widetilde{\theta}$, $(\widehat{\theta}, T(\widehat{\theta})) = (\widehat{\theta}, \widehat{t})$ is preferred to $(\widetilde{\theta}, T(\widetilde{\theta})) = (\widetilde{\theta}, t_{\mathrm{SB}}(\widetilde{\theta}))$. Indeed, above the second-best taxes, the differential equation selects globally worst choices rather that globally best choices (see Global Incentive Compatibility). We can conclude that the policy is not incentive compatible for type $\widetilde{\theta}$. By necessity, we have $t(\overline{\theta}) = t_{\mathrm{SB}}(\overline{\theta})$.

Now we prove that associating belief $\overline{\theta}$ to any tax rate *above* $t_{\mathrm{SB}}(\overline{\theta})$ does not induce deviations. Take any $t > t_{\mathrm{SB}}(\overline{\theta})$. Assume that there is a type of government $\widehat{\theta}$ for which $(\overline{\theta}, t)$ is a better policy than $(\widehat{\theta}, t_{\mathrm{SB}}(\widehat{\theta}))$ (contradiction with incentive compatibility). By continuity of the preferences w.r.t. to beliefs, there exists a type $\widetilde{\theta} \in (\widehat{\theta}, \overline{\theta})$ such that $(\overline{\theta}, t))$ and $(\widetilde{\theta}, t(\widetilde{\theta}))$ are equivalent. Call I the corresponding indifference curve. This curve cannot intersect the set of policies that are better than or equivalent to $(\overline{\theta}, t_{\mathrm{SB}}(\overline{\theta}))$ (for type $\overline{\theta}$) since this would imply double crossing of the frontier of this upper contour, which is excluded from Proposition 2. Now remark that both $(\widetilde{\theta}, t_{\mathrm{SB}}(\widetilde{\theta}))$ and $(\overline{\theta}, t_{\mathrm{SB}}(\overline{\theta}))$ are on the

same side of I, meaning that they both provide more welfare that policies in I. This contradicts incentive compatibility for type $\widetilde{\theta}$: $(\overline{\theta}, t_{\text{SB}}(\overline{\theta}))$ is preferred to $(\widetilde{\theta}, t(\widetilde{\theta}))$.

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 $(\underline{\theta}, t(\underline{\theta}))$.

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

A.3.2 Any Fully Revealing Equilibrium is Differentiable

Take a fully-revealing equilibrium. For all consumer beliefs, the government's preferences are single-peaked with respect to t (a direct consequence of Convexity), and there are a maximum of two tax rates per θ, $t_L(\theta)$ and $t_U(\theta)$ that can implement the equilibrium welfare. More precisely, $t_L(\theta) \leqslant t_{\text{SB}}(\theta) \leqslant t_U(\theta)$). 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 (L) and the upper (U) selection respectively. Notice that $\Theta_L \cup \Theta_U = [\underline{\theta}, \overline{\theta}]$ and $\Theta_L \cap \Theta_U = \emptyset$ (mixed strategies are not considered). 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 $(\theta_n)_{n \geqslant 1}$ in Θ_U converging to θ_0. We prove that for all sequences $(\theta_n)_{n \geqslant 1}$,

$$(27) \qquad \lim_{n \to \infty} \lambda \frac{t_n - t_0}{\theta_n - \theta_0} = -(1 + \lambda) \frac{X_{\text{E}}'(\theta_0)}{X_{\text{E}}(\theta_0)} \left(t_0 - \frac{\eta}{1 + \lambda}\right),$$

where t_n denotes $t_U(\theta_n)$. Indeed, incentive constraints (θ_n should not mimic θ_0, and vice-versa) imply that:

$$(28) \quad U[X_{\text{E}}(\theta_n)] - (\eta + \theta_n - \lambda t_n) X_{\text{E}}(\theta_n) \;\geqslant\; U[X_{\text{E}}(\theta_0)] - (\eta + \theta_n - \lambda t_0) X_{\text{E}}(\theta_0),$$
$$(29) \quad U[X_{\text{E}}(\theta_n)] - (\eta + \theta_0 - \lambda t_n) X_{\text{E}}(\theta_n) \;\leqslant\; U[X_{\text{E}}(\theta_0)] - (\eta + \theta_0 - \lambda t_0) X_{\text{E}}(\theta_0).$$

Therefore, taking a first-order approximation around t_0 and θ_0 yields

$$(30)$$
$$0 \;\geqslant\; X_{\text{E}}'(\theta_0)((1 + \lambda)t - \eta)(\theta_n - \theta_0) + \lambda X_{\text{E}}(\theta_0)(t_n - t_0) + o(\theta_n - \theta_0) + o(t_n - t_0),$$
$$(31)$$
$$0 \;\leqslant\; X_{\text{E}}'(\theta_0)((1 + \lambda)t - \eta)(\theta_n - \theta_0) + \lambda X_{\text{E}}(\theta_0)(t_n - t_0) + o(\theta_n - \theta_0) + o(t_n - t_0).$$

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, and that it is driven by (22).

18

A solution of the differential equation (22) situated above the second-best taxes is not incentive compatible at any point, because the second-order condition is never satisfied (see paragraph on local incentive compatibility). 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, being the complementary set (in an interval) of a set Θ_U which is not dense anywhere, is dense in $[\underline{\theta}, \overline{\theta}]$. Consequently, t_L satisfies the differential equation (22) 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. If not, there would be an accumulation point θ in Θ_U around which (27) would be satisfied, meaning that policies in the neighborhood of θ follow the differential equation (22). But in this region of the plan, we proved that policies along (22) are never incentive compatible (Global Incentive Compatibility), which proves that Θ_U must not contain accumulation points. In other words, Θ_U contains a finite number of points.
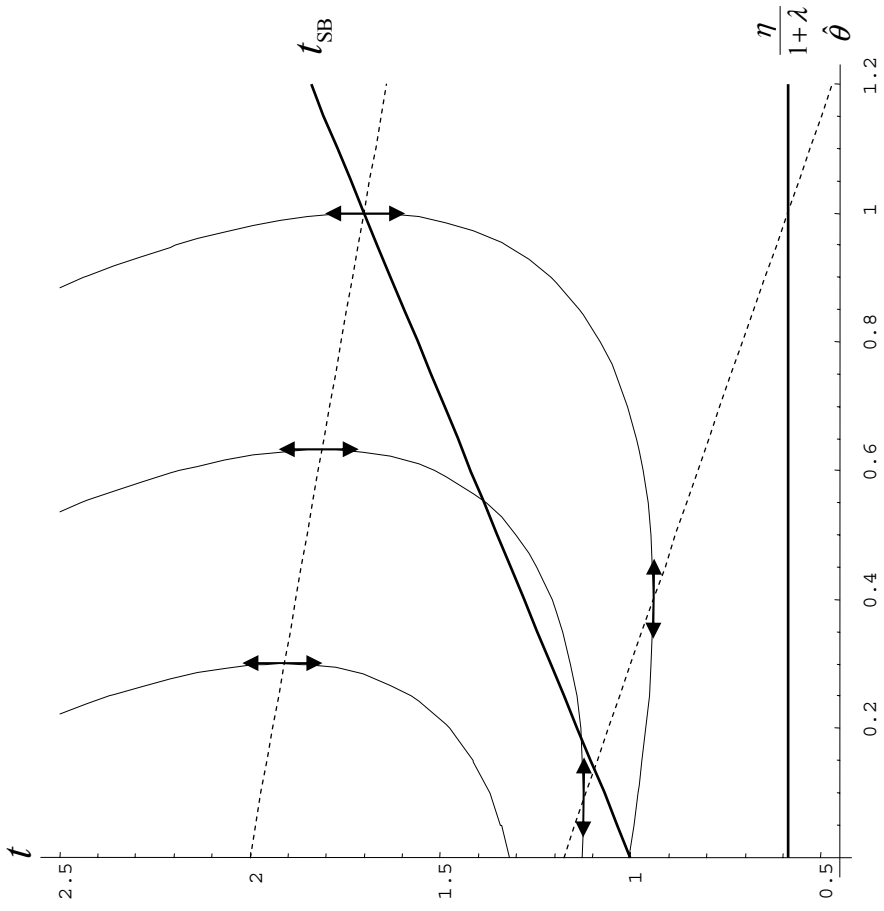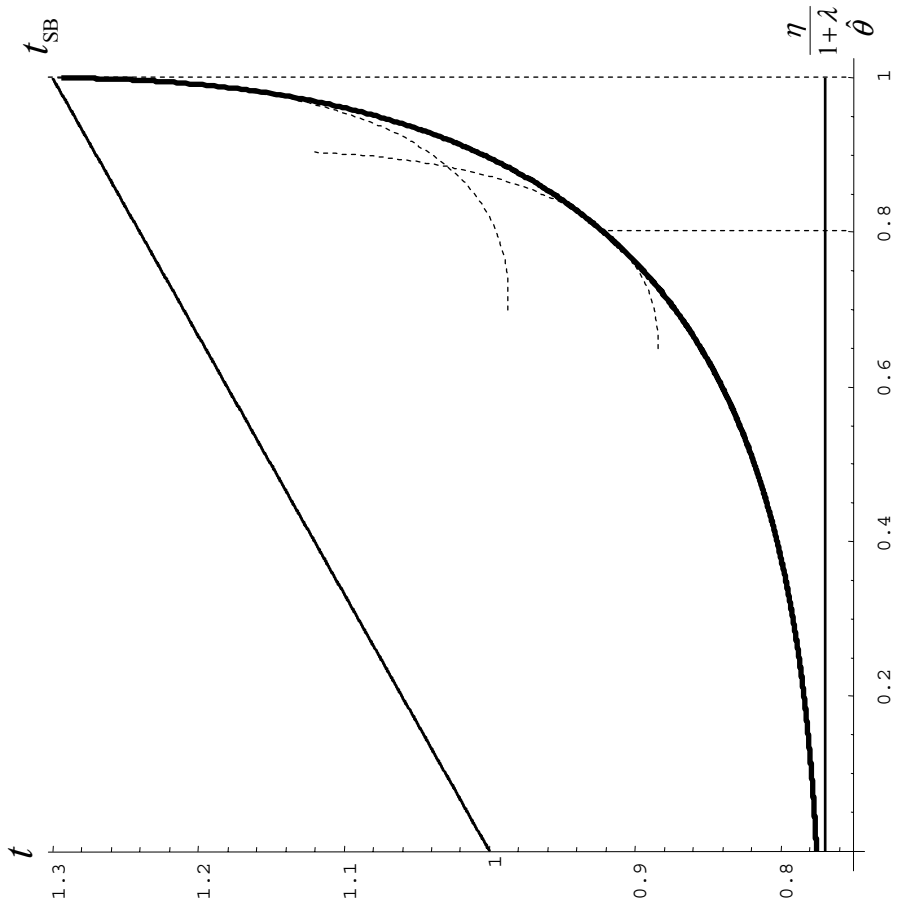
Figure 1

Figure 2