CSEF
Centre for Studies in
Economics and Finance

WORKING PAPER NO. 536

Education, Taxation and the Perceived Effects of Sin Good Consumption

Giovanni Immordino, Anna Maria C. Menichini, Maria Grazia Romano

July 2019

University of Naples Federico II | University of Salerno | Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS – UNIVERSITY OF NAPLES
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it

WORKING PAPER NO. 536

Education, Taxation and the Perceived Effects of Sin Good Consumption

Giovanni Immordino[*], Anna Maria C. Menichini[**], Maria Grazia Romano[**]

Abstract
In a setting in which an agent has a behavioral bias that causes an underestimation or an overestimation of the health consequences of sin goods consumption, the paper studies how a social planner can affect the demand of such goods through education initiatives and/or taxation. When only optimistic consumers are present, depending on the elasticity of demand of the sin good with respect to taxation and the relative efficiency of educational measures, the two instruments can be used as substitutes or complements. When both optimistic and pessimistic consumers coexist, the correcting effect that taxation has on optimistic consumers has unintended distorting effects on pessimistic ones. In this framework, educational measures, by aligning both consumers' perceptions closer to the true probability of health damages, are more effective than taxation.

JEL Classification: D03, H21, L51.

Keywords: Overoptimism, Taxation, Educational initiatives, Sin goods.

[*] Università di Napoli Federico II and CSEF. Address: Dipartimento di Scienze Economiche e Statistiche, Università di Napoli Federico II, Via Cintia, Complesso Monte Sant'Angelo, 80126 Napoli, Italy. E-mail: giovanni.immordino@unina.it.

[**] Università di Salerno and CSEF. Contact address: Dipartimento di Scienze Economiche e Statistiche, Università d Salerno, Via Giovanni Paolo II, 84084 Fisciano (SA), Italy. E-mails: amenichini@unisa.it and maromano@unisa.it.

Table of contents

Abstract

In a setting in which an agent has a behavioral bias that causes an underestimation or an overestimation of the health consequences of sin goods consumption, the paper studies how a social planner can affect the demand of such goods through education initiatives and/or taxation. When only optimistic consumers are present, depending on the elasticity of demand of the sin good with respect to taxation and the relative efficiency of educational measures, the two instruments can be used as substitutes or complements. When both optimistic and pessimistic consumers coexist, the correcting effect that taxation has on optimistic consumers has unintended distorting effects on pessimistic ones. In this framework, educational measures, by aligning both consumers' perceptions closer to the true probability of health damages, are more effective than taxation.

Keywords: Overoptimism, Taxation, Educational initiatives, Sin goods.

JEL classification: D03, H21, L51.

1 Introduction

It is widely believed that overoptimism is a common trait of human beings. An individual with such characteristic is typically inclined to think that favorable events are more likely or more positive than they actually are. In investment decisions, for example, an optimistic agent tends to overestimates a project's future returns. Although much literature has focused on the effects of overoptimistim on economic and financial decisions (Roll, 1986; Malmendier and Tate, 2005, 2008; Bénabou, 2012), less attention has been paid to the impact that the minority of pessimistic agents, i.e., of those who think that favorable events are less likely than they actually are, has on such decisions. In this paper we argue that, however small, this fraction of agents may reduce the effectiveness of the policies aimed at correcting the negative effects of an overoptimistic bias.

We construct a model in which a risk averse agent has a behavioral bias regarding the future effects related to the consumption of sin goods, i.e., goods which are enjoyable to consume but have future health consequences. This bias causes a misperception of the health damages, namely, an underestimation in optimistic consumers or an overestimation in pessimistic ones. Within this setting, we study the way in which a social planner can affect the demand of such goods through taxation and education initiatives. The latter are referred to information/awareness campaigns, such as warning labels on tobacco, alcohol or fatty food, and have the effect of increasing the consumers' awareness and align their perception of the health damages related to the consumption of such goods to the actual one.[1]

We first analyse the benchmark case in which there are only optimistic consumers and determine the optimal mix of education initiatives and taxation needed to correct the consumer's behavioral bias. We then introduce also pessimistic consumers and show that the correcting effect that taxation has on optimistic consumers has a side effect of depressing the consumption of the pessimistic ones. Thus, introducing pessimistic consumers in a setting in which consumers are mainly optimistic highlights one possible flaw of taxation, that, while correcting the excess consumption of optimistic consumers, worsens the underconsumption problem of pessimistic ones. Education, instead, by aligning the consumers perception of the health damages to the

[1] An description of education initiatives and of the literature testing their effectiveness is provided in the next section.

actual one, not only contrasts the consumer's behavioral bias, but also mitigates the negative effects of taxation.

In modeling taxation and education measures, we assume that they are both costly. In particular, the cost of taxation is related to the inefficiency of the fiscal system translating one euro of tax revenues in less than one euro transferred to consumers. Education initiatives, instead, generate a disutility on consumers for managing and processing the information received.

Focusing on a simple quasi-linear economy in which there is a composite good and a sin good, we analyze the two benchmark cases in which one of two instruments, taxation and education, is used in isolation. We find that if education is the only instrument, its level is always strictly positive. When only taxation is available, instead, its level is strictly positive so long the inefficiency of the tax system is not too high. This is because, unlike the marginal cost of education, the marginal cost of taxation is always strictly positive and equal to the deadweight loss implied by the inefficiency of the tax system. Clearly, when this is sufficiently high, it may be optimal for a regulator not to rely on taxation to correct consumers' behavior.

When both instruments are available, we find the conditions under which taxation and education are both relied upon. Focusing on the case in which the fiscal system is not too inefficient so that both instruments are used at the optimum, we show that taxation and education can be substitute or complement, depending on the elasticity of the demand of the sin good with respect to taxation and the relative effectiveness of education relative to taxation in reducing the demand of the sin good. In particular, when the demand of the sin good is inelastic with respect to taxation, the marginal cost of taxation is positive and increasing in the degree of inefficiency of the fiscal system. An increased inefficiency of the fiscal system makes room for education initiatives. When instead the demand of the sin good is elastic with respect to taxation, the marginal cost of taxation is negative and decreasing in the degree of inefficiency of the fiscal system. An increased inefficiency of the fiscal system should then call for an increased reliance on taxation relative to education. However, whether this is so depends now on the relative effectiveness of the two instruments in reducing the consumption of the sin good. In particular, when taxation is more (less) effective than education, an increased inefficiency of the fiscal system calls for an increased (reduced) reliance on taxation relative to education.

We then introduce the possibility that some consumers may be pessimistic, i.e., the perception of the health damages caused by sin good consumption exceeds the actual likelihood. In this framework, we show our main result, i.e., that the presence of pessimistic agents reduces the effectiveness of taxation in mitigating overconsumption and induces the social planner to boost up educational measures. This is because the correcting effect of taxation on the overconsumption of optimistic consumers has unintended depressing effects on the choices of pessimistic ones. This asymmetry is not at work when education is used because of its effectiveness in aligning both optimistic and pessimistic consumers' perception to the true probability of health damages.

The paper is related to the literature on time-inconsistency and hyperbolic discounting (Ainslie, 1992), which studies the welfare effects of sin good regulation (O'Donoghue and Rabin, 2003, 2006; Gruber and Koszegi, 2001, 2004; Gruber and Mullainathan, 2005; Immordino, Menichini and Romano, 2019). With respect to these contributions we study a different motivation for government intervention, namely, to correct the distortions induced by the misperception of the health damages due to sin good consumption.

The paper is also related to the literature on overconfidence and overoptimism, and the effects that managers with behavioral biases have on corporate decision making. For instance, Goel and Thakor (2008) show that overconfident managers are more likely to be promoted to CEOs than perfectly rational ones because they perceive less risk and so take more chances. Moreover, a moderate degree of overconfidence is beneficial, mitigating the problem of underinvestment that plagues strictly rational managers. The benefits of overconfidence and moderate overoptimism are also highlighted by Gervais, Heaton, and Odean (2003), who show that both these traits help offset the excessive prudence induced by risk aversion, inducing managers to take investment decisions less hesitantly. Unlike the above cited literature, in the present paper agents overoptimism is always detrimental to agents as they underestimate the probability of health damages and are induced to consume too much unhealthy goods.

The paper is organized as follows. In the next section, we provide a description of education initiatives. In Section 3, we set up the model focusing on optimistic consumers, and develop the two benchmark cases in which only education initiatives or only taxation can be used in

Subsections 3.1 and 3.2, respectively. In Section 4, we consider the optimal mix of the two instruments when they can be jointly used. In Section 5, we analyze how this mix is affected by the presence of pessimistic consumers. All the proofs are in the Appendix.

2 Evidence on education initiatives

Everyday life provides us with several examples of initiatives aimed at forbidding, limiting or deterring the consumption of sin goods. These initiatives can be categorized into two groups: educational initiatives, aimed at increasing the awareness of the health effects of the consumption of sin goods, such as information/awareness campaigns, and policy initiatives, aimed at directly limiting the availability and provision of sin goods, such as consumption restrictions.Educational initiatives are widespreadly used. In many countries, for example, health warning labels appear on fatty food, tobacco or alcohol products pointing to the health risks associated with their consumption, and are often mandatory. In some cases they can take the form of recommendations, like in responsible drinking campaigns to prevent alcoholism or drunk driving. Recommendations to moderate/responsible consumption also close advertisements campaigns of alcohol products and gambling (e.g., gamble responsibly). On the packaging of cigarettes and other tobacco products a variety of textual and pictorial warnings appear, covering, within a black frame, a large part of the surface of the pack and concerning the health effects of tobacco products consumption. For fatty food, some countries (e.g., UK) have developed a system of front of pack nutritional labels that associates colors with information on fat, salt, sugar, and calories contained in food products, to help people making healthier choices.

Education initiatives seem to be effective in increasing the consumers' knowledge and attitude about the health consequences of the consumption of sin goods. For tobacco and alcohol products, for example, comprehensive review studies have provided evidence showing the effectiveness of strategies and interventions aimed at preventing smoking uptakes (Thomas et al., 2013) or alcohol related problems (Babor et al., 2003). Although much more limited, some evidence is also available for food products. For example, a study by Cioffi et al. (2015) has shown that the introduction of food labels on a sample of pre-packaged food items results in a reduction of the average calories purchased from the labelled foods.

3 The Model

We study a setting in which a representative optimistic risk averse agent has to choose the optimal level of consumption of a sin good, i.e., a good that is enjoyable to consume but may create negative health consequences, such as alcohol, cigarettes, potato chips. The agent's utility is quasi-linear with respect to the sin good (x) and a composite good which acts as a numeraire (z). Specifically,

$$u(x,z) = \begin{cases} v(x) - c(x) + z & \text{with probability } q \\ v(x) + z & \text{with probability } 1-q \end{cases}$$

The function $v(\cdot)$ represents the the benefit from sin good consumption and satisfies Inada conditions. The function $c(\cdot)$ represents the uncertain negative health consequences from sin good consumption and is such that $c_x > 0$, $c_{xx} > 0$ and small, and $c_x(x) = 0$ when $x = 0$.[2] The parameter $q \in (0, 1]$ is the (true) probability that the sin good causes health damages. To preserve the empirically desirable feature of decreasing absolute risk aversion for any q, we also assume that $v_{xxx} \geq 0$ and $c_{xxx} \leq 0$.[3]

Agents are optimistic in that they underestimate the probability of health damages. We denote by $q^o \leq q$ the agent's perceived probability that the sin good causes health damages. Thus, the agent maximizes the *optimistic* expected utility function

$$U^o \equiv v(x) - q^o c(x) + z, \tag{1}$$

subject to the budget constraint $I = p_x x + p_z z$, where I is the exogenous income earned by the consumer, p_x and p_z are the prices of the sin good and of the numeraire, respectively. We assume that there is no borrowing or lending, that markets are competitive and that the marginal cost of producing both goods is equal to one, so that the price of each good is also one.[4]

Because of overoptimism, if $q^o < q$ the consumer does not maximize his own expected welfare, measured by the *actual* expected utility function

$$U = v(x) - qc(x) + z. \tag{2}$$

[2] Those assumptions guarantee that the consumer's problem is well-behaved and are made to simplify the exposition. In particular, the assumptions that $\lim_{x\to 0} v_x(x) = \infty$, $\lim_{x\to\infty} v_x(x) = 0$, and $c_x(x) = 0$ when $x = 0$ ensure that the sin good demand is strictly positive for any price $p_x < \infty$. Moreover, $c_{xx} > 0$ and small ensures concavity of the social planner optimization problem.

[3] For the empirical relevance of this assumption, see Guiso and Paiella (2008).

[4] We assume that I is large relative to the sin good consumption, so as to avoid corner solutions for x.

In the following, we will call optimistic utility the expected utility function corresponding the choice of the agent (U^o), and actual utility the expected utility function that correctly reflects the expected welfare of the agent (U).

The first-best consumption, which we denote by (x^{FB}, z^{FB}), maximizes the actual utility U, subject to the budget constraint $I = x + z$, and it is such that x^{FB} satisfies the first order condition $v_x(x^{FB}) - qc_x(x^{FB}) = 1$ and $z^{FB} = I - x^{FB}$.

In the absence of policy measures aimed at affecting the consumers' behavior, the agent's consumption of the sin good, x^o, satisfies the first order condition $v_x(x^o) - q^o c_x(x^o) = 1$. Since v_x is decreasing in x. c_x is increasing in x and $v_x(x) - qc_x(x)$ is lower than $v_x(x) - q^o c_x(x)$ for any x, x^o is larger than x^{FB}. Moreover, $z^o = I - x^o$ is lower than z^{FB}. Thus, because of overoptimism, the agent consumes too much of the sin good and too little of the numeraire.

We next introduce the two benchmark cases in which one of two alternative measures — education or taxation — is introduced to correct the overconsumption of sin goods implied by overoptimism.

3.1 Education initiatives

In this section we consider education initiatives, i.e., all the regulatory measures aimed at creating the conditions for consumers to voluntarily limit consumption, like communication, awareness and education campaigns. To model them, we assume that the government adopts an informational campaign aimed at increasing the awareness of the expected negative health consequences of sin good consumption and thereby affecting the same. Such campaigns may also have negative effects, as they may generate a disutility on consumers, for instance in absorbing and processing information. The level of information provision is captured by the parameter $\gamma \in [0, \overline{\gamma}]$, where $\gamma = 0$ means absence of education measures and $\gamma = \overline{\gamma}$ implies full awareness of the health damages of sin good consumption, i.e., $q^o = q$. Formally:

Assumption 1 *The benefit of information provision is given by a continuous function $q^o(\gamma)$ defined on $[0, \bar{\gamma}]$, with $q^o(\gamma)$ strictly increasing and concave, and equal to q^o when $\gamma = 0$ and equal to q when $\gamma = \bar{\gamma}$. The disutility of regulation is given by a continuous function $b(\gamma)$ defined on $(0, \bar{\gamma}]$, with $b(\gamma)$ strictly increasing and convex, and such that* $\lim_{\gamma \to 0} b(\gamma) = 0$.

When educational initiatives are introduced, the agent's actual utility becomes

$$U(\gamma) = v(x) - q\,c(x) - b(\gamma) + z, \tag{3}$$

and the agent's optimistic utility becomes

$$U^o(\gamma) = v(x) - q^o(\gamma)\,c(x) - b(\gamma) + z. \tag{4}$$

An increase in γ has a negative effect on the level of both actual and optimistic utility ($U_\gamma(\gamma) = -b_\gamma < 0$ and $U^o_\gamma(\gamma) = -b_\gamma - q^o_\gamma\,c(x) < 0$), and on the marginal optimistic utility ($U^o_{x\gamma}(\gamma) = -q^o_\gamma\,c_x < 0$).

The optimal agent's consumption bundle, which we denote by $(x(\gamma), z(\gamma))$, maximizes (4) subject to $z = I - x$. Then, the consumption rule of the sin good, $x(\gamma)$, satisfies the first order condition

$$v_x(x) - q^o(\gamma)\,c_x(x) = 1. \tag{5}$$

Clearly, $z(\gamma) = I - x(\gamma)$. Since regulation improves the consumer's awareness $q^o(\gamma)$, the consumption of the sin good decreases and that of the numeraire increases.[5] More generally, the agent's sin good consumption is decreasing in γ. Indeed, the first derivative of $x(\gamma)$ with respect to γ is $x_\gamma(\gamma) \equiv \frac{q^o_\gamma(\gamma)c_x(x(\gamma))}{v_{xx}(x(\gamma)) - q^o(\gamma)c_{xx}(x(\gamma))}$, that is negative.

The social planner chooses the level of information provision γ that maximizes the actual utility (3) subject to the budget constraint $I - x = z$ and the consumption rule $x(\gamma)$ defined by condition (5).

By substituting the budget constraint and the consumption rule in the actual utility function, the latter can be written as the difference between the benefit and the cost of regulation, that is

$$\Omega(\gamma) = \underbrace{[v(x(\gamma)) - q\,c(x(\gamma)) + I - x(\gamma)]}_{BI(\gamma)} - \underbrace{b(\gamma)}_{CI(\gamma)}. \tag{6}$$

The first term in square brackets, $BI(\gamma)$, is the expected benefit from information provision, i.e., the expected utility that would be obtained by inducing a level of consumption $x(\gamma) < x^o$ if there was no disutility associated with information provision. The second term, $CI(\gamma)$, represents the

[5]Indeed, because information provision increases the weight of the marginal health cost from sin good consumption ($q^o(\gamma)$) and because the utility function is concave, it follows that $x(\gamma)$ is lower than x^o and $z(\gamma)$ is higher than z^o for any γ.

cost induced by the inefficiency of information provision, namely, the reduction in the immediate benefit of current consumption due to the disutility it generates.

The social planner's problem is to choose $\hat{\gamma}$ that maximizes the distance between the benefits and the costs of information provision. To determine the optimal level of information provision, the social planner compares the inefficiency of the regulatory instrument, $b_\gamma(\gamma)$, with its effectiveness in reducing the utility loss associated to overconsumption of the sin good, $(q - q^o(\gamma))\,|c_\gamma|$. The effectiveness of the information provision is measured by the impact of γ on the health costs, $|c_\gamma| \equiv c_x\,|x_\gamma|$, weighted by the residual overoptimism upon regulation, i.e., the difference between the actual and the optimistic likelihood of health damages, $q - q^o(\gamma)$. A higher awareness reduces the consumption of the sin good, and thus its cost. The impact of such reduction is larger, the larger is the distance between the actual and the subjective/optimistic likelihood of health damages.

Proposition 1 *Assume that $q^o < q$. The optimal level of information provision is $\hat{\gamma} \in (0, \bar{\gamma})$ such that*

$$(q - q^o(\hat{\gamma}))\,|c_\gamma(x(\hat{\gamma}))| = b_\gamma(\hat{\gamma}).$$

Proposition 1 states that if $q^o < q$, the agent's consumption is too high and the optimal information provision, $\hat{\gamma}$, is always strictly positive. Moreover, the optimal regulation is such that the disutility generated by information provision is offset by its benefits. An increase in information provision generates a trade-off between the higher inefficiency related to the costs of absorbing and processing information, and the lower health damages due to the higher awareness that reduces the consumption of the sin good. At the optimum, $\hat{\gamma}$ is such that the inefficiency of the information provision equals its effectiveness.

3.2 Taxes

In this section we study the second of the two benchmark cases in which a linear tax τ aimed at reducing the over-consumption of the sin good is used. Such case is isomorphic to the analysis developed by Immordino, Menichini, and Romano (2019) for time-inconsistent individuals. As in that paper, we assume that the tax proceeds τx are redistributed in a lump sum way to consumers and that one euro tax revenues translates in less than one euro transfer for consumers due to the inefficiency of the fiscal system.[6] Formally:

Assumption 2 *The per-capita transfer l from tax proceeds is given by:*

$$l = (1-\lambda)\tau x, \tag{7}$$

where $\lambda \in (0, 1/2)$ is the direct inefficiency of the tax system, reflecting the loss in the economy from collecting one euro tax revenues.[7]

In the case of a linear tax τ and a lump sum transfer l, the agent's budget constraint becomes $I + l = (1+\tau)\,x + z$, and the consumption rule of the sin good, $x(\tau)$, satisfies the first order condition

$$v_x(x) - q^o c_x(x) = 1 + \tau. \tag{8}$$

Clearly, $z(\tau) = I + l - (1+\tau)\,x(\tau)$ and, from the concavity of the utility function, $x(\tau)$ is lower than x^o and decreasing in τ. Indeed, the first derivative of $x(\tau)$ with respect to τ is $x_\tau(\tau) \equiv \frac{1}{v_{xx}(x(\tau)) - q^o c_{xx}(x(\tau))}$, which is negative for all τ.

The social planner chooses the level of taxation $\hat{\tau}$ that maximizes the actual utility (2) subject to the budget constraint $z = I + l - (1+\tau)\,x$, the lump-sum transfer constraint (7) and the consumption rule $x\,(\tau)$ defined by condition (8). By substituting the budget constraint, the lump-sum transfer constraint and the consumption rule in the actual utility function, the objective function reads as

$$\Omega\,(\tau) = \underbrace{[v(x\,(\tau)) - qc(x\,(\tau)) + I - x\,(\tau)]}_{BT(\tau)} - \underbrace{\lambda \tau x\,(\tau)}_{CT(\tau)}. \tag{9}$$

The term $BT\,(\tau)$ represents the benefit of taxation and is given by the utility that would be obtained by inducing a level of consumption $x\,(\tau) \leq x^o$ if there was no inefficiency associated with taxation ($\lambda = 0$). The second term, $CT\,(\tau)$, represents the resources lost due to the inefficiency

[6]These can be both administrative and compliance costs. The first, which are incurred by the tax authority to collect taxes and enforce fiscal laws, have been estimated to be 0.5% of net revenue collection for US, with a median of about 1% for OECD countries (OECD, 2011). For compliance costs, a study by Pricewaterhouse Coopers (2015) for 189 countries across the world reports an average of 99 hours spent by consumers to comply with sales tax and VAT.

[7]The assumption that $\lambda \in (0, 1/2)$, along with the assumption of decreasing absolute risk aversion, guarantees the concavity of the optimization problem. However, our qualitative results also hold in a more general setting where these assumptions are relaxed.

of taxation that reduce the consumption of the numeraire. The social planner's problem is to choose $\hat{\tau}$ that maximizes the distance between the benefits and costs of taxation. The optimal level of taxation will depend on the trade-off between the inefficiency of the fiscal system and the capability of taxation in reducing the expected utility loss due to the overoptimism of the agent.

When there is no efficiency loss associated with taxation ($\lambda = 0$), the social planner's problem (9) simplifies to maximizing $BT(\tau)$ and the optimal tax chosen by the social planner $\hat{\tau}$ coincides with the level $\tau^{FB} = (q - q^o)c_x(x^{FB})$ that induces the agent to consume the first-best level of the sin good. This can be better understood by noticing that in the social planner's problem (9) the benefit from taxation $BT(\tau)$ is maximum when the agent consumes the first-best level of the sin good, x^*.[8] However, when $\lambda > 0$, the cost component $CT(\tau)$ of the social planner's problem (9) is positive. It turns out that, when $\lambda > 0$, whether the optimal tax rate exceeds or falls short of the first-best depends on the elasticity of the demand of the sin good with respect to taxation, i.e., $\eta_{x,\tau} = \frac{x_\tau(\tau)\tau}{x(\tau)}$.[9] The intuition is the following. When the inefficiency of taxation is strictly positive, an increase in the sin good tax has two opposite effects on the taxation cost $CT(\tau) = \lambda\tau x(\tau)$: a negative direct effect due to the higher price paid on each unit of sin good purchased, and a positive indirect effect due to the distortionary impact of taxation on quantities. When the demand is highly elastic ($\eta_{x,\tau} < -1$), the positive effect prevails on the negative one and the optimal taxation exceeds the first best ($\hat{\tau} > \tau^{FB}$). Conversely, when the elasticity is low ($\eta_{x,\tau} > -1$), the negative effect prevails on the positive one and taxation has a very negative impact on the consumption of the numeraire. To mitigate such impact, taxation has to be set lower than its first-best level ($\hat{\tau} < \tau^{FB}$). Finally, if $\eta_{x,\tau} = -1$ the optimal tax $\hat{\tau}$ is τ^{FB}, regardless of λ.

Proposition 2 shows that if Assumption (2) holds, then the optimal tax is strictly positive and can be both higher or lower than the first-best taxation, τ^{FB}, depending on the elasticity of demand of the sin good with respect to taxation, $\eta_{x,\tau}$.

Proposition 2 *Assume that $q^o < q$. If $\lambda < \lambda^M$, with $\lambda^M = \min\left\{\frac{(q-q^o)\,|c_\tau(x^o)|}{x^o}, \frac{1}{2}\right\}$, the optimal level of taxation is $\hat{\tau} > 0$ such that $(q - q^o)\,|c_\tau(x(\hat{\tau}))| + \hat{\tau}x_\tau(\hat{\tau}) = \lambda\left(1 + \eta_{x,\tau}\right)x(\hat{\tau})$.*

[8] Indeed, $BT'(\tau) = [v_x(x(\tau)) - qc_x(x(\tau)) - 1]\, x_\tau(\tau) = 0$ when $x(\tau) = x^{FB}$.

[9] For a formal proof of this result see Immordino, Menichini, and Romano (2019).

The left hand side represents the net marginal benefit of taxation. It is equal to the effectiveness of sin taxes, i.e., the reduction in the health costs due to the lower sin good consumption induced by taxation, weighted by the overoptimism $((q - q^o)\,|c_\tau|)$, net of the contraction of the sin tax revenues produced by the reduction in the sin good demand (τx_τ). The right hand side represents the marginal cost of taxation and is related to the inefficiency of the fiscal system, λ. It may be positive or negative depending on how the tax proceeds vary with taxation $\left(1 + \eta_{x,\tau}\right)$. If the elasticity of the demand of the sin good with respect to taxation is low $(\eta_{x,\tau} > -1)$, the demand of the sin good varies less than proportionally with taxation and the tax proceeds increase with τ. This in turn implies that more resources are available for distribution and a higher loss due to the inefficiency of taxation arises. If instead the elasticity of the demand of the sin good with respect to taxation is high $(\eta_{x,\tau} < -1)$, the demand of the sin good varies more than proportionally with taxation and the tax proceeds decrease with τ. This implies that the resources to redistribute are lower and so is the loss due to the inefficiency of taxation.

4 Education initiatives and taxes

In this section we study the case where both educational measures and taxation are available to the social planner to correct consumers' overoptimism. Again, the effect of educational initiatives is to provide information useful to reduce the distance between q^o and q, while the effect of taxation is to limit the consumption of the sin good by increasing its price, that becomes $p_x = 1 + \tau$.

In the case of educational initiatives γ, linear tax τ, and lump sum transfer l, the optimal agent's consumption level maximizes the agent's optimistic utility $v(x) - b(\gamma) - q^o(\gamma)\,c(x) + z$ subject to the budget constraint $z = I + l - (1 + \tau)\,x$. Then, the optimal sin good consumption satisfies the first order condition

$$v_x(x) - q^o(\gamma)\,c_x(x) = (1 + \tau)\,. \tag{10}$$

Clearly, $z = I + l - (1 + \tau)\,x$.

Let $x(\gamma, \tau)$ be the agent's consumption rule of the sin good defined from condition (10). Since taxation increases the price of the sin good, the agent's consumption rule $x(\gamma, \tau)$ is lower than $x(\gamma)$, for any policy pair (γ, τ).

The social planner chooses the level of education, γ^*, and the level of taxation, τ^*, that maximize the actual utility function (3), subject to the budget constraint $z = I + l - (1+\tau)\,x$, the lump-sum transfer constraint $l = (1-\lambda)\tau x$, and the consumption rule defined by condition (10).

By substituting the budget constraint, the lump-sum transfer constraint, and the consumption rule in the objective function, the latter reads as

$$\Omega(\gamma,\tau) = \underbrace{\left[v\left(x\left(\gamma,\tau\right)\right) - qc(x\left(\gamma,\tau\right)) + I - x\left(\gamma,\tau\right)\right]}_{B(\gamma,\tau)} - \underbrace{\lambda\tau x\left(\gamma,\tau\right)}_{CT(\gamma,\tau)} - \underbrace{b\left(\gamma\right)}_{CI(\gamma)}. \tag{11}$$

For any policy pair (γ, τ), the term in square brackets represents the benefit of consumption, $B(\gamma,\tau)$. The second term, $CT(\gamma,\tau)$, represents the cost of taxation in terms of reduced consumption of numeraire due to the inefficiency of taxation, while the third term, $CI(\gamma,\tau)$, represents the cost of educational initiatives.

Proposition 3 states that information provision is always Pareto improving and, for λ sufficiently low, the regulator prefers to use both instruments.

Proposition 3 *Assume that $q^o < q$. There exists a threshold*

$$\bar{\lambda} \equiv \frac{b_\gamma}{x\left(\hat{\gamma}\right)}\frac{x_\tau}{x_\gamma} \tag{12}$$

such that if $\lambda < \min\left\{\lambda^M, \bar{\lambda}\right\}$, the optimal policy involves both education and taxation, i.e., $\gamma^* > 0$ and $\tau^* > 0$. Moreover, $\gamma^* < \hat{\gamma}$ and $\tau^* < \hat{\tau}$.

As for the case in which only taxation can be used, the use of both instruments rests on taxation being not too inefficient. If this is the case, the reliance on taxation allows the regulator to reduce the reliance on education and save on education costs. This in turn implies that optimally a mix of the two instruments is used.

When $\min\left\{\lambda^M, \bar{\lambda}\right\} = \bar{\lambda}$, the threshold $\bar{\lambda}$ has an economic interpretation. Indeed (12) can be rewritten as $\frac{\bar{\lambda}x\left(\hat{\gamma}\right)}{x_\tau} = \frac{b_\gamma}{x_\gamma}$, where the left hand side can be interpreted as the ratio between the marginal cost and the marginal effectiveness of (introducing) taxation, while the right hand side as the ratio between the marginal cost and the marginal effectiveness of education.[10]

[10]Notice that the threshold value of λ is computed for the quantity corresponding to the optimal level of education $\hat{\gamma}$ when no taxation is used $(x\left(\hat{\gamma}\right))$.

However, within the range of values in which both instruments are used, Proposition 4 shows that the way in which they are relied upon, in conjunction or in alternative to each other, depends on the elasticity of the demand of the sin good with respect to taxation, $\eta_{x,\tau}$, and on the relative effectiveness of the two instruments in reducing overconsumption., namely:

Proposition 4 *Assume that* $\lambda < \min\left\{\lambda^M, \bar{\lambda}\right\}$. *Then:*

- *if* $\eta_{x,\tau} > -1$, *then* $\partial\tau^*/\partial\lambda \leq 0$ *and* $\partial\gamma^*/\partial\lambda \geq 0$*;*
- *if* $\eta_{x,\tau} < -1$, *then there exists a lower bound* $\underline{\alpha}$ *and an upper bound* $\overline{\alpha}$ *such that:*

   1. $\partial\tau^*/\partial\lambda \leq 0$ *and* $\partial\gamma^*/\partial\lambda > 0$ *if* $\frac{x_\gamma}{x_\tau} \geq \overline{\alpha}$,
   2. $\partial\tau^*/\partial\lambda > 0$ *and* $\partial\gamma^*/\partial\lambda > 0$ *if* $\underline{\alpha} < \frac{x_\gamma}{x_\tau} < \overline{\alpha}$,
   3. $\partial\tau^*/\partial\lambda > 0$ *and* $\partial\gamma^*/\partial\lambda \leq 0$ *if* $\frac{x_\gamma}{x_\tau} \leq \underline{\alpha}$.

The intuition behind the above results is the following. When the demand of the sin good is inelastic ($\eta_{x,\tau} > -1$), the marginal cost of taxation (i.e., the deadweight loss of an increase in taxation) is positive and increasing in λ. In particular, because the demand varies less than proportionally with taxation, an increase in taxation increases both the proceeds and the deadweight loss of taxation. Since the deadweight loss is linear in λ, an increase in λ increases also the marginal cost of taxation, lowering the reliance on taxation and increasing the reliance on education. Thus, the two instruments are substitutes.

When the demand of the sin good with respect to taxation is elastic ($\eta_{x,\tau} < -1$), the marginal cost of taxation is negative and decreasing in λ. One could then expect an increased reliance on taxation and a decreased reliance on education. However, whether this is so depends on the kick in of a second factor, namely, the effectiveness of education relative to taxation in reducing the consumption of the sin good, measured by $\frac{x_\gamma}{x_\tau}$.

When the effectiveness of education relative to taxation is high ($\frac{x_\gamma}{x_\tau} \geq \overline{\alpha}$), the benefit of the reduced marginal cost of taxation (driven by the high elasticity of demand) is overcome by the effectiveness of education and the two instruments are substitute again. As in the case of inelastic demand, an increase in λ calls for a decreased reliance on taxation and an increased reliance on education.

When the effectiveness of education relative to taxation is low ($\frac{x_\gamma}{x_\tau} \leq \underline{\alpha}$), the benefit of the reduced marginal cost of taxation (driven by the high elasticity of demand) overcomes the effectiveness of education and the two instruments are also substitutes. However, unlike the case of inelastic demand, an increase in λ calls now for an increased reliance on taxation and a lowered reliance on education.

When education is mildly effective relative to taxation ($\underline{\alpha} < \frac{x_\gamma}{x_\tau} < \overline{\alpha}$), both instruments are effective in reducing the consumption of the sin good and their use increases with λ. Only in this case they are complements.

5 Pessimistic consumers

So far we have assumed that agents underestimate the likelihood of health damages caused by sin good consumption. There is some evidence in support of this assumption. Krosnick et al. (2017), for example, in a study investigating people's perception of getting lung cancer if they smoke, show that 54.6% of the respondents vastly underestimate it. However, although in prevalence people underestimate their risk, there are in general also cases of people who overestimate it. For instance, 23.9% of the respondents of this same study overestimate their chance of getting lung cancer.[11]

In this section, within a setting in which consumers are mainly optimistic, we introduce pessimistic consumers in order to highlight a possible flaw of taxation. To this aim, we assume that a fraction μ of agents is pessimistic and overestimate the probability of health damages.

We denote by $q^p \geq q$ the perceived probability that the sin good causes health damages for a pessimistic agent and by $q^o \leq q$ the perceived probability that the sin good causes health damages for an optimistic one. Moreover, we assume that the misperception of optimistic and pessimistic agents is symmetric, i.e., $|q^p - q| = |q^o - q| \equiv \overline{\delta}$, with $\overline{\delta} < \min\{q, 1-q\}$.[12]

An agent who overestimates the probability of health damages maximizes his *pessimistic* expected utility function $U^p \equiv v(x) - q^p c(x) + z$, subject to the budget constraint $I = p_x x + p_z z$. In the absence of policy measures aimed at affecting the consumers' behavior, the sin good

[11] Only about 1.5% of respondents perceive relative risk approximately correctly.

[12] A conservative stance in this regard seems reasonable since we ignore the extent of the misperception suffered by optimistic and pessimistic consumers.

consumption for a pessimistic agent, x^p, satisfies the first order condition $v_x(x^p) - q^p c_x(x^p) = 1$. Since v_x is decreasing in x, c_x is increasing in x and $v_x(x) - qc_x(x)$ is greater than $v_x(x) - q^p c_x(x)$ for any x, we have that x^p is lower than x^{FB}. Moreover, $z^p = I - x^p$ is larger than z^{FB}. Thus, because of pessimism, the agent consumes too little of the sin good and too much of the numeraire.

In what follows, we consider two settings: one in which only educational policies are available to correct consumers' misbehavior, and a second one in which both educational measures and taxation can be used.

Education initiatives. The aim of education is to provide information useful to improve the consumer's awareness by reducing agents' misperception. Formally:

Assumption 3 *The benefit of information provision is given by a continuous function*

$$q^a(\gamma) \equiv \begin{cases} q^o + \delta(\gamma) & \text{if the agent is optimistic (a=o)} \\ q^p - \delta(\gamma) & \text{if the agent is pessimistic (a=p)} \end{cases}$$

with $\delta(\cdot)$ *defined on* $[0, \bar{\gamma}]$, *strictly increasing and concave, and equal to* 0 *when* $\gamma = 0$ *and to* $\overline{\delta}$ *when* $\gamma = \bar{\gamma}$. *The disutility of regulation is given by a continuous function* $b(\gamma)$ *defined on* $(0, \bar{\gamma}]$ *with* $b(\gamma)$ *strictly increasing and convex, and such that* $\lim_{\gamma \to 0} b(\gamma) = 0$.

Under Assumption 3, when regulation is imposed, the agent's optimistic utility becomes

$$U^o(\gamma) = v(x) - \left(q - \overline{\delta} + \delta(\gamma)\right) c(x) - b(\gamma) + z, \tag{13}$$

and the agent's pessimistic utility becomes

$$U^p(\gamma) = v(x) - \left(q + \overline{\delta} - \delta(\gamma)\right) c(x) - b(\gamma) + z. \tag{14}$$

An increase in γ has a negative effect on the marginal optimistic utility and a positive effect on the marginal pessimistic utility ($U^o_{x\gamma}(\gamma) = -\delta_\gamma \, c_x < 0 < U^p_{x\gamma}(\gamma) = \delta_\gamma \, c_x$).

In the case of information provision γ, the optimal consumption bundle of a type a agent, which we denote by $(x^a(\gamma), \, z^a(\gamma))$ with $a \in \{o, p\}$, maximizes U^a subject to $z = I - x$. Then, the consumption rule of the sin good, $x^a(\gamma)$, satisfies the first order condition

$$v_x(x) - q^a(\gamma) \, c_x(x) = 1. \tag{15}$$

The social planner chooses the level of information provision $\hat{\gamma}^{\mu}$ that maximizes the actual utility

$$\begin{aligned} U^{\mu}(\gamma) &= \mu\left(v\left(x^{p}(\gamma)\right)-\left(q+\overline{\delta}-\delta(\gamma)\right)c\left(x^{p}(\gamma)\right)+z^{p}(\gamma)\right)+ \\ &\quad +(1-\mu)\left(v\left(x^{o}(\gamma)\right)-\left(q-\overline{\delta}+\delta(\gamma)\right)c\left(x^{o}(\gamma)\right)+z^{o}(\gamma)\right)-b(\gamma) \end{aligned}$$

subject to the budget constraints $I-x^{o}(\gamma)=z^{o}(\gamma)$, and $I-x^{p}(\gamma)=z^{p}(\gamma)$.

By substituting the budget constraints in the actual utility function, the latter can be written as

$$\begin{aligned} \Omega^{\mu}(\gamma) &= [\mu\underbrace{\left(v\left(x^{p}(\gamma)\right)-qc\left(x^{p}(\gamma)\right)-x^{p}(\gamma)+I\right)}_{BI^{p}(\gamma)}+ \\ &\quad (1-\mu)\underbrace{\left(v\left(x^{o}(\gamma)\right)-qc\left(x^{o}(\gamma)\right)-x^{o}(\gamma)+I\right)}_{BI^{o}(\gamma)}]-\underbrace{b(\gamma)}_{CI^{\mu}(\gamma)}. \end{aligned}$$

The term in square brackets, $BI^{\mu}(\gamma)\equiv\mu BI^{p}(\gamma)+(1-\mu)BI^{o}(\gamma)$, is the benefit of information provision, defined as the expected utility that would be obtained by inducing levels of consumption closer to the first-best, i.e., $x^{p}\in(x^{p},x^{FB})$ and $x^{o}(\gamma)\in(x^{FB},x^{o})$. The second term, $CI^{\mu}(\gamma)$, represents the cost induced by the disutility of information provision.

The social planner's problem is to choose $\hat{\gamma}^{\mu}$ that maximizes the distance between the benefits and the costs associated to information provision.

Proposition 5 *Suppose there is a small fraction $\mu>0$ of pessimistic consumers in the population. If only education is available, the optimal level of information provision is $\hat{\gamma}^{\mu}<\hat{\gamma}$.*

By reducing the distance between perceived and actual likelihoods of health damages (caused by sin good consumption) education makes the agents' consumption closer to the first-best and increases utility. In equilibrium, the marginal cost of information provision equals the average marginal benefit, $BI^{\mu}_{\gamma}\equiv\mu BI^{p}_{x}x^{p}_{\gamma}+(1-\mu)BI^{o}_{x}x^{o}_{\gamma}$, where $BI^{a}_{x}\equiv v_{x}(x)-q\,c_{x}(x)-1$ is the social planner's marginal utility evaluated at $x=x^{a}(\gamma)$ and $x^{a}_{\gamma}\equiv\frac{q^{a}_{\gamma}c_{x}(x^{a}(\gamma))}{v_{xx}(x^{a}(\gamma))-q^{a}(\gamma)c_{xx}(x^{a}(\gamma))}$, with $a\in\{o,p\}$, is the effectiveness of information provision in aligning optimistic and pessimistic consumption to the first-best. With respect to the case where all consumers are optimistic, the presence of pessimistic consumers modifies both i) the social planner's marginal utility and ii) the effectiveness of information provision. In particular, as shown in the proof of Proposition 5, the social marginal benefit obtained by reducing the consumption of an optimistic agent, $|BI^{o}_{x}|$, is larger than the one obtained by increasing the consumption of a pessimistic agent, $|BI_x^p|$. As a consequence, the introduction of a fraction of pessimistic consumers lowers the social planner's marginal utility.[13] Similarly, the effectiveness of information provision decreases since educational measures are more effective in reducing the consumption of optimistic agents than in increasing that of pessimistic ones. Hence, the introduction of some pessimistic consumers leads to a level of information provision lower than that stated in Proposition 1, that is, $\hat{\gamma}^{\mu} < \hat{\gamma}$.

Education initiatives and taxes. Consider now the case in which taxation is also available to correct consumers' misbehavior. Notice that, when consumers are heterogeneous, implementing the first-best outcome would require individual-specific taxes and subsidies.[14] However, since this is unrealistic because of informational constraints, implementation costs and the like, we limit our analysis to a uniform tax. Recall that the goal of taxation is to limit the consumption (of optimistic individuals) by increasing the sin good price. However, this has the undesired effect of further reducing consumption also for the pessimistic agents.[15]

The social planner chooses the level of education, $\gamma^{\mu *}$, and the level of taxation, $\tau^{\mu *}$ maximizing the distance between the expected benefits and costs, i.e.,

$$B^{\mu}(\gamma, \tau) = \mu \underbrace{\left(v(x^p) - q\, c(x^p) + I - x^p\right)}_{B^p(\gamma,\tau)} + (1-\mu) \underbrace{\left(v(x^o) - q\, c(x^o) + I - x^o\right)}_{B^o(\gamma,\tau)}$$

and

$$CT^{\mu}(\gamma, \tau) = \mu \underbrace{\left(\lambda \tau x^p + b(\gamma)\right)}_{C^p(\gamma,\tau)} + (1-\mu) \underbrace{\left(\lambda \tau x^o + b(\gamma)\right)}_{C^o(\gamma,\tau)},$$ [16]

where $x^a \equiv x^a(\gamma, \tau)$ for all $a \in \{o, p\}$.

Proposition 6 states the main result of our paper, i.e. that the presence of a fraction, however small, of pessimistic agents lowers the effectiveness of taxation in mitigating overconsumption thereby inducing the social planner to boost up educational measures.

[13] The proof relies on the distance $\left|x^o(\gamma) - x^{FB}\right|$ being larger than $\left|x^p(\gamma) - x^{FB}\right|$, so that, from the concavity of the social planner utility function, the marginal benefit of reducing the sin good consumption of an optimistic agent, $|BI_x^o|$, is larger than the marginal benefit of increasing the sin good consumption of a pessimistic agent, $|BI_x^p|$.

[14] Clearly this is true even if the efficiency cost of taxation is absent, i.e., for $\lambda = 0$.

[15] Although in principle a large fraction of pessimistic agents in the sample could imply a negative tax (i.e., a subsidy), the proof of our main result assumes that the fraction of pessimistic agents is small, in line with the evidence, implicitly ruling out $\tau^* < 0$.

[16] We use a social welfare function that puts equal weight on all agents, that is, the expectation of the individual experienced utility.

Proposition 6 *Suppose there is a small fraction $\mu > 0$ of pessimistic consumers in the population. If both education and taxation are available, the optimal level of taxation is $\tau^{\mu *} < \tau^*$ and the optimal level of information provision is $\gamma^{\mu *} > \gamma^*$.*

The intuition behind this result can be grasped by considering that the optimal uniform tax for pessimistic agents would be negative. Thus, when both optimistic and pessimistic agents coexist the optimal tax level is smaller than the one prevailing when agents are all optimistic. The previous result also shows that the level of education is higher than the one that obtains with optimistic agents only, and the two instruments are substitutes. This can be ascribed to the fact that while education initiatives always get both optimistic and pessimistic consumers' perceptions closer to the true probability of health damages, taxation, although effective for optimistic types, distorts the choice of pessimistic ones. Moreover, relative to the case in which only educational measures are available, information provision is more valuable when taxation is also available. This is because taxation increases the distance between the pessimistic and the first-best consumption level, thus increasing the social benefit of inducing pessimistic agents to consume more.

Finally, notice that this result does not rest on the inefficiency of the tax system. lndeed, even with a fully efficient fiscal system, a tax levied on sin goods would still affect the consumption choice of pessimistic consumers in the "unwanted/undesired" direction, thus calling for the corrective effect of educational measures.

6 Conclusions

The paper studies the role of taxation and informational measures in a setting in which an optimistic agent underestimates the probability of health damages and thus consumes too much unhealthy goods. Depending on the elasticity of demand of the sin good with respect to taxation and the relative efficiency of educational measures, the paper shows that these two instruments can be used as substitutes or complements. However, when both optimistic consumers (who underestimate the probability of health damages and thus consume too much) and pessimistic consumers (who overestimate the probability of health damages and thus consume too little) coexist, the correcting effect that taxation has on optimistic consumers has unintended distorting effects on the choices of pessimistic ones. In this framework, educational measures, by aligning both optimistic and pessimistic consumers' perception closer to the true probability of health damages, are more effective than taxation and should be preferred. Thus, besides the relative efficiency of each instrument, the paper points to an additional and unexpected advantage of educational measures in correcting the distorting effects of taxation.

Appendix

Proof of Proposition 1. For all $\gamma \in [0, \overline{\gamma}]$, as long as v and c are thrice differentiable, $\Omega(\gamma)$ is continuous and twice differentiable. If strictly positive and lower than $\overline{\gamma}$, $\hat{\gamma}$ satisfies the first order condition $\Omega_\gamma(\hat{\gamma}) = \Omega_x(\hat{\gamma})x_\gamma(\hat{\gamma}) - b_\gamma(\hat{\gamma}) = 0$, where $\Omega_x(\hat{\gamma}) = [v_x(x(\hat{\gamma})) - q^o(\hat{\gamma})c_x(x(\hat{\gamma})) - 1]$. From (5), we can derive $\Omega_x = -(q - q^o(\gamma))c_x(x(\gamma)) < 0$. When $\gamma = 0$, $\Omega_\gamma = -(q - q^o)c_x(x^o)x_\gamma(0) > 0$. Hence, $\gamma = 0$ cannot be a corner solution of the social planner maximization problem. Moreover, $q(\overline{\gamma}) = q$ implies that when $\gamma = \overline{\gamma}$, $\Omega_x = 0$ and $\Omega_\gamma = -b_\gamma(\overline{\gamma}) < 0$. Hence, by the continuity of $\Omega(\gamma)$, there exists at least one $\hat{\gamma} \in (0, \overline{\gamma})$ satisfying condition $\Omega_\gamma(\hat{\gamma}) = 0$. Finally, $\Omega_{\gamma\gamma} < 0$. Indeed, $\Omega_{\gamma\gamma} = \Omega_x x_{\gamma\gamma} + \Omega_{x\gamma}x_\gamma - b_{\gamma\gamma} = \Omega_x x_{\gamma\gamma} + \Omega_{xx}x_\gamma^2 - b_{\gamma\gamma}$, with $\Omega_{xx} = v_{xx} - qc_{xx}$ and $x_{\gamma\gamma} = \frac{(q^o_{\gamma\gamma}c_x + q^o_\gamma c_{xx}x_\gamma)(v_{xx} - q^o(\gamma)c_{xx}) - (v_{xxx} - q^o(\gamma)c_{xxx})x_\gamma q^o_\gamma c_x + (q^o_\gamma)^2 c_x c_{xx}}{(v_{xx} - q^o(\gamma)c_{xx})^2}$. Since by assumption $q^o_\gamma > 0$, $c_x > 0$, $v_{xx} - q^o c_{xx} < 0$, $b_{\gamma\gamma} \geq 0$, $q^o_{\gamma\gamma} \leq 0$, and $v_{xxx} - q^o c_{xxx} \geq 0$, then $\Omega_{xx} < 0$, $x_{\gamma\gamma} > 0$, and then $\Omega_{\gamma\gamma} < 0$. ■

Proof of Proposition 2. For all $\tau \geq 0$, as long as v and c are thrice differentiable, $\Omega(\tau)$ is continuous and twice differentiable. If strictly positive, $\hat{\tau}$ satisfies the first order condition $\Omega_\tau(\hat{\tau}) = \Omega_x(\hat{\tau})x_\tau(\hat{\tau}) - \lambda x(\hat{\tau}) = 0$, where $\Omega_x(\hat{\tau}) = [v_x(x(\hat{\tau})) - qc_x(x(\hat{\tau})) - 1 - \lambda\hat{\tau}]$. From (8) we derive $\Omega_x = \tau(1-\lambda) - (q - q^o)c_x(x(\tau))$. When $\tau = 0$, $\Omega_\tau = -(q - q^o)c_x(x^o)/(v_{xx}(x^o) - q^o c_{xx}(x^o)) - \lambda x^o$, which is positive for all $\lambda < \lambda^M$. Moreover, Inada conditions for $v(x)$ together with $c_x(0) = 0$ imply $\lim_{\tau\to\infty} x(\tau) = 0$ and $\lim_{\tau\to\infty} \Omega_\tau = -\infty$. Hence, by continuity of $\Omega(\tau)$, there exists at least one $\hat{\tau} > 0$ satisfying the first order condition. Finally, $\Omega_{\tau\tau} = (\Omega_{xx}x_\tau - 2\lambda)\, x_\tau + \Omega_x x_{\tau\tau} < 0$ in $\tau = \hat{\tau}$. Indeed, $x_{\tau\tau} = -(v_{xxx} - q^o c_{xxx})x_\tau^3 > 0$ since $(v_{xxx} - q^o c_{xxx}) > 0$ by assumption, $\Omega_x(\hat{\tau}) < 0$ by the first order condition, and $\Omega_{xx}x_\tau - 2\lambda > 0$ since $\Omega_{xx}x_\tau = \frac{v_{xx} - qc_{xx}}{v_{xx} - q^o c_{xx}} > 1$ for all $q < q^o$, and $\lambda < 1/2$. ■

Proof of Proposition 3. By Kuhn-Tucker necessary conditions, if (γ^*, τ^*) maximizes the objective function (11) and $\gamma^* < \overline{\gamma}$, then the following equations and inequalities are satisfied:

$$\begin{cases} ((v_x(x^*) - qc_x(x^*) - 1 - \lambda\tau^*)x_\tau - \lambda x^*)\tau^* = 0 \\ ((v_x(x^*) - qc_x(x^*) - 1 - \lambda\tau^*)x_\gamma - b_\gamma)\gamma^* = 0 \\ (v_x(x^*) - qc_x(x^*) - 1 - \lambda\tau^*)x_\tau - \lambda x^* \leq 0 \qquad (KT1) \\ (v_x(x^*) - qc_x(x^*) - 1 - \lambda\tau^*)x_\gamma - b_\gamma \leq 0 \qquad (KT2) \\ \tau^* \geq 0 \\ \gamma^* \geq 0 \end{cases}, \tag{16}$$

with $x^* \equiv x(\gamma^*, \tau^*)$. We will show that there exists $(\gamma^*, \tau^*) \in (0, \hat{\gamma}) \times (0, \hat{\tau})$ that satisfies (16) in 4 Steps.

Step 1: $\Omega_{\tau\gamma} < 0$ *for all* (γ,τ) *such that* $v_x(x(\gamma,\tau)) - q\, c_x(x(\gamma,\tau)) - 1 - \lambda\tau < 0$. Let be $A \equiv \{(\gamma,\tau) \in (0,\hat{\gamma}) \times (0,\tau^{FB}) : v_x(x(\gamma,\tau)) - q\, c_x(x(\gamma,\tau)) - 1 - \lambda\tau < 0\}$. The cross derivative of $\Omega(\gamma,\tau)$ is

$$\Omega_{\tau\gamma} = (v_{xx} - q\, c_{xx})\, x_\gamma x_\tau + (v_x - q\, c_x - 1 - \lambda\tau)\, x_{\gamma\tau} - \lambda x_\gamma =$$

$$= x_\gamma((v_{xx} - q\, c_{xx})\, x_\tau - \lambda) + (v_x - q\, c_x - 1 - \lambda\tau)\, x_{\gamma\tau},$$

where $x_{\tau\gamma} = -x_\tau^2 q_\gamma^o((v_{xxx} - q^o(\gamma)\, c_{xxx}) x_\tau c_x - c_{xx}) > 0$ since $(v_{xxx} - q^o(\gamma)\, c_{xxx}) > 0$ and c_{xx} positive and near to 0 by assumption, and $((v_{xx} - q\, c_{xx})\, x_\tau - \lambda) > 0$ since $(v_{xx} - q\, c_{xx})\, x_\tau = \frac{v_{xx} - q\, c_{xx}}{v_{xx} - q^o(\gamma) c_{xx}} > 1$ for all $q^o(\gamma) < q$. Hence, $\Omega_{\tau\gamma} < 0\ \forall(\gamma,\tau) \in A$.

Step 2: Let be $\tau_1(\gamma)$ *the function implicitly defined by equation* $KT1(\gamma,\tau) \equiv (v_x(x(\gamma,\tau)) - qc_x(x(\gamma,\tau)) - 1 - \lambda\tau)x_\tau - \lambda x(\gamma,\tau) = 0$. *Then,* $\tau_1(0) = \hat{\tau}$ *and* $\tau_1(\hat{\gamma}) \in (0,\hat{\tau})$.

By substituting $\gamma = 0$ in $KT1(\gamma,\tau)$, gives $\tau_1(0) = \hat{\tau}$ by definition of $\hat{\tau}$. Moreover, by Step 1, $KT1(\hat{\gamma},\hat{\tau}) < 0$ and, by definition of $\hat{\gamma}$, $KT1(\hat{\gamma},0) = (v_x(x(\hat{\gamma},0)) - qc_x(x(\hat{\gamma},0)) - 1)x_\tau - \lambda x(\hat{\gamma},0) = \frac{b_\gamma(\hat{\gamma})}{q_\gamma^o(\hat{\gamma})\, c_x(x(\hat{\gamma}))} - \lambda x(\hat{\gamma})$, that is positive $\forall \lambda < \bar{\lambda}$. This implies that $\exists\ \tau' \in (0,\hat{\tau}) : KT1(\hat{\gamma},\tau') = 0$, and $\tau_1(\hat{\gamma}) = \tau'$.

Step 3: Let be $\tau_2(\gamma)$ *the function implicitly defined by equation* $KT2(\gamma,\tau) \equiv (v_x(x(\gamma,\tau)) - qc_x(x(\gamma,\tau)) - 1 - \lambda\tau)x_\gamma - b_\gamma = 0$. *Then,* $\tau_2(0) \in (\hat{\tau},\tau^{FB})$ *and* $\tau_2(\hat{\gamma}) = 0$.

By substituting $\gamma = \hat{\gamma}$ in $KT2(\gamma,\tau)$, gives $\tau_2(\hat{\gamma}) = 0$ by definition of $\hat{\gamma}$. Moreover, by definition of $\hat{\tau}$, $KT2(0,\hat{\tau}) = (v_x(x(0,\hat{\tau})) - qc_x(x(0,\hat{\tau})) - 1 - \lambda\hat{\tau})x_\gamma - b_\gamma(0) = \lambda x(\hat{\tau}) q_\gamma^o(0)\, c_x(x(\hat{\tau})) > 0$ and, by definition of τ^{FB}, $KT2(0,\tau^{FB}) = (v_x(x(0,\tau^{FB})) - qc_x(x(0,\tau^{FB})) - 1 - \lambda\tau^{FB})x_\gamma - b_\gamma(0) = -\lambda\tau^{FB} < 0$. This implies that $\exists\ \tau'' \in (\hat{\tau},\tau^{FB}) : KT2(0,\tau'') = 0$, and $\tau_2(0) = \tau''$.

Step 4: There exists $\gamma^* \in (0,\hat{\gamma})$ *such that* $\tau_1(\gamma^*) = \tau_2(\gamma^*)$. *Moreover,*$\tau_1(\gamma^*) \in (0,\hat{\tau})$.

Let be $\Delta(\gamma) \equiv \tau_1(\gamma) - \tau_2(\gamma)$. $\Delta(0) = \hat{\tau} - \tau'' < 0$ and $\Delta(\hat{\gamma}) = \tau' - 0 > 0$. Hence, by Bolzano's theorem, there exists $\gamma^* \in (0,\hat{\gamma})$ such that $\Delta(\gamma^*) = 0$. Moreover, $\tau_1(\gamma^*) < \tau_1(0) = \hat{\tau}$

To conclude the proof we will show that $\Omega(\gamma,\tau)$ is concave in $(\gamma,\tau) = (\gamma^*,\tau^*)$, that is $\Omega_{\tau\tau} \leq 0$, $\Omega_{\gamma\gamma} \leq 0$, and $\Omega_{\tau\tau}\Omega_{\gamma\gamma} \geq \Omega_{\tau\gamma}^2$. $\Omega_{\tau\tau} = (v_{xx} - q\, c_{xx})\, x_\tau^2 - 2\lambda x_\tau + (v_x - q\, c_x - 1 - \lambda\tau)\, x_{\tau\tau} = x_\tau \left(\frac{v_{xx} - q\, c_{xx}}{v_{xx} - q^o(\gamma) c_{xx}} - \lambda\right) + (v_x - q\, c_x - 1 - \lambda\tau)\, x_{\tau\tau} < 0$ for all $\lambda < 1/2$, $q^o(\gamma) < q$, and $(\gamma,\tau) \in A$. $\Omega_{\gamma\gamma} = (v_{xx} - q\, c_{xx})\, x_\gamma^2 - b_{\gamma\gamma} + (v_x - q\, c_x - 1 - \lambda\tau)\, x_{\gamma\gamma} < 0$ for all $(\gamma,\tau) \in A$. ■

Proof of Proposition 4. If the social planner optimization problem defined on page 12 involves an interior solution, (γ^*, τ^*) satisfies the following equations:

$$\begin{cases} (v_x - qc_x - 1 - \lambda\tau)x_\tau - \lambda x = 0 \\ (v_x - qc_x - 1 - \lambda\tau)x_\gamma - b_\gamma = 0. \end{cases} \tag{17}$$

From the implicit function theorem,

$$\begin{pmatrix} \partial\tau/\partial\lambda \\ \partial\gamma/\partial\lambda \end{pmatrix} = -\,\mathbf{H^{-1}}\,\mathbf{D_\lambda}, \tag{18}$$

where $\mathbf{H^{-1}}$ is the inverse of the Hessian matrix of $\Omega(\gamma, \tau)$ evaluated at (γ^*, τ^*), that is,

$$\mathbf{H^{-1}} = \frac{1}{\Omega_{\tau\tau}\Omega_{\gamma\gamma} - \Omega_{\tau\gamma}^2} \begin{pmatrix} \Omega_{\gamma\gamma} & -\Omega_{\tau\gamma} \\ -\Omega_{\tau\gamma} & \Omega_{\tau\tau} \end{pmatrix},$$

and $\mathbf{D_\lambda}$ is the vector of the first derivatives of $\nabla\Omega$ with respect to λ, evaluated at (γ^*, τ^*), that is,

$$\mathbf{D_\lambda} = \begin{pmatrix} \Omega_{\tau\lambda} \\ \Omega_{\gamma\lambda} \end{pmatrix} = \begin{pmatrix} -(\frac{1}{\eta_{x,\tau}} + 1)\tau x_\tau \\ -\tau x_\gamma \end{pmatrix} = -\tau x_\tau \begin{pmatrix} (\frac{1}{\eta_{x,\tau}} + 1) \\ q_\gamma^o c_x \end{pmatrix}.$$

Substituting in (18), gives

$$\begin{pmatrix} \partial\tau/\partial\lambda \\ \partial\gamma/\partial\lambda \end{pmatrix} = \frac{-1}{\Omega_{\tau\tau}\Omega_{\gamma\gamma} - \Omega_{\tau\gamma}^2} \begin{pmatrix} \Omega_{\gamma\gamma} & -\Omega_{\tau\gamma} \\ -\Omega_{\tau\gamma} & \Omega_{\tau\tau} \end{pmatrix} (-\tau x_\tau) \begin{pmatrix} (\frac{1}{\eta_{x,\tau}} + 1) \\ q_\gamma^o c_x \end{pmatrix} =$$

$$= \frac{\tau x_\tau}{\Omega_{\tau\tau}\Omega_{\gamma\gamma} - \Omega_{\tau\gamma}^2} \begin{pmatrix} \Omega_{\gamma\gamma}(\frac{1}{\eta_{x,\tau}} + 1) - \Omega_{\tau\gamma} q_\gamma^o c_x \\ -\Omega_{\tau\gamma}(\frac{1}{\eta_{x,\tau}} + 1) + \Omega_{\tau\tau} q_\gamma^o c_x \end{pmatrix}.$$

From the proof of Proposition 3, we know that $\Omega_{\gamma\tau} < 0$. From the concavity of $\Omega(\gamma, \tau)$, it follows $\Omega_{\gamma\gamma} < 0$, $\Omega_{\tau\tau} < 0$, and $\Omega_{\tau\tau}\Omega_{\gamma\gamma} - \Omega_{\tau\gamma}^2 > 0$ or, equivalently, $\frac{\Omega_{\gamma\gamma}}{\Omega_{\tau\gamma}} > \frac{\Omega_{\tau\gamma}}{\Omega_{\tau\tau}}$. Moreover, $x_\tau < 0$, $q_\gamma^o c_x > 0$, $\eta_{x,\tau} < 0$, and $(\frac{1}{\eta_{x,\tau}} + 1) > 0$ iff $\eta_{x,\tau} < -1$. Hence,

$$\begin{cases} \partial\tau/\partial\lambda \geq 0 \quad \textit{iff} \quad q_\gamma^o c_x \leq \frac{\Omega_{\gamma\gamma}}{\Omega_{\tau\gamma}}(\frac{1}{\eta_{x,\tau}} + 1) \\ \partial\gamma/\partial\lambda \geq 0 \quad \textit{iff} \quad q_\gamma^o c_x \geq \frac{\Omega_{\tau\gamma}}{\Omega_{\tau\tau}}(\frac{1}{\eta_{x,\tau}} + 1). \end{cases}$$

If $\eta_{x,\tau} > -1$, then $\frac{\Omega_{\gamma\gamma}}{\Omega_{\tau\gamma}}(\frac{1}{\eta_{x,\tau}} + 1) < 0$, and $\frac{\Omega_{\tau\gamma}}{\Omega_{\tau\tau}}(\frac{1}{\eta_{x,\tau}} + 1) < 0$. Since $q_\gamma^o c_x > 0$, this implies $\partial\tau/\partial\lambda \leq 0$ and $\partial\gamma/\partial\lambda \geq 0$.

If $\eta_{x,\tau} < -1$, then $\frac{\Omega_{\gamma\gamma}}{\Omega_{\tau\gamma}}(\frac{1}{\eta_{x,\tau}} + 1) > \frac{\Omega_{\tau\gamma}}{\Omega_{\tau\tau}}(\frac{1}{\eta_{x,\tau}} + 1) > 0$. Let be $\overline{\alpha} \equiv \frac{\Omega_{\gamma\gamma}}{\Omega_{\tau\gamma}}(\frac{1}{\eta_{x,\tau}} + 1)$ and $\underline{\alpha} \equiv \frac{\Omega_{\tau\gamma}}{\Omega_{\tau\tau}}(\frac{1}{\eta_{x,\tau}} + 1)$. Then,

$$\left\{ \begin{array}{l} \partial\tau/\partial\lambda > 0 \\ \partial\gamma/\partial\lambda < 0 \end{array} \right. \quad \textit{iff} \ q_\gamma^o c_x < \underline{\alpha}$$

$$\begin{cases} \partial\tau/\partial\lambda > 0 \\ \partial\gamma/\partial\lambda \geq 0 \end{cases} \textit{iff } \underline{\alpha} \leq q^o_\gamma c_x < \overline{\alpha}$$

$$\begin{cases} \partial\tau/\partial\lambda \leq 0 \\ \partial\gamma/\partial\lambda > 0 \end{cases} \textit{iff } q^o_\gamma c_x \geq \overline{\alpha}.$$

To conclude the proof notice that $q^o_\gamma c_x = \frac{x_\gamma}{x_\tau}$. ■

Proof of Proposition 5. For all $\gamma \in [0, \overline{\gamma}]$, as long as v, and c are thrice differentiable, $\Omega^\mu(\gamma)$ is continuous and twice differentiable. Moreover, $\Omega^\mu_\gamma = \mu\Omega^p_x x^p_\gamma + (1-\mu)\,\Omega^o_x x^o_\gamma - b_\gamma$, with $\Omega^a_x = [v_x(x^a(\gamma)) - qc_x(x^a(\gamma)) - 1]$ for any $a \in \{o, p\}$. From (15), we can derive $\Omega^o_x(\gamma) = -(\overline{\delta} - \delta(\gamma))c_x(x^o(\gamma)) < 0$, and $\Omega^p_x(\gamma) = (\overline{\delta} - \delta(\gamma))c_x(x^p(\gamma)) > 0$. Moreover, $x^o_\gamma < 0$ and $x^p_\gamma > 0$ since $q^o_\gamma = \delta_\gamma > 0$, $q^p_\gamma = -\delta_\gamma < 0$, $c_x > 0$, and $v_{xx} - q^a c_{xx} < 0$ for any $a \in \{o, p\}$ by assumption. Hence, both $\Omega^p_x x^p_\gamma$ and $\Omega^o_x x^o_\gamma$ are positive.

When $\gamma = 0$, $\Omega^\mu_\gamma = \mu\overline{\delta}c_x(x^p)x^p_\gamma - (1-\mu)\,\overline{\delta}c_x(x^o)x^o_\gamma > 0$ if $\mu \to 0$. Moreover, $q^p(\overline{\gamma}) = q^o(\overline{\gamma}) = q$ implies that when $\gamma = \overline{\gamma}$, $\Omega^\mu_x(\overline{\gamma}) = 0$ and $\Omega^\mu_\gamma = -b_\gamma(\overline{\gamma}) < 0$. Hence, $\gamma = 0$ and $\gamma = \overline{\gamma}$ cannot be corner solutions of the social planner maximization problem, and there exists $\hat{\gamma}^\mu \in (0, \overline{\gamma})$ satisfying condition $\Omega^\mu_\gamma(\hat{\gamma}^\mu) = 0$. When $\mu = 0$, $\hat{\gamma}^\mu = \hat{\gamma}$. From the implicit function theorem, into a neighborhood of $\gamma = \hat{\gamma}$ and $\mu = 0$,

$$\partial\hat{\gamma}^\mu/\partial\mu = -\frac{(\overline{\delta} - \delta(\hat{\gamma}))(c_x(x^p(\hat{\gamma}))x^p_\gamma(x^p(\hat{\gamma})) + c_x(x^o(\hat{\gamma}))x^o_\gamma(x^o(\hat{\gamma})))}{\Omega^o_{\gamma\gamma}},$$

which is lower than zero iff $c_x(x^p(\hat{\gamma}))x^p_\gamma(x^p(\hat{\gamma})) < -c_x(x^o(\hat{\gamma}))x^o_\gamma(x^o(\hat{\gamma}))$, that is always if c_{xx} is positive and small enough. Indeed, $c_x(x^p(\hat{\gamma})) < c_x(x^o(\hat{\gamma}))$ since $c_{xx} > 0$ and $x^p(\hat{\gamma}) < x^o(\hat{\gamma})$. Moreover, if the health cost function is not too convex, then the concavity of the consumer's utility functions implies $\left|x^o_\gamma(\gamma)\right| \geq \left|x^p_\gamma(\gamma)\right|$ for any γ. Indeed, the effectiveness of information provision of optimistic and pessimistic agents differ for two reasons. On the one hand, for any level of sin good consumption x, $\frac{\delta_\gamma c_x(x)}{|v_{xx}(x) - q^o(\gamma)c_{xx}(x)|} < \frac{\delta_\gamma c_x(x)}{|v_{xx}(x) - q^p(\gamma)c_{xx}(x)|}$ since $q^p > q^p$ and $v_{xx}(x) - q^o(\gamma)c_{xx}(x) < v_{xx}(x) - q^p(\gamma)c_{xx}(x) < 0$. On the other hand, both $\frac{\delta_\gamma c_x(x)}{|v_{xx}(x) - q^o(\gamma)c_{xx}(x)|}$ and $\frac{\delta_\gamma c_x(x)}{|v_{xx}(x) - q^p(\gamma)c_{xx}(x)|}$ are increasing in x and $x^o(\gamma) > x^p(\gamma)$. If c_{xx} is small, the second effect prevails on the first. As a consequence, pessimistic consumers lower the social planner's marginal benefit of information provision and reduce the optimal γ. ■

Proof of Proposition 6. By Kuhn-Tucker necessary conditions, if $(\gamma^{\mu*}, \tau^{\mu*}) > 0$ maximize $\Omega^\mu(\gamma, \tau) \equiv B^\mu(\gamma, \tau) - CT^\mu(\gamma, \tau)$ and $\gamma^{\mu*} < \overline{\gamma}$, then the following equations and inequalities are satisfied:

$$\begin{cases} (1-\mu)\Omega^o_\tau(\gamma^{\mu*},\, \tau^{\mu*}) + \mu\Omega^p_\tau(\gamma^{\mu*},\, \tau^{\mu*}) = 0 \\ (1-\mu)\Omega^o_\gamma(\gamma^{\mu*},\, \tau^{\mu*}) + \mu\Omega^p_\gamma(\gamma^{\mu*},\, \tau^{\mu*}) = 0 \end{cases} \tag{19}$$

with $\Omega^a_\tau(\gamma^{\mu*},\, \tau^{\mu*}) \equiv (v_x(x^{a*}) - qc_x(x^{a*}) - 1 - \lambda\tau^{\mu*})x^{a*}_\tau - \lambda x^{a*}$, $\Omega^a_\gamma(\gamma^{\mu*},\, \tau^{\mu*}) \equiv (v_x(x^{a*}) - qc_x(x^{a*}) - 1 - \lambda\tau^*)x^{a*}_\gamma - b_\gamma)$, $x^{a*}_\gamma \equiv \frac{q^a_\gamma(\gamma^{\mu*})c_x(x^{a*})}{v_{xx}(x^{a*})-q^a(\gamma^{\mu*})c_{xx}(x^{a*})}$, $x^{a*}_\tau \equiv \frac{1}{v_{xx}(x^{a*})-q^a(\gamma^{\mu*})c_{xx}(x^{a*})}$, and $x^{a*} \equiv x^{a*}(\gamma^{\mu*}, \tau^{\mu*})$ for any $a \in \{o,\, p\}$.

Let be $\Omega^\mu(\gamma,\tau) \equiv \Omega(\gamma,\tau,\mu)$. We will prove the proposition in 3 steps.

Step 1: $\Omega_{\tau\mu}(\gamma^*,\tau^*,0) < 0$.

$$\Omega_{\tau\mu}(\gamma^*,\tau^*,0) = \Omega^p_\tau(\gamma^*,\tau^*) - \Omega^o_\tau(\gamma^*,\tau^*) = \Omega^p_x x^p_\tau - \lambda x^p(\gamma^*,\tau^*)) < 0,$$

with $\Omega^p_x = ((q^p(\gamma^*) - q)c_x(x^p(\gamma^*,\tau^*)) + (1-\lambda)\tau^*)$. Indeed, $\Omega^o_\tau(\gamma^*,\tau^*) = 0$ by definition of (γ^*,τ^*), $q^p(\gamma^*) > q$ by assumption, and $x^p_\tau < 0$.

Step 2: $\Omega_{\gamma\mu}(\gamma^*,\tau^*,0) > 0$. Let be $\Omega^a_x = ((q^a(\gamma^*) - q)c_x(x^a(\gamma^*,\tau^*)) + (1-\lambda)\tau^*)$ for any $a \in \{o,\, p\}$.

$$\Omega_{\gamma\mu}(\gamma^*,\tau^*,0) = \Omega^p_\gamma(\gamma^*,\tau^*) - \Omega^o_\gamma(\gamma^*,\tau^*) =$$

$$= \frac{-\Omega^p_x\delta_\gamma(\gamma^*)c_x(x^p(\gamma^*,\tau^*))}{v_{xx}(x^p(\gamma^*,\tau^*)) - q^p(\gamma^*)c_{xx}(x^p(\gamma^*,\tau^*))} - \frac{\Omega^o_x\delta_\gamma(\gamma^*)c_x(x^o(\gamma^*,\tau^*))}{v_{xx}(x^o(\gamma^*,\tau^*)) - q^o(\gamma^*)c_{xx}(x^o(\gamma^*,\tau^*))},$$

which is higher than zero iff

$$\frac{v_{xx}(x^p(\gamma^*,\tau^*)) - q^p(\gamma^*)c_{xx}(x^p(\gamma^*,\tau^*))}{v_{xx}(x^o(\gamma^*,\tau^*)) - q^o(\gamma^*)c_{xx}(x^o(\gamma^*,\tau^*))} < \frac{c_x(x^p(\gamma^*,\tau^*))}{c_x(x^o(\gamma^*,\tau^*))}\frac{\Omega^p_x}{-\Omega^o_x}. \tag{20}$$

The left hand side of (20) is lower than 1 since $\frac{\partial(v_{xx}-q\,c_{xx})}{\partial q} = (v_{xxx} - q\,c_{xxx})x_q - c_{xx} < 0$. The right hand side of (20) is higher than 1 since c_{xx} is near to 0 by assumption and then the distance between $c_x(x^p(\gamma^*,\tau^*))$ and $c_x(x^o(\gamma^*,\tau^*))$ is small. Indeed, if $|c_x(x^p(\gamma^*,\tau^*)) - c_x(x^o(\gamma^*,\tau^*))| < \epsilon$, with $\epsilon > 0$ and small enough,

$$\frac{\Omega^p_x}{-\Omega^o_x} = \frac{(\bar\delta - \delta(\gamma))c_x(x^p(\gamma^*,\tau^*)) + (1-\lambda)\tau^*}{(\bar\delta - \delta(\gamma))c_x(x^o(\gamma^*,\tau^*)) - (1-\lambda)\tau^*} > 1$$

since $2\,(1-\lambda)\tau^* > ((\bar\delta - \delta(\gamma))(c_x(x^o(\gamma^*,\tau^*)) - c_x(x^p(\gamma^*,\tau^*)))$, and $\frac{c_x(x^p(\gamma^*,\tau^*))}{c_x(x^o(\gamma^*,\tau^*))}$ is near to 1.

Step 3: $\tau^{\mu*} < \tau^*$ *and* $\gamma^{\mu*} > \gamma^*$. The implicit function theorem into a neighborhood of $\gamma = \gamma^*$, $\tau = \tau^*$, and $\mu = 0$ gives:

$$\begin{cases} \partial\gamma^{\mu*}/\partial\mu \propto \Omega^o_{\tau\gamma}\Omega^p_\tau - \Omega^o_{\tau\tau}\Omega^p_\gamma & > & 0 \\ \partial\tau^{\mu*}/\partial\mu \propto \Omega^o_{\tau\gamma}\Omega^p_\gamma - \Omega^o_{\gamma\gamma}\Omega^p_\tau & < & 0 \end{cases}$$

from Steps 1 and 2, and because $\Omega^o_{\tau\gamma} < 0$, $\Omega^o_{\gamma\gamma} < 0$ and $\Omega^o_{\tau\tau} < 0$. This concludes the proof. ■

References

[1] Ainslie, George (1992), Picoeconomics. Cambridge University Press.

[2] Babor, Thomas, Caetano, Raul, Casswell, Sally, Edwards, Griffith, Giesbrecht, Norman, Graham, Kathryn, Grube, Joel, Hill, Linda, Holder, Harold, Homel, Ross, Livingston, Michael, Österberg, Esa, Rehm, Jürgen, Room, Robin, and Rossow, Ingeborg (2003), "Alcohol: No Ordinary Commodity. Research and Public Policy," New York: Oxford University Press.

[3] Bénabou, Roland (2012) "Groupthink: Collective delusions in organizations and markets," *Review of Economic Studies*, **80**, 429-462.

[4] Cioffi, Catherine E., Levitsky, David A., Pacanowski, Carly R., and Bertz, Fredrik (2015), "A Nudge in a Healthy Direction: the effect of nutrition labels on food purchasing behaviors in university dining facilities," *Appetite*, **92**.

[5] Gervais, Simon, Heaton, J. B., and Odean, Terrance (2003), "Overconfidence, Investment Policy, and Executive Stock Options,"mimeo.

[6] Goel, A. M., and Thakor, Anjan V. (2008), "Overconfidence, CEO Selection and Corporate Governance," *Journal of Finance*, **53**, pp. 2737–2784.

[7] Guiso, Luigi and Paiella, Monica (2008), "Risk Aversion, Wealth, and Background Risk". *Journal of the European Economic Association*, **6**, 1109-1150.

[8] Gruber, Jonathan, and Koszegi, Botond (2001), "Is addiction 'rational?' Theory and evidence," *Quarterly Journal of Economics*, **116**, 1261-1305.

[9] Gruber, Jonathan, and Koszegi, Botond (2004), "Tax incidence when individuals are time-inconsistent: The case of cigarette excise taxes," *Journal of Public Economics*, **88**, 1959-1987.

[10] Gruber, Jonathan, and Mullainathan, Sendhil (2005), "Do cigarette taxes make smokers happier?" *Advances in Economic Analysis and Policy*, **5**.

[11] Immordino, Giovanni, Menichini, Anna Maria C., and Romano, Maria Grazia (2019), "Taxing and Regulating Vices," *Scandinavian Journal of Economics*, forthcoming (https://onlinelibrary.wiley.com/doi/epdf/10.1111/sjoe.12334).

[12] Krosnick, Jon, Malhotra, Neil, Mo, Cecilia, Bruera, Eduardo, Chang, LinChiat, Pasek, John, and Thomas, Randall K. (2017), "Perceptions of health risks of cigarette smoking: A new measure reveals widespread misunderstanding". PloS one, **12**.

[13] Malmendier, Ulrike and Tate, Geoffrey (2005), "CEO Overconfidence and Corporate Investment," *Journal of Finance*, **60**, 2661-2700.

[14] Malmendier, Ulrike and Tate, Geoffrey (2008), "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction," *Journal of Financial Economics*, **89**, 20-43.

[15] O'Donoghue, Ted and Rabin, Matthew (2003), "Studying optimal paternalism, illustrated by a model of sin taxes," *American Economic Review*, Papers and Proceedings, **93**, 186-191.

[16] O'Donoghue, Ted and Rabin, Matthew (2006), "Optimal sin taxes," *Journal of Public Economics*, **90**, 1825–49.

[17] OECD (2011), Tax Administration in OECD and Selected Non-OECD Countries: Comparative Information Series (2010), OECD Publishing, Paris.

[18] PricewaterhouseCoopers (2015), *Paying Taxes: The global picture*, World Bank Group.

[19] Roll, Richard (1986), "The Hubris Hypothesis of Corporate Takeovers," *Journal of Business*, **59**, 197-216.

[20] Thomas, Roger, McLellan, Julie, and Perera, Rafael (2013), "School-based programmes for preventing smoking," *Cochrane Database of Systematic Review*, **30**.