DISCUSSION PAPER SERIES

No. 6777

TAXING SIN GOODS AND SUBSIDIZING HEALTH CARE

Helmuth Cremer, Philippe De Donder,
Darío Maldonado and Pierre Pestieau

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP6777.asp**

ISSN 0265-8003

TAXING SIN GOODS AND SUBSIDIZING HEALTH CARE

Helmuth Cremer, Toulouse School of Economics and CEPR
Philippe De Donder, Toulouse School of Economics and CEPR
Darío Maldonado, Universidad de Rosario
Pierre Pestieau, Université de Liege; Université Catholique de Louvain; PSE and CEPR

Discussion Paper No. 6777
April 2008

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 6777

April 2008

ABSTRACT

Taxing Sin Goods and Subsidizing Health Care*

We consider a two-period model. In the first period, individuals consume two goods: one is sinful and the other is not. The sin good brings pleasure but has a detrimental effect on second period health and individuals tend to underestimate this effect. In the second period, individuals can devote part of their saving to improve their health status and thus compensate for the damage caused by their sinful consumption. We consider two alternative specifications concerning this second period health care decision: either individuals acknowledge that they have made a mistake in the first period out of myopia or ignorance, or they persist in ignoring the detrimental effect of their sinful consumption. We study the optimal linear taxes on sin good consumption, saving and health care expenditures for a paternalistic social planner. We compare those taxes in the two specifications. We show under which circumstances the first best outcome can be decentralized and we study the second best taxes when saving is unobservable.

JEL Classification: H21 and I18
Keywords: behavioral economics, dual vs single self and paternalism

Helmuth Cremer
School of Economics
Manufacture des Tabacs
Aile Jean-Jacques Laffont
Accueil MF 404, 21 Allée de Brienne
F-31000 Toulouse
FRANCE
Email: helmut@cict.fr

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=124993

Philippe De Donder
IDEI & GREMAQ-CNRS
School of Economics
Manufacture des Tabacs, Office: MF418
21 Allée de Brienne
31000 Toulouse
FRANCE
Email: dedonder@cict.fr

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=157241

Darío Maldonado
Department of Economics
Universidad del Rosario
Calle 14 No. 4-69
Bogotá
Colombia
Email: dario.maldonadoca@urosario.edu.co

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=158200

Pierre Pestieau
Centre de Recherche en Economie Publique
et de la Population (CREPP)
Université de Liège
Boulevard du Rectorat, 7 (B31)
B-4000 Liège
BELGIUM
Email: p.pestieau@ulg.ac.be

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=105854

* We thank participants to Queen's University's 2007 Public Economics Workshop for helpful comments and suggestions. Part of this paper was written while the second author was visiting Yale University. He thanks Yale Economics Department for its hospitality.

Submitted 26 March 2008

1 Introduction

In our everyday life we consume a number of goods that all bring us utility. For most of them, that is all. For some, today's consumption can have some effects on tomorrow's health. For example, smoking leads to shorter lives or excess sugar to diabetes. To the extent that we impose costs on ourselves, there is no need for government action except if out of ignorance or myopia we do not take into account the delayed damage done to our health.[1] If this is the case, then there is a "paternalistic" mandate for public action, assuming that the government has a correct perception of the health damage generated by our sinful consumption.

Optimal sin taxes have been studied by O'Donoghue and Rabin (2003, 2006). They model an economy where individuals have hyperbolic preferences and differ in both their taste for the sin good and in their degree of time-inconsistency. They show how (heterogeneity in) time inconsistency affects the optimal (Ramsey) consumption tax policy. Their main insight is that, "although taxes create consumption distortion for fully self-controlled people, such distortions are second-order relative to the benefits from reducing over-consumption by people with self-control problems" (O'Donoghue and Rabin, 2006, p. 1827). Gruber and Koszegi (2001) study a Pigouvian tax used to counteract over-consumption due to self-control problems, and apply their model to the determination of optimal cigarette taxes. Gruber and Koszegi (2004) also study cigarette taxation with self-control problems, but their focus is the tax incidence for different income groups rather than optimal taxes.

O'Donoghue and Rabin (2003, 2006) are representative of the literature studying present-biased preferences (such as Laibson (1997)) in that 1) they assume that all non-biased individuals disapprove or regret their past consumption decisions and 2) there is nothing that agents can do to mitigate the current impact of past consumption decisions. Our objective in this paper is to lift those two assumptions and study their consequences on optimal sin taxes.

We model a two-period situation where individuals consume a sin good, with positive

[1] We are not concerned here by two important issues: addiction and externalities associated with sinful activities.

immediate gratification but negative impact on second stage health status. In the second stage, individuals may invest in health care services that have a positive impact on their health status. Individuals differ in income and in how much they take into account the link between sin good consumption and health care on the one hand, and health status on the other hand.

We contrast two possibilities. In the first one, individuals realize in the second period the mistake they did previously, regret their past high sin good consumption, and invest in health care understanding its correct impact on their health and utility. In other words, individuals suffer from myopia in the first period, but use their true or correct preferences later on. They thus exhibit "dual selves", using a term coined in the behavioral literature. A second possibility is that these individuals never take into account the true impact of sin good consumption and health expenditures on health status. They thus act upon the same (mistaken) preferences in the two periods of their lives. Up to their last days, these people stay ignorant of (or unwilling to act upon) these effects.

Which preferences should the social planner use when assessing optimal taxes/subsidies on sin good consumption, saving and health care expenditures? The recent literature on paternalism has studied the impact of the introduction of behavioral considerations on the objective of the planner. Thaler and Sunstein (2003) make a strong case for "libertarian paternalism", which applies to settings in which no coercion is involved, such as when the planner has to choose the default option (such as the automatic enrollment in 401(k) employee savings plans in the US). Other papers go further and study situations where the planner's decision coerces people, such as when it taxes certain goods or even prevents them from being consumed. The literature has focused upon the case where people differ in their degree of non-rationality. These contributions advocate the use by the planner of "cautious" (O'Donoghue and Rabin (1999)) or "asymmetric" (Camerer et al. (2003)) paternalism, which trades-off the benefits of paternalistic interventions for people making mistakes against the costs for fully rational individuals. This literature usually shows that some intervention is called for by this kind of paternalism, since deviations from laissez-faire impose second-order costs on rational individuals, but first-order gains for non-rational people. Moreover, they show (O'Donoghue and Rabin (2003, 2006)) that even a small probability (or proportion) of people making mistakes can have dramatic effects for optimal policy.

We add to this literature by departing from it in two ways. First, we study the consequences of adopting a paternalistic objective when individuals are adamant in the mistakes they make – i.e., when they never realize (because for instance of ignorance or cognitive dissonance) that they base their decisions on wrong premises. Second, rather than mixing rational and non-rational individuals, we contrast the results obtained when all individuals are repentant in the second period or when none is (in the two cases, we allow for heterogeneity in the degree of myopia, but not for whether myopia persists or not in the second stage). Obviously, paternalism is easier to defend in our setting when individuals have dual selves. We nevertheless think that a case can be made in favor of paternalism even when individuals never realize their mistake and thus always use preferences different from those used by the planner. Our approach in that case is reminiscent of the older literature on "merit goods" (Musgrave (1959), more recently Besley (1988)), where we add that the reason for the difference between the planner's and the individuals' preferences resides in the (unrecognized) mistakes made by individuals. We consider it interesting to study and contrast the results obtained in the two scenarios even if the argument for paternalism is less convincing in the latter one.[2]

We obtain the following results. We first show that the first best outcome can be decentralized with individualized linear taxes and subsidies in the two scenarios (persistent error and dual self). In the first one, it is necessary to tax the sin good consumption while subsidizing health care expenditures. There is no need to influence saving. The second scenario is more complex, since the social planner faces a problem with changing preferences. The planner has to intervene in the first period by taxing the sin good while subsidizing savings. There is no need to influence health care expenditures, which are optimally chosen provided that first period choices are optimal. Comparing the sin

[2] To be complete, in the conclusion we discuss what would happen if we adopt a welfarist approach in the case of persistent myopia/ignorance.

tax in the two scenarios, we obtain that it is smaller in the dual self case if and only if the marginal effect of health care on health status increases with sin good consumption.

We then turn to the second best setting where the planner observes neither income, preferences nor savings. In the single self case, optimal linear sin taxes and health expenditure subsidies depend upon two terms: a (classical) covariance term reflecting distributive considerations and a "Pigouvian" term that reflects the "internalities" an individual imposes on himself. In the dual self case, optimal tax formulas contain a third term, which is linked to the inability to control savings. This additional term calls for higher tax on sin good/subsidies on health care provided that this tax/subsidy encourages savings.

The rest of the paper is organized as follows. In section 2, the model, the first-best solutions and the decentralization conditions are presented for the two specifications. Then in section 3 we turn to the second-best problem when individuals persist in their ignorance. In section 4, we study the alternative second-best problem, that is when individuals realize having made a mistake. A final section concludes.

2 First-best and decentralization

2.1 Model

We consider a society consisting of a number of types of individuals i. Each type is characterized by a wealth endowment w_i and a subjective and objective health parameter α_i and β_i. Each individual's life spans over two periods. In the first one, he consumes a numeraire good c_i and a sin good x_i. He also saves s_i for future expenses. In the second period, he consumes an amount d_i of the numeraire and he invests e_i in health improvement. In this second period, he enjoys a quality of health $\beta_i h(x_i, e_i)$, on which x_i has a negative effect and e_i a positive effect. For reason of ignorance or myopia, the individual has a perception of this function that underestimates the impact of both arguments. In other words, he perceives a health function equal to $\alpha_i h(x_i, e_i)$ with $\alpha_i < \beta_i$.[3]

[3] An alternative specification could be $h(\beta_i x_i, e_i)$ or $h(\alpha_i x_i, e_i)$, in which case myopia only concerns the sin good and not health care. This would not change the qualitative nature of our results.

His two-period utility function can be written as:

$$U_i = u(c_i) + \varphi(x_i) + u(d_i) + \alpha_i h(x_i, e_i), \quad (1)$$

with budget constraints:

$$\begin{aligned} w_i &= (1+\tau) s_i + (1+\theta) x_i + c_i - a, \\ d_i &= s_i - (1+\sigma) e_i, \end{aligned}$$

where τ, θ, σ are tax rates and a is a demogrant. For simplicity reasons, we assume a zero time discount rate and a zero rate of interest.

2.2 First-best

We assume that the government is a paternalistic utilitarian. In other words, it adopts an objective made of the sum of utilities (1) in which β_i replaces α_i. As a benchmark, we derive the first-best (FB) conditions by solving the following Lagrangian:

$$\begin{aligned} \mathcal{L}_1 &= \sum n_i \left[u(c_i) + \varphi(x_i) + u(d_i) + \beta_i h(x_i, e_i) \right. \\ &\quad \left. -\mu(c_i + x_i + d_i + e_i - w_i) \right], \end{aligned}$$

where μ is the Lagrangian multiplier associated with the resource constraints and n_i the relative number of type i's individuals. The FOCs yield:

$$u'(c_i) = u'(d_i) = \varphi'(x_i) + \beta_i h_x(x_i, e_i) = \beta_i h_e(x_i, e_i) = \mu,$$

with $h_x < 0$ and $h_e > 0$. Denote the first best solution by c_i^*, x_i^*, d_i^* and e_i^*. We can also define $s_i^* = d_i^* + e_i^*$, the (implicit) individual savings at the first-best solution.

The utilitarian planner equalizes marginal utility of consuming the numeraire good in both periods. Since preferences for this good are the same for all individuals, this calls for c_i^* and d_i^* to be equal and the same for all. Marginal utility for numeraire and sin goods are also equalized, with the latter composed of the immediate marginal gratification and of the (true) delayed marginal impact on health. Finally, the planner also equalizes second period marginal utility from consuming the sin good and from consuming health care. If the marginal impact of sin good consumption on health is the same for all individuals ($\beta_i = \beta$), then x_i^* and e_i^* are also identical for all.

It is interesting to contrast these conditions with the *laissez-faire* (LF) ones, which are obtained by maximizing:

$$\begin{aligned} U_i &= u\left(w_i - (1+\tau)\, s_i - (1+\theta)\, x_i + a_i\right) \\ &\quad + \varphi\left(x_i\right) + \alpha_i h\left(x_i, e_i\right) + u\left(s_i - (1+\sigma)\, e_i\right). \end{aligned}$$

In the LF, $\tau = \sigma = \theta = a_i = 0$ and we have:

$$u'\left(c_i\right) = u'\left(d_i\right) = \varphi'\left(x_i\right) + \alpha_i h_x\left(x_i, e_i\right) = \alpha_i h_e\left(x_i, e_i\right).$$

Marginal utility of consuming the numeraire and the sin goods are also equalized (though they differ across agents if there is heterogeneity in α_i), but not at the correct level since individuals make a mistake when assessing the impact of both sin good and health care consumption on their second period utility (health status).

2.3 Decentralization with persisting errors

To decentralize the above optimum, we need individualized redistributive lump sum taxes a_i and individualized corrective taxes or subsidies on the sin good and health expenditure.

$$\theta_i = \frac{\left(\alpha_i - \beta_i\right) h_x\left(x_i^*, e_i^*\right)}{\beta_i h_e\left(x_i^*, e_i^*\right)} > 0, \tag{2}$$

and

$$\sigma_i = \frac{\left(\alpha_i - \beta_i\right)}{\beta_i} < 0. \tag{3}$$

The tax on sin good consumption forces the individual to internalize the full impact of his sin good consumption on his health. It is proportional to the share of the impact that he does not spontaneously internalize (given by the difference between β and α), to the marginal impact of sin good on health, and decreases with the marginal impact of health care on health status. It is also necessary to subsidize health care, since individuals underestimate its impact on health. Intuitively, the subsidy rate is equal to the percentage of underestimation by the individual $(\beta - \alpha)/\beta$. There is no need to influence saving, since individuals do not exhibit time-inconsistent preferences.

Those taxes and subsidies are individualized. Naturally, with $\alpha_i = \alpha$ and $\beta_i = \beta$, they would be identical for all.

2.4 Decentralization with dual self

Up to now we have assumed that individuals stick to their beliefs in the second period. Let us now make the reasonable assumption that in the second period they realize that they have made a mistake out of ignorance or myopia and will accordingly modify their decision concerning health care. In behavioral economics, one then speaks of dual self.

When the "reasonable" self prevails in the second period, the choice of e_i is determined by the equality

$$(1+\sigma)u'\left(s_i - (1+\sigma)e_i\right) = \beta_i h_e\left(x_i, e_i\right). \tag{4}$$

On the other hand, this level of e is not the one that the individual envisioned when he chose his sin good consumption and saving in the first period. Rather, the amount of health care that the individual planned to buy later on, denoted by e_i^P, is given by

$$(1+\sigma)u'(s_i - (1+\sigma)e_i^P) = \alpha_i h_e(x_i, e_i^P). \tag{5}$$

The choice of x_i and s_i then satisfies the following first-order conditions:

$$\begin{aligned} &-(1+\tau_i)u'(c_i) + u'(s_i - e_i^P) = 0, \\ &-(1+\theta_i)u'(c_i) + \varphi'\left(x_i\right) + \alpha_i h_x\left(x_i, e_i^P\right) = 0. \end{aligned}$$

Is it possible to decentralize the first-best optimum in these conditions with our linear instruments that are chosen in the first period? In fact, this is possible using τ_i and θ_i plus a_i. With these instruments, and denoting optimal values with a *, one obtains x_i^* and s_i^*, which then imply e_i^*. Defining e_i^{P*} as the planned level of e_i when the tax instruments are set to decentralize the first-best, we obtain that

$$\tau_i = \frac{u'(s_i^* - e_i^{P*}) - u'(c_i^*)}{u'(c_i^*)}, \tag{6}$$

$$\theta_i = \frac{\varphi'\left(x_i^*\right) + \alpha_i h_x\left(x_i^*, e_i^{P*}\right) - u'(c_i^*)}{u'(c_i^*)} = \frac{\alpha_i h_x\left(x_i^*, e_i^{P*}\right) - \beta_i h_x\left(x_i^*, e_i^*\right)}{u'(c_i^*)}. \tag{7}$$

Equation (4) shows that the individual will take the optimal health care decision in the second stage, provided that he chose the optimal values of x and s in the first stage. Influencing health care decision is then unnecessary, provided that tax instruments on saving and sin good consumption decentralize these two optimal choices. The sin tax is proportional to the mistake made by the individual. This mistake comes from two sources: under-estimation of the impact of sin good on health (since $\alpha < \beta$) and misplanning of the future amount of health care consumed (e_i^P v e_i). Since individuals misplan their future health care need, it is also necessary to influence their saving decision, as shown by (6). Note that assuming $\alpha_i = \alpha$ and $\beta_i = \beta$, we obtain $\tau_i = \tau$ and $\theta_i = \theta$. In words, a_i makes everyone identical and the Pigouvian tax and subsidy rates are identical.

To illustrate this point, assume a single individual with $\alpha = 0 < \beta = 1$. We then have $e_i^{P*} = 0$ so that the implementing tax rates or subsidy are

$$\tau = \frac{u'(s^*)}{u'(c^*)} - 1 < 0$$
$$\theta = \frac{-h_x(x^*, e^*)}{u'(c^*)} > 0.$$

2.5 Comparison of sin taxes in the two specifications

It is interesting to compare the sin taxes obtained under the two specifications. To make the comparison easier, we assume that $\alpha_i = \alpha > 0$ and $\beta_i = \beta$. We thus have (with S for single self and D for dual self):

$$\theta^S = \frac{(\alpha - \beta) h_x(x^*, e^*)}{u'(c^*)}$$

and

$$\theta^D = \frac{(\alpha - \beta) h_x(x^*, e^*) + \alpha \left[h_x\left(x^*, e^{P*}\right) - h_x(x^*, e^*)\right]}{u'(c^*)}. \quad (8)$$

In the two cases, the sin tax is proportional to the error made in the first stage when evaluating the damage of sin good consumption on health, measured at the optimal sin good and health care consumptions. An additional term is present in the dual self case, which is proportional to the second mistake made by the individual in that case. Since this individual misestimates how much health care he will buy at the optimum, he is also mistaken in his assessment of the marginal damage done by the optimal amount of sin good consumption, as measured by the function $h(x, e)$. The sign of this impact depends on the cross-derivative of this function. Assume for instance that it is positive. Since the individual under-estimates how much health care he will buy, he then over-estimates how bad the marginal impact of sin good will be (as measured by the function $h(.)$). This calls for decreasing the tax on sin good, compared to a "single self" individual. Assuming that h_{xe} has everywhere the same sign, we then obtain that

$$\theta^S \gtreqless \theta^D \quad \text{if and only if} \quad h_{xe} \gtreqless 0.$$

The sign of this cross derivative depends upon the kind of sin good we are considering. For instance, it seems reasonable to assume that it is positive if the sin good is sugar: the more you eat, the more medications designed to treat diabetes may be helpful to you. With this assumption the sin tax is smaller when the individual acknowledges his mistake in the second period of his life. The opposite assumption can be make for smoking: heavy smokers increase their probability of getting lung cancer, for which there is up to now no efficient cure in the majority of cases. Put bluntly, there is not much utility that you can get from consuming health care if you end up with lung cancer following heavy smoking...

3 Second-best in the case of persistent errors

We now turn to the second-best setting with linear tax instruments and uniform demogrant. We assume in the remaining of this paper that $\beta_i = \beta > \alpha_i$. In other words, the objective effect of both e and x on health is the same for all, but individuals vary in their degree of myopia (as well as in income). We also assume that taxes/subsidies on saving are not available anymore (either because saving is not observable, or because elements not modelled, like international mobility of capital, prevent saving from being taxed or subsidized).

We first consider the case when the individuals never acknowledge that the true health parameter is β. In that case, restricting the instruments to linear taxes and uniform demogrant we write the new Lagrangian as:

$$\begin{aligned}\mathcal{L}_2 &= \sum n_i \left[u\left(w_i - s_i(1+\tau) - x_i\left(1+\theta\right) + a\right) + \varphi\left(x_i\right) + u\left(s_i - \left(1+\sigma\right)e_i\right) \right. \\ &\quad \left. + \beta h\left(x_i, e_i\right) - \mu\left(a - \tau s_i - \theta x_i - \sigma e_i\right)\right],\end{aligned}$$

where s_i, x_i and e_i are functions of a, θ, τ and σ and are obtained from the following optimal conditions for individual choices:

$$-u'\left(c\right)\left(1+\tau\right) + u'\left(d\right) = 0 \tag{9}$$

$$-u'\left(c\right)\left(1+\theta\right) + \varphi'\left(x\right) + \alpha h_x\left(x, e\right) = 0 \tag{10}$$

$$-u'\left(d\right)\left(1+\sigma\right) + \alpha h_e\left(x, e\right) = 0. \tag{11}$$

Assuming interior solutions, the FOCs of the social problem are given by:

$$\begin{aligned}\frac{\partial \mathcal{L}_2}{\partial a} =& Eu'\left(c\right) + E\left[h_x\left(x,e\right)\frac{\partial x}{\partial a} + h_e\left(x,e\right)\frac{\partial e}{\partial a}\right]\left(\beta-\alpha\right) \\ & -\mu E\left[1 - \tau\frac{\partial s}{\partial a} - \theta\frac{\partial x}{\partial a} - \sigma\frac{\partial e}{\partial a}\right] = 0, \\ \frac{\partial \mathcal{L}_2}{\partial \tau} =& -Eu'\left(c\right)s + E\left[h_x\left(x,e\right)\frac{\partial x}{\partial \tau} + h_e\left(x,e\right)\frac{\partial e}{\partial \tau}\right]\left(\beta-\alpha\right) \\ & +\mu E\left[s + \tau\frac{\partial s}{\partial \tau} + \theta\frac{\partial x}{\partial \tau} + \sigma\frac{\partial e}{\partial \tau}\right] = 0, \\ \frac{\partial \mathcal{L}_2}{\partial \theta} =& -Eu'\left(c\right)x + E\left[h_x\left(x,e\right)\frac{\partial x}{\partial \theta} + h_e\left(x,e\right)\frac{\partial e}{\partial \theta}\right]\left(\beta-\alpha\right) \\ & +\mu E\left[x + \tau\frac{\partial s}{\partial \theta} + \theta\frac{\partial x}{\partial \theta} + \sigma\frac{\partial e}{\partial \theta}\right] = 0, \\ \frac{\partial \mathcal{L}_2}{\partial \sigma} =& -Eu'\left(d\right)e + E\left[h_x\left(x,e\right)\frac{\partial x}{\partial \sigma} + h_e\left(x,e\right)\frac{\partial e}{\partial \sigma}\right]\left(\beta-\alpha\right) \\ & +\mu E\left[e + \tau\frac{\partial s}{\partial \sigma} + \theta\frac{\partial x}{\partial \sigma} + \sigma\frac{\partial e}{\partial \sigma}\right] = 0.\end{aligned}$$

In these expressions, we have used the operator E for $\sum n_i$.

In compensated terms, these expressions can be written as:[4]

$$\frac{\partial \tilde{\mathcal{L}}_2}{\partial \tau} = -Eu'(c)\,s + Eu'(c)\,Es + E\left[h_x(x,e)\frac{\partial \tilde{x}}{\partial \tau} + h_e(x,e)\frac{\partial \tilde{e}}{\partial \tau}\right](\beta - \alpha)$$
$$+\mu E\left[\tau\frac{\partial \tilde{s}}{\partial \tau} + \theta\frac{\partial \tilde{x}}{\partial \tau} + \sigma\frac{\partial \tilde{e}}{\partial x}\right] = 0,$$
$$\frac{\partial \tilde{\mathcal{L}}_2}{\partial \theta} = -Eu'(c)\,x + Eu'(c)\,Ex + E\left[h_x(x,e)\frac{\partial \tilde{x}}{\partial \theta} + h_e(x,e)\frac{\partial \tilde{e}}{\partial \theta}\right](\beta - \alpha)$$
$$+\mu E\left[\tau\frac{\partial \tilde{s}}{\partial \theta} + \theta\frac{\partial \tilde{x}}{\partial \theta} + \sigma\frac{\partial \tilde{e}}{\partial \theta}\right] = 0,$$
$$\frac{\partial \tilde{\mathcal{L}}_2}{\partial \sigma} = -Eu'(c)\,e + Eu'(c)\,Ee + E\left[h_x(x,e)\frac{\partial \tilde{x}}{\partial \sigma} + h_e(x,e)\frac{\partial \tilde{e}}{\partial \sigma}\right](\beta - \alpha)$$
$$-\mu E\left[\tau\frac{\partial \tilde{s}}{\partial \sigma} + \theta\frac{\partial \tilde{x}}{\partial \sigma} + \sigma\frac{\partial \tilde{e}}{\partial \sigma}\right] = 0.$$

It is important to note that we here use the concept of average compensation and not that of the standard Slutsky term. Using a tilde for the former, a hat for the latter, we have:

$$\frac{\partial \tilde{x}}{\partial \sigma} = \frac{\partial x}{\partial \sigma} + \frac{\partial x}{\partial a}Ee = \frac{\partial \hat{x}}{\partial \sigma} + \frac{\partial x}{\partial a}(Ee - e)$$

and similarly,

$$\frac{\partial \tilde{x}}{\partial \theta} = \frac{\partial \hat{x}}{\partial \theta} + \frac{\partial x}{\partial a}(Ex - x)\,.$$

In Appendix A we provide the FOC's in term of the standard Slutsky effects. Our approach is simple, but the signs of the compensated terms have to be interpreted with caution. For example, we know that $\partial \hat{x}/\partial \theta < 0$, but if x is much smaller than the average, Ex, $\partial \tilde{x}/\partial \theta$ could be positive.

In interpreting the above FOCs, we assume that these compensated derivative are negative. We also observe that with either identical individuals or individualized lump sum transfers a_i, the first-best optimum is obtained with just σ and θ. As mentioned

[4] Defining

$$\frac{\partial \tilde{\mathcal{L}}_2}{\partial \tau} = \frac{\partial \mathcal{L}_2}{\partial \tau} + \frac{\partial \mathcal{L}_2}{\partial a}Es,$$
$$\frac{\partial \tilde{\mathcal{L}}_2}{\partial \theta} = \frac{\partial \mathcal{L}_2}{\partial \theta} + \frac{\partial \mathcal{L}_2}{\partial a}Ex,$$
$$\frac{\partial \tilde{\mathcal{L}}_2}{\partial \sigma} = \frac{\partial \mathcal{L}_2}{\partial \sigma} + \frac{\partial \mathcal{L}_2}{\partial a}Ee.$$

at the beginning of the section, we assume that taxation/subsidization of saving is not available to the planner, so that $\tau = 0$. With this assumption, we have

$$\mu\theta = \frac{\operatorname{cov}\left(u'\left(c\right),x\right)E\dfrac{\partial\tilde{e}}{\partial\sigma} - \operatorname{cov}\left(u'\left(c\right),e\right)E\dfrac{\partial\tilde{e}}{\partial\theta} - E\left(\beta-\alpha\right)H_{\theta}E\dfrac{\partial\tilde{e}}{\partial\sigma} + E\left(\beta-\alpha\right)H_{\sigma}E\dfrac{\partial\tilde{e}}{\partial\theta}}{E\dfrac{\partial\tilde{x}}{\partial\theta}E\dfrac{\partial\tilde{e}}{\partial\sigma} - E\dfrac{\partial\tilde{e}}{\partial\theta}E\dfrac{\partial\tilde{x}}{\partial\sigma}},$$

$$\mu\sigma = \frac{\operatorname{cov}\left(u'\left(c\right),e\right)E\dfrac{\partial\tilde{x}}{\partial\theta} - \operatorname{cov}\left(u'\left(c\right),x\right)E\dfrac{\partial\tilde{x}}{\partial\sigma} - E\left(\beta-\alpha\right)H_{\sigma}E\dfrac{\partial\tilde{x}}{\partial\theta} + E\left(\beta-\alpha\right)H_{\theta}E\dfrac{\partial\tilde{x}}{\partial\sigma}}{E\dfrac{\partial\tilde{x}}{\partial\theta}E\dfrac{\partial\tilde{e}}{\partial\sigma} - E\dfrac{\partial\tilde{e}}{\partial\theta}E\dfrac{\partial\tilde{x}}{\partial\sigma}},$$

where

$$H_{\theta} = h_x\frac{\partial\tilde{x}}{\partial\theta} + h_e\frac{\partial\tilde{e}}{\partial\theta} \qquad \text{and} \qquad H_{\sigma} = h_x\frac{\partial\tilde{x}}{\partial\sigma} + h_e\frac{\partial\tilde{e}}{\partial\sigma}.$$

If we assume that the cross derivatives are negligible, namely that $\partial\tilde{x}/\partial\sigma \rightarrow 0$ and $\partial\tilde{e}/\partial\theta \rightarrow 0$, we obtain:

$$\theta = \frac{E\left(\beta-\alpha\right)h_x\left(x,e\right)\dfrac{\partial\tilde{x}}{\partial\theta} - \operatorname{cov}\left(u'\left(c\right),x\right)}{-\mu E\dfrac{\partial\tilde{x}}{\partial\theta}}, \tag{12}$$

$$\sigma = \frac{E\left(\beta-\alpha\right)h_e\left(x,e\right)\dfrac{\partial\tilde{e}}{\partial\sigma} - \operatorname{cov}\left(u'\left(c\right),e\right)}{-\mu E\dfrac{\partial\tilde{e}}{\partial\sigma}}, \tag{13}$$

where $\mu > 0$.

The first term of the numerator of (12) and (13) is the Pigouvian term found in (2) and (3) summed over all individuals with weights equal to the effect of the tax on individual demands of either x or e. With both derivatives of demand functions negative, this term calls for a tax on sin good and a subsidy on health care. The second term of the numerator of (12) and (13) reflects redistributive considerations. It depends on the concavity of u, the initial inequality of earnings and the correlation between α_i and w_i. With identical individuals, this term disappears and (12) and (13) reduce to (2) and (3). With different individuals and no correlation between α_i and w_i, the covariance will be negative in both equations, since richer people consume more of all goods (x, c and e) than poorer people with the same degree of myopia. This tends to increase the tax on sin goods and decrease the subsidy on health care, compared to the case with identical individuals. With a positive correlation, $\operatorname{cov}(u'\left(c\right),x)$ will tend to increase (since less myopic people, those with a larger α, consume less sin good) which leads to a smaller sin tax than with zero correlation. The impact of a positive correlation on $\text{cov}(u'(c), e)$ is less clear: smarter people buy less sin good, but they also better realize the importance of health care, so that the net impact on the amount of e consumed is not easy to determine.

4 Second-best with dual self

Now we assume that the individuals realize after one period that they made a mistake and that the only corrective decision they can make is the choice of health expenditure. As in section 2.4, we thus distinguish between the planned investment $e^P$and the *ex post* choice e. The indirect utility function used by the social planner in its welfare maximization has to take into account these two values of e which yield two values of d.

In the first period, the functions $x(\tau,\theta,\sigma,a)$, $s(\tau,\theta,\sigma,a)$ and $e^P(\tau,\theta,\sigma,a)$ are obtained as the solution to:

$$-u'(c)(1+\tau)+u'\left(d^P\right)=0, \tag{14}$$

$$-u'(c)(1+\theta)+\varphi'(x)+\alpha h_x\left(x,e^P\right)=0, \tag{15}$$

$$-u'\left(d^P\right)(1+\sigma)+\alpha h_e\left(x,e^P\right)=0. \tag{16}$$

where $d^P=s-e^P(1+\sigma)>d=s-e(1+\sigma)$.

In the second period we have the effective demand for e defined by (4)

$$e=f\left(s,x,e^P\right)=e(\tau,\theta,\sigma,a). \tag{17}$$

The Lagrangian is given by:

$$\begin{aligned}\mathcal{L}_3 \;=\; & \sum n_i\left[u\left(w_i-s_i(1+\tau)-x_i(1+\theta)+a\right)+\varphi(x_i)+u\left(s_i-(1+\sigma)e_i\right)\right.\\ & \left.+\beta h(x_i,e_i)-\mu\left(a-\tau s_i-\theta x_i-\sigma e_i\right)\right],\end{aligned}$$

which is similar to $\mathcal{L}_2$ except that individual choices are now determined by (14), (15) and (17). The FOCs are given by

$$\begin{aligned}
\frac{\partial \mathcal{L}_3}{\partial a} =& Eu'(c) + E\left\{u'[s-e(1+\sigma)] - u'[s-e^P(1+\sigma)]\right\}\frac{\partial s}{\partial a} \\
&+ E\left[\beta h_x(x,e) - \alpha h_x\left(x,e^P\right)\right]\frac{\partial x}{\partial a} - \mu E\left[1 - \tau\frac{\partial s}{\partial a} - \theta\frac{\partial x}{\partial a} - \sigma\frac{\partial e}{\partial a}\right] = 0, \\
\frac{\partial \mathcal{L}_3}{\partial \tau} =& -Eu'(c)s + E\left\{u'[s-e(1+\sigma)] - u'[s-e^P(1+\sigma)]\right\}\frac{\partial s}{\partial \tau} \\
&+ E\left[\beta h_x(x,e) - \alpha h_x\left(x,e^P\right)\right]\frac{\partial x}{\partial \tau} + \mu E\left[s + \tau\frac{\partial s}{\partial \tau} + \theta\frac{\partial x}{\partial \tau} + \sigma\frac{\partial e}{\partial \tau}\right] = 0, \\
\frac{\partial \mathcal{L}_3}{\partial \theta} =& -Eu'(c)x + E\left\{u'[s-e(1+\sigma)] - u'[s-e^P(1+\sigma)]\right\}\frac{\partial s}{\partial \theta} \\
&+ E\left[\beta h_x(x,e) - \alpha h_x\left(x,e^P\right)\right]\frac{\partial x}{\partial \theta} + \mu E\left[x + \tau\frac{\partial s}{\partial \theta} + \theta\frac{\partial x}{\partial \theta} + \sigma\frac{\partial e}{\partial \theta}\right] = 0, \\
\frac{\partial \mathcal{L}_3}{\partial \sigma} =& -Eu'(s-e(1+\sigma))e + E\left\{u'[s-e(1+\sigma)] - u'[s-e^P(1+\sigma)]\right\}\frac{\partial s}{\partial \sigma} \\
&+ E\left[\beta h_x(x,e) - \alpha h_x\left(x,e^P\right)\right]\frac{\partial x}{\partial \sigma} + \mu E\left[e + \tau\frac{\partial s}{\partial \sigma} + \theta\frac{\partial x}{\partial \sigma} + \sigma\frac{\partial e}{\partial \sigma}\right] = 0.
\end{aligned}$$

As in the previous section, we assume $\tau = 0$ and use $\partial\mathcal{L}_3/\partial a$ to obtain the compensated expressions of $\partial\mathcal{L}_3/\partial\theta$ and $\partial\mathcal{L}_3/\partial\sigma$.

$$\begin{aligned}
\frac{\partial \tilde{\mathcal{L}}_3}{\partial \theta} =& -\operatorname{cov}\left(u'(c),x\right) + E\left(u'(d) - u'\left(d^P\right)\right)\frac{\partial \tilde{s}}{\partial \theta} \\
&+ E\left(\beta h_x(x,e) - \alpha h_x\left(x,e^P\right)\right)\frac{\partial \tilde{x}}{\partial \theta} + \mu\left(\theta E\frac{\partial \tilde{x}}{\partial \theta} + \sigma E\frac{\partial \tilde{e}}{\partial \theta}\right) = 0, \\
\frac{\partial \tilde{\mathcal{L}}_3}{\partial \sigma} =& -\operatorname{cov}\left(u'(c),e\right) + E\left(u'(d) - u'\left(d^P\right)\right)\frac{\partial \tilde{s}}{\partial \sigma} \\
&+ E\left(\beta h_x(x,e) - \alpha h_x\left(x,e^P\right)\right)\frac{\partial \tilde{x}}{\partial \sigma} + \mu\left(\theta E\frac{\partial \tilde{x}}{\partial \sigma} + \sigma E\frac{\partial \tilde{e}}{\partial \sigma}\right) = 0.
\end{aligned}$$

It is clear from the above that even with identical individuals, one cannot achieve the first-best with θ and σ as instruments. Solving for θ and σ, we obtain

$$\begin{aligned}
\Delta\mu\theta =& \operatorname{cov}\left(u'(c),x\right)E\frac{\partial \tilde{e}}{\partial \sigma} - \operatorname{cov}\left(u'(c),e\right)E\frac{\partial \tilde{e}}{\partial \theta} - E\tilde{H}\frac{\partial \tilde{x}}{\partial \theta}E\frac{\partial \tilde{e}}{\partial \sigma} + E\tilde{H}\frac{\partial \tilde{x}}{\partial \sigma}E\frac{\partial \tilde{e}}{\partial \theta} \\
&- E\left(u'(d) - u'\left(d^P\right)\right)\frac{\partial \tilde{s}}{\partial \theta}E\frac{\partial \tilde{e}}{\partial \sigma} + E\left(u'(d) - u'\left(d^P\right)\right)\frac{\partial \tilde{s}}{\partial \sigma}E\frac{\partial \tilde{e}}{\partial \theta}, \\
\Delta\mu\sigma =& \operatorname{cov}\left(u'(c),e\right)E\frac{\partial \tilde{x}}{\partial \theta} - \operatorname{cov}\left(u'(c),x\right)E\frac{\partial \tilde{x}}{\partial \sigma} - E\tilde{H}\frac{\partial \tilde{x}}{\partial \sigma}E\frac{\partial \tilde{x}}{\partial \theta} + E\tilde{H}\frac{\partial \tilde{x}}{\partial \theta}E\frac{\partial \tilde{x}}{\partial \sigma} \\
&- E\left(u'(d) - u'\left(d^P\right)\right)\frac{\partial \tilde{s}}{\partial \sigma}E\frac{\partial \tilde{x}}{\partial \theta} + E\left(u'(d) - u'\left(d^P\right)\right)\frac{\partial \tilde{s}}{\partial \theta}E\frac{\partial \tilde{x}}{\partial \sigma},
\end{aligned}$$

where

$$\tilde{H} = \beta h_x(x,e) - \alpha h_x\left(x,e^P\right) \qquad \text{and} \qquad \Delta = E\frac{\partial \tilde{x}}{\partial \theta}E\frac{\partial \tilde{e}}{\partial \sigma} - E\frac{\partial \tilde{e}}{\partial \theta}E\frac{\partial \tilde{x}}{\partial \sigma}.$$

These can be rewritten as follows if we assume that the cross price effects are negligible:

$$\theta = \frac{-\operatorname{cov}\left(u'\left(c\right), x\right) + E\left(u'\left(d\right) - u'\left(d^P\right)\right)\dfrac{\partial \tilde{s}}{\partial \theta} + E\left[\beta h_x\left(x, e\right) - \alpha h_x\left(x, e^P\right)\right]\dfrac{\partial \tilde{x}}{\partial \theta}}{-\mu E \dfrac{\partial \tilde{x}}{\partial \theta}}, \quad (18)$$

$$\sigma = \frac{-\operatorname{cov}\left(u'\left(c\right), e\right) + E\left(u'\left(d\right) - u'\left(d^P\right)\right)\dfrac{\partial \tilde{s}}{\partial \sigma}}{-\mu E \dfrac{\partial \tilde{e}}{\partial \sigma}}. \quad (19)$$

The covariance terms in both equations are identical to those in the single self scenario and reflect the equity concern of public policy. The third term of the numerator of (18) is the Pigouvian term found in equation (7) summed over all individuals with weights equal to the effect of the tax on individual demands of x. With $\alpha = 0$ or $h_{xe} < 0$, it is definitively positive (since $\partial \tilde{x}/\partial \theta < 0$). The second term of the numerator of both equations has the sign of either $\partial \tilde{s}/\partial \sigma$ or $\partial \tilde{s}/\partial \theta$. The intuition for this term goes as follows: individuals over-estimate their second period consumption ($d^P > d$) since they under-estimate their health care needs. As a consequence, they do not save enough. If either tax has a positive effect on saving, then it should have a relatively higher value.

In the case of identical individuals, (18) and (19) can be rewritten as:

$$\mu\theta = \frac{\left[u'\left(d^P\right) - u'\left(d\right)\right]\dfrac{\partial \tilde{s}}{\partial \theta}}{\dfrac{\partial \tilde{x}}{\partial \theta}} + \left(\alpha - \beta\right) h_x\left(x, e\right) + \alpha\left[h_x\left(x, e^P\right) - h_x(x, e)\right],$$

$$\mu\sigma = \left[u'\left(d^P\right) - u'\left(d\right)\right]\frac{\partial \tilde{s}}{\partial \sigma}\frac{\partial \tilde{e}}{\partial \sigma}.$$

The second part of the sin tax is the familiar Pigouvian term, expressed as in the first-best decentralization equation (8). The other term in both equations comes from our inability to control saving directly, which would be necessary to decentralize the first-best optimum. Saving can be indirectly controlled through the use of both θ and σ. If any of these instruments stimulate saving, this makes using it more desirable.

5 Conclusion

In this paper we have considered the case of sin goods that have delayed negative effects that individuals ignore at the time of consumption but acknowledge later. Individuals have then the possibility of partially compensating those negative effects by investing in health care. Assuming a paternalistic government, we show that the first-best could be decentralized with a sin tax, a subsidy on saving and individualized lump sum transfers (or alternatively, by assuming identical individuals). In the second-best, individualized lump sum transfers are not available and the only available instruments are a linear sin tax and a linear subsidy on health care. We discuss the optimal second-best tax subsidy policy wherein distributive and corrective Pigouvian considerations are mixed.

We also consider the case of what we call persistent error, namely the case where individuals acknowledge the negative effects of their sinful consumption when it is too late to take any corrective action (i.e., after e_i has been chosen). From an individuals perspective this latter case is formally equivalent to yet another setting which corresponds to what can be called persisting ignorance, where the individual never acknowledges the negative effects of his consumption. However, under this interpretation, the use of a paternalistic approach become questionable; we return to the traditional, "old" paternalism. One could thus argue that the right approach is welfarist and not paternalist in that setting. We present the welfarist solution to the case of persisting ignorance in Appendix B. Comparing this to the results obtain in Section 2.3., we show that with paternalism the planner wants the good of people against their own will not only in the first, but also in the second period of their life. With welfarism, we add different utilities, which is also questionable.

In this paper we have focused on sinful consumption. Our method could be used for other problems. For example, lack of physical exercises or hygiene in the first period of life which has delayed detrimental effects. These effects can be partially offset in the second period. Another example is overtime or moonlighting that lead to early disability. A fully rational individual would understand the importance of not abusing one's body when young to avoid regretful consequences later on in lifetime. The ingredients of these various situations are: behavior with delayed detrimental effects, myopia and possibility of partial compensation.

Appendix

A Expressions with standard Slutsky terms

We only consider the case of persistent errors:

$$\begin{aligned}\mathcal{L}_2 &= \sum n_i \left[u\left(w_i - s_i\left(1+\tau\right) - x_i\left(1+\theta\right) + a\right) + \varphi\left(x_i\right) + u\left(s_i - \left(1+\sigma\right)e_i\right)\right. \\ &\quad \left. -\mu\left(a - \tau s_i - \theta x_i - \sigma e_i\right)\right].\end{aligned}$$

The FOCs are given by

$$\frac{\partial \mathcal{L}_2}{\partial a} = Eu'\left(c\right) + E\left[h_x\left(x,e\right)\frac{\partial x}{\partial a} + h_e\left(x,e\right)\frac{\partial e}{\partial a}\right]\left(\beta-\alpha\right) - \mu E\left[1 - \tau\frac{\partial s}{\partial a} - \theta\frac{\partial x}{\partial a} - \sigma\frac{\partial e}{\partial a}\right] = 0,$$

$$\frac{\partial \mathcal{L}_2}{\partial \tau} = -Eu'\left(c\right)s + E\left[h_x\left(x,e\right)\frac{\partial x}{\partial \tau} + h_e\left(x,e\right)\frac{\partial e}{\partial \tau}\right]\left(\beta-\alpha\right) - \mu E\left[-s - \tau\frac{\partial s}{\partial \tau} - \theta\frac{\partial x}{\partial \tau} - \sigma\frac{\partial e}{\partial \tau}\right] = 0,$$

$$\frac{\partial \mathcal{L}_2}{\partial \theta} = -Eu'\left(c\right)x + E\left[h_x\left(x,e\right)\frac{\partial x}{\partial \theta} + h_e\left(x,e\right)\frac{\partial e}{\partial \theta}\right]\left(\beta-\alpha\right) - \mu E\left[-x - \tau\frac{\partial s}{\partial \theta} - \theta\frac{\partial x}{\partial \theta} - \sigma\frac{\partial e}{\partial \theta}\right] = 0,$$

$$\frac{\partial \mathcal{L}_2}{\partial \sigma} = -Eu'\left(d\right)x + E\left[h_x\left(x,e\right)\frac{\partial x}{\partial \sigma} + h_e\left(x,e\right)\frac{\partial e}{\partial \sigma}\right]\left(\beta-\alpha\right) - \mu E\left[-e - \tau\frac{\partial s}{\partial \sigma} - \theta\frac{\partial x}{\partial \sigma} - \sigma\frac{\partial e}{\partial \sigma}\right] = 0.$$

Using the compensated terms, we obtain:

$$\begin{aligned}\frac{\partial \tilde{\mathcal{L}}_2}{\partial \theta} &= -\operatorname{cov}\left(u'\left(c\right), x\right) + \mu E\left[\tau\frac{\partial \hat{s}}{\partial \theta} + \theta\frac{\partial \hat{x}}{\partial \theta} + \sigma\frac{\partial \hat{e}}{\partial \theta}\right] + \Lambda_\theta, \\ \frac{\partial \tilde{\mathcal{L}}_2}{\partial \sigma} &= -\operatorname{cov}\left(u'\left(c\right), e\right) + \mu E\left[\tau\frac{\partial \hat{s}}{\partial \sigma} + \theta\frac{\partial \hat{x}}{\partial \sigma} + \sigma\frac{\partial \hat{e}}{\partial \sigma}\right] + \Lambda_\sigma.\end{aligned}$$

where variables with hat ($\hat{s}, \hat{x}$ and $\hat{e}$) are the standard compensated demands and

$$\Lambda_\theta = E\left[h_x\left(x,e\right)\frac{\partial x}{\partial \theta} + h_e\left(x,e\right)\frac{\partial e}{\partial \theta}\right]\left(\beta-\alpha\right) + E\left[h_x\left(x,e\right)\frac{\partial x}{\partial a} + h_e\left(x,e\right)\frac{\partial e}{\partial a}\right]\left(\beta-\alpha\right)Ex,$$

$$\Lambda_\sigma = E\left[h_x\left(x,e\right)\frac{\partial x}{\partial \sigma} + h_e\left(x,e\right)\frac{\partial e}{\partial \sigma}\right]\left(\beta-\alpha\right) + E\left[h_x\left(x,e\right)\frac{\partial x}{\partial a} + h_e\left(x,e\right)\frac{\partial e}{\partial a}\right]\left(\beta-\alpha\right)Ee,$$

we can rewrite Λ_θ and Λ_σ as

$$\begin{aligned}\Lambda_\theta &= E\left(\beta-\alpha\right)\hat{H}_\theta + cov\left(\left(\beta-\alpha\right)\hat{H}_\theta, x\right), \\ \Lambda_\sigma &= E\left(\beta-\alpha\right)\hat{H}_\sigma + cov\left(\left(\beta-\alpha\right)\hat{H}_\sigma, e\right),\end{aligned}$$

with

$$\begin{aligned}\hat{H}_\theta &= h_x\left(x,e\right)\frac{\partial \hat{x}}{\partial \theta} + h_e\left(x,e\right)\frac{\partial \hat{e}}{\partial \theta}, \\ \hat{H}_\sigma &= h_x\left(x,e\right)\frac{\partial \hat{x}}{\partial \sigma} + h_e\left(x,e\right)\frac{\partial \hat{e}}{\partial \sigma}.\end{aligned}$$

We finally get:

$$\mu\theta = \frac{\text{cov}\left(u'\left(c\right),x\right)E\frac{\partial\hat{e}}{\partial\sigma} - cov\left(u'\left(c\right),e\right)E\frac{\partial\hat{e}}{\partial\theta}}{E\frac{\partial\hat{x}}{\partial\theta}E\frac{\partial\hat{e}}{\partial\sigma} - E\frac{\partial\hat{e}}{\partial\theta}E\frac{\partial\hat{x}}{\partial\sigma}} + \frac{-\Lambda_\theta E\frac{\partial\hat{e}}{\partial\sigma} + \Lambda_\sigma E\frac{\partial\hat{e}}{\partial\theta}}{E\frac{\partial\hat{x}}{\partial\theta}E\frac{\partial\hat{e}}{\partial\sigma} - E\frac{\partial\hat{e}}{\partial\theta}E\frac{\partial\hat{x}}{\partial\sigma}},$$
$$\mu\sigma = \frac{\text{cov}\left(u'\left(c\right),e\right)E\frac{\partial\hat{x}}{\partial\theta} - cov\left(u'\left(c\right),x\right)E\frac{\partial\hat{x}}{\partial\sigma}}{E\frac{\partial\hat{x}}{\partial\theta}E\frac{\partial\hat{e}}{\partial\sigma} - E\frac{\partial\hat{e}}{\partial\theta}E\frac{\partial\hat{x}}{\partial\sigma}} + \frac{-\Lambda_\sigma E\frac{\partial\hat{x}}{\partial\theta} + \Lambda_\theta E\frac{\partial\hat{x}}{\partial\sigma}}{E\frac{\partial\hat{x}}{\partial\theta}E\frac{\partial\hat{e}}{\partial\sigma} - E\frac{\partial\hat{e}}{\partial\theta}E\frac{\partial\hat{x}}{\partial\sigma}}.$$

In the case of negligible cross price effects the optimal tax are given by

$$\mu\theta = \frac{\text{cov}\left(u'\left(c\right),x\right)}{E\frac{\partial\hat{x}}{\partial\theta}} - \frac{E\left(\beta-\alpha\right)\hat{H}_\theta + cov\left(\left(\beta-\alpha\right)\hat{H}_\theta, x\right)}{E\frac{\partial\hat{x}}{\partial\theta}},$$
$$\mu\sigma = \frac{\text{cov}\left(u'\left(c\right),e\right)}{E\frac{\partial\hat{e}}{\partial\sigma}} - \frac{E\left(\beta-\alpha\right)\hat{H}_\sigma + cov\left(\left(\beta-\alpha\right)\hat{H}_\sigma, e\right)}{E\frac{\partial\hat{e}}{\partial\sigma}}.$$

B Welfarist solution in the case of persistent ignorance

In that case the first-best problem of the social planner can be expressed as

$$\begin{aligned}\mathcal{L}_5 &= \sum n_i\left[u\left(c_i\right)+\varphi\left(x_i\right)+u\left(d_i\right)+\alpha_i h\left(x_i,e_i\right)\right] \\ &\quad -\mu\left(c_i+x_i+d_i+e_i-w_i\right)\end{aligned}$$

and the FOC's are simply

$$\begin{aligned}u'\left(c_i\right) &= u'\left(d_i\right)=\varphi'\left(x_i\right)+\alpha_i h\left(x_i,c_i\right) \\ &= \alpha_i h_e\left(x_i,c_i\right)=\mu.\end{aligned}$$

To achieve such an optimum, one does not need any Pigouvian instrument. Lump sum individualized transfers suffice. The redistribution goes for sure from high income to low income agents. As to the other characteristic α_i it is not clear whether the more myopic benefit from redistribution. Individuals with low α_i tend to consume more x_i but less e_i than individuals with high α_i. Without further assumption, it is difficult to know the direction of redistribution.

As an example, assume two types with identical income, $\alpha_1 = 0$ and $\alpha_2 = \beta = 1$. Further, we take

$$h\left(x,e\right) = h\left(e - \nu x^2/2\right).$$

In the laissez-faire,

$$u'(c_1) = u'(d_1) = \varphi'(x_1);\ e_1 = 0.$$
$$u'(c_2) = u'(d_2) = \varphi'(x_2) - h'\left(e_2 - \nu x_2^2/2\right) \nu x_2 = h'\left(e_2 - \nu x_2^2/2\right).$$

If ν is very small, type 2 will benefit from redistribution; if ν is high (strong sin effect), type 1 will benefit from redistribution.

References

[1] Besley, T. (1988), A simple model for merit good arguments, *Journal of Public Economics* 35, 371-383.

[2] Camerer, C., S. Issacharoff, G. Loewenstein, T. O'Donoghue and M. Rabin (2003), Regulation for conservatives: behavioral economics and the case for "asymmetric paternalism", University of Pennsylvania Law Review 151(3), 1211-1254.

[3] Gruber, J. (2007), Commentary: sin taxes, unpublished.

[4] Gruber, J. and B. Koszegi (2001), Is addiction 'rational'? theory and evidence, Quarterly Journal of Economics 116, 1261-1303.

[5] Gruber J. and B. Koszegi (2004), Tax incidence when individuals are time-inconsistent: the case of cigarette excise taxes, Journal of Public Economics 88, 1959-1987.

[6] Laibson, D. (1997), Hyperbolic discounting and golden eggs, Quarterly Journal of Economics 112 (2), 443-477.

[7] Musgrave, R. (1959), The Theory of Public Finance, McGraw-Hill, New York.

[8] O'Donoghue, T. and M. Rabin (1999), Procrastination in preparing for retirement, in Behavioral Dimensions of Retirement Economics 125, Henry J. Aaron ed.

[9] O'Donoghue, T. and M. Rabin (2003), Studying optimal paternalism, illustrated by a model of sin taxes, American Economic Review (Papers and Proceedings) 93 (2), 186-191.

[10] O'Donoghue, T. and M. Rabin (2006), Optimal sin taxes, *Journal of Public Economics* 90, 1825-1849.

[11] Thaler, R. and C. Sunstein (2003), Behavioral economics, public policy, and paternalism, American Economic Review (Papers and Proceedings) 93 (2), 175-179.