The Dark Side of "Sin" Prices: Welfare Loss in Correcting Myopic Behavior *

Laura Abrardi[†], Carlo Cambini[‡]

October 2017
Preliminary version

Abstract

Can financial incentives promote the sustainability of good habits in the long run? We analyze the optimal price that a paternalistic social planner should impose to correct the behavior of time-inconsistent consumers. We focus on the situation in which consumers can obtain a financial reward by exerting effort in a virtuous behavior consistently over time. We find that financial incentives may have perverse effects on behavior when the effort must be exerted steadily over a long time horizon. Policies directed at increasing future self-control may even undermine the participation to programs directed at improving behavior, despite the complementarity between the participation and self-control decisions. As a consequence, it is socially preferable to reduce the size of the incentives both when myopia is low and when it is high. Moreover, the optimal level of incentives should be higher for naive consumers if myopia is high, and for sophisticated ones if myopia is low.

Keywords: Optimal paternalism, hyperbolic discounting, sin taxes, commitment, present-bias.
JEL codes: D11, D83, L51

*We would like to thank Luigi Balletta, Maria Bigoni, Christos Genakos, Marit Hinnosaar, Giovanni Immordino, Elisabetta Iossa, David E.M. Sappington, Yossi Spiegel, as well as participants to IIOC conference (Boston) for valuable discussion and suggestions to previous versions of the paper.

[†]Politecnico di Torino, DIGEP. Email: laura.abrardi@polito.it.
[‡]Politecnico di Torino, DIGEP; IEFE-Bocconi, Milan; and EUI. Email: carlo.cambini@polito.it.

1 Introduction

In March 2010, the US Patient Protection and Affordable Care Act allowed employers to vary the health insurance premium charged to employees by as much as 30 percent, based on meeting certain health status factors such as body mass index, tobacco use, physical fitness or activity levels. In 2013, such workplace wellness programs were offered by 75% of US employers (Kaiser Family Foundation, 2013), and by more than 90% of those with more than 50,000 employees (Mattke et al., 2013). Despite the size of financial incentives were substantial, uptake of worksite wellness programs remained limited, with participation rates below 50% of the eligible employees (Mattke et al., 2015; Persson et al., 2013; Robroek et al., 2009). In addition, even when employees did participate, the success of such programs was disappointing. In a year-long randomized controlled trial run among 197 obese employees of the University of Pennsylvania, a financial reward of 550$ for achieving a weight loss goal did not produce any significant effect on weight reduction (Patel et al., 2016)[1].

Financial incentives are increasingly being used by policymakers to motivate people to change their behavior, often as part of schemes aimed at reducing rates of obesity, smoking, alcoholism and other harmful habits. However, the evidence provided shows that people seem hesitant to uptake programs involving long run behavioral change, even when such programs are strongly incentivized. Moreover, when they do uptake such programs, their effort fades over time: people seem unable to follow up on their good intentions when long time horizons are involved.

The inconsistency of individuals' behavior over time arises some important questions. What is the role of financial incentives in fostering program uptake *and* subsequent effort? Why do they partially fail? And what is their optimal level to maximize the impact on behavior, while accounting for the social cost of such programs? In this paper we address these questions, that inherently are of great interest for policymakers.

From a theoretical point of view, the disappointing evidence on participation rates and program failures is particularly puzzling. In principle, the rates of program uptake and the probability to succesfully conclude it are expected to be positively correlated, as incentives motivate both the participation and the effort. The correlation should be even stronger if we account for the positive externality existing between the two: the expectation of higher effort should induce higher participation rates. Notwithstanding the financial incentives,

[1] A similar evidence can be observed in the repeated failed quit attempts of smokers (Pallonen et al. 1990; Chassin et al. 1996). In England, a national survey conducted in 2010 revealed that 63% of 11-15 year old regular smokers had unsuccesfully tried to quit, despite they typically smoke less and are more sensitive to cigarette prices than older smokers (Fuller, 2011).

the inexplicably low participation and success rates suggest that the interaction between participation and effort decision could be more tangled than we usually believe. This paper sheds light on this interplay between intertemporal behavior and incentives.

From a policymaking point of view, the fact that the consumers' myopia leads to a misbehavior, rather than to a misconsumption decision, has important implications in terms of welfare effects of the incentive schemes. The crux of the problem is that habits, differently from consumption, are not marketed. In the absence of a market price, it is not possible to directly impose a tax/subsidy on effort. Moreover, in practice the application of direct punishment or reward mechanisms conditional on the outcome of behavior is not feasible on large scale and outside the realm of experiments, either because its outcome is not observable, or because these mechanisms would be discriminatory. For instance, the excessive consumption of unhealthy food could be easily discouraged by taxing it, while the insufficient practice of physical activity cannot be punished directly in real contexts. Policymakers can only encourage this effort indirectly, for example by increasing the health insurance premium. The use of direct rather than indirect forms of incentives has deep welfare consequences. In the case of misconsumption (which is the focus of the literature on sin taxes), tax proceeds can be returned in a fixed form to consumers and no welfare loss occurs if consumers are homogeneuos in their myopia. However, we show that in case of misbehavior some welfare loss is unavoidable, as the correction of misbehavior necessarily requires the distortion of the consumption of another good. As a consequence, incentives to correct myopic behavior decisions are always socially costly, and policymakers need to determine their optimal level by trading-off the social cost of incentives and their effectiveness.

In our model, we focus on the non-marketed component of behavior that is not observable and not contractible. Because of these characteristics, only indirect (and socially costly) incentives can be applied to correct behavior in the form of a price distortion of a good whose consumption is affected by the consumer's actions.

In particular, we assume that a myopic consumer can exert quantity-saving efforts, which allow her to achieve some level of utility but with lower consumption, hence lower expenditure. However, to reap any benefit, the effort must be exerted consistently over two different periods, i.e. the efforts are complements. The first-stage effort can be seen as related to participation decisions, while the second-stage effort can be interpreted as related to self-control decisions. For instance, the effort of joining a worksite wellness program and then actually attending to it allows future savings in healthcare costs. Given that the benefit of the two efforts in terms of savings is accrued in the future, while its cost is paid in advance, a present-biased consumer misbehaves by underexerting effort in both participation and in self-control activities. To induce the consumer to exert more effort, a paternalistic social planner should increase the value of its future benefits. This is achieved by raising the price of the good whose consumption is reduced by the effort. Then, a trade-off arises between the deadweight loss of the raise in price and the effect it has on misbehavior.

Our analysis points out to which extent the level of incentives should be raised as a function of the levels of consumer's myopia and awareness of her present-bias problem. We identify the drivers undelying the consumer's decision and provide a theory that offers insights on the existing theoretical puzzle and empirical evidence.

We find that in some cases incentives may undermine effort decisions both in the long and in the short run for three reasons.

First, a very myopic consumer has actually more to gain from an increase in self-control or participation effort, but she is also more difficult to convince to exert it, as the benefits would only be collected in the future, which she values less because of her strong present bias. As a consequence, policies meant to increase efforts by very myopic consumers are less effective. A myopic consumer is less responsive to any nudge.

Second, if the consumer is very myopic, she will later refrain from actually exerting effort, thus wasting the cost of participation. Then, not only a very myopic consumer is more difficult to convince to participate, but, on top of that, participation may be worthless.

Third, an unexpected intertemporal effect arises between incentives on self-control and incentives on participation. When the consumer is nearly rational (i.e., her myopia is weak), incentives for increasing self-control dilutes the impact of such incentives on participation. Intuitively, if the consumer's myopia is weak, the self-control problem is minor. Even though financial incentives further increase self-control effort, the marginal benefit of doing so decreases, due to the higher marginal cost of effort. Anticipating the lower marginal benefits from self-control, the nearly rational consumer has weaker incentives to increase participation. In this situation, the impact of incentives on self-control decreases the impact of incentives on participation. Interestingly, the interplay between the effectiveness of incentives on self-control and on participation arises despite the fact that self-control effort and participation are complements, so that intuition would mistakenly lead to believe that incentives addressing the two should reinforce each other. This effect, that resembles a crowding-out phenomenon, is entirely originated by the intertemporal externalities of the efforts exerted at different times, and not because psychological factors such as the reduction of intrinsic motivation.

Because of the first two factors, policies addressing very myopic consumers could be excessively costly, and policymakers might want to target more aggressively consumers with a more moderate degree of present-bias. Because of the third factor, policies addressing the participation of weakly myopic consumers are made ineffective by the low effects on future self-control effort.

The bi-directional interaction between participation and self-control incentives has important policy implications, as it suggests that the optimal size of incentives should be non-monotonic in the degree of myopia. In particular, it should be low both when myopia is low (to preserve participation incentives) and when it is high (to account for the consumer's unresponsiveness and for the undermining effects of low self-control on participation).

The effectiveness of incentives is also strongly affected by the consumer's awareness of her own degree of myopia. We distinguish between *sophisticated consumers*, who are well aware of their myopia, and *naive ones*, who conversely do not anticipate their self-control problems (see Della Vigna and Malmendier, 2004; O'Donoghue and Rabin, 2003). We find that the degree of awareness has two opposite effects on the optimal price distortion. On the one hand, if the consumer is aware that she will later under-exert effort, she is less willing to invest in participation (as she knows that a low future effort dilutes the benefit of participation); then, the optimal level of incentives is higher to correct the deep social repercussions of the severe participation problem. On the other hand, if the consumer is aware that she will later under-exert effort, it is more difficult to convince her to invest in participation. From this perspective, since incentives on participation are less effective, they should instead be lowered from a social point of view.

This work contributes to the debate on the optimal policy in the presence of time-inconsistent consumers and it is related to three main strands of literature, corresponding to the two types of instruments (financial and not) available to policymakers to correct the consumers' behavior, and to the motivation for the crowding out effects emerging from our analysis.

The first strand of literature investigates the economic incentives, mainly in the form of sin taxes applied on a consumption good (see, e.g., O'Donoghue and Rabin, 1999a, 1999b, 2003 and 2006; Haavio and Kotakorpi, 2011). Since the consumer's utility depends on the quantity alone, raising the price (and then returning the proceeds to the consumer by means of a fixed subsidy) fully corrects the distortion in the consumption decision caused by the consumer's present bias and the first best level can be implemented. We add to this literature by modelling consumers' behavior. We do this by introducing an effort which could be exerted before the consumption decision, and which has an impact on the demand of the good. A typical result of the literature on sin taxes is that if all consumers are affected by self-restraint problems to the same degree, then a price distortion allows to achieve the first best, with no welfare loss. We show that this is no longer true in a framework in which the consumer's mistake pertains to a behavior. A similar conclusion is obtained by Lerner (1970) and Dixit (1970), when studying the optimal taxes when one sector is not taxable (as in the case of effort). Note, however, that their focus is on the optimal allocation of taxes between the taxable and non-taxable sector, while ours is related to the impact of incentives on the consumer's behavior in the long run.

The second strand of literature examines the effects of financial incentives on intrinsic motivation. This literature argues that providing extrinsic rewards can be counterproductive, because it may destroy people's intrinsic motivation, ultimately reducing effort (see, e.g., Kreps, 1997; Deci et al., 1999). The present work complements this literature, as we provide an additional reason for the observed crowding-out effects of financial incentives on effort. The effect that we identify is entirely unrelated to psychological factors, as it pertains exclusively to the intertemporal externality between long and short run effort decisions. Furthermore, note that the crowding-out effect due to intrinsic motivation only applies to cases in which intrinsic motives are actually present, and is thus inadequate to explain the ineffectiveness of incentives in many other situations. For example, the participation to worksite wellness programs is higher for healthier and wealthier employees (Mattke et al., 2015), i.e. individuals who have a high intrinsic motivation to stay healthy, as they exerted effort even outside the context of an incentivized program[2]. However, the low participation rates by less healthy individuals cannot be explained simply by the crowding out of intrinsic motivation.

The third strand of literature pertains to the non-pricing instruments whose aim is to reduce the consumer's choice set (Strotz, 1955; Gul and Pesendorfer, 2001; Ashraf et al., 2006; Laibson, 2015). Their purpose is to curb from the outset any possibility to succumb to the temptation of misbehavior. The social planner must then induce the consumer to commit to a restricted set of alternatives. The problem with commitment is that it is a problematic prediction, as very little of it can be observed in the economy. A growing experimental literature (see, e.g., Giné et al., 2010; Kaur et al., 2015; Laibson, 2015) finds that only a minority of subjects choose to tie their own hands, and they rarely express a willingness to pay to have their choice-set reduced. While commitment and participation can be seen as opposite phenomena (the former restricting the choice set, the second enlarging it), they actually represent two sides of the same coin, as they entail a decision that changes the course of a future behavior. In fact, in our model the participation decision could be symmetrically

[2]By definition, an individual is intrinsically motivated if he performs an action even in the absence of extrinsic incentives. Therefore, the status of health and physical fitness are usually positively correlated with the levels of intrinsic motivation.

be seen as the commitment not to misbehave in the future.

This work can be seen as an attempt to bridge the gap between the literature on economic incentives and that on commitment. In fact, we look into the economic incentives to induce the consumer to commit to a situation in which the possibilities of harmful behavior are diminished.

The rest of the paper is organized as follows. In Section 2 we introduce a general model to present the central idea and relate it to the literature on sin taxes. In Section 3 we introduce the model and Section 4 characterizes the social optimum, i.e. our benchmark. Section 5 analyzes the socially optimal price to induce a time-inconsistent consumer to invest in effort. We specifically identify how incentives drive the participation and self-control decisions. Section 6 examines the role on conusmer's myopia and how it affects the effectiveness of incentives on participation and self-control. Section 7 study the role of consumer's awareness on incentive design and Section 8 concludes. In Appendix 1 we extend the basic model by assuming that the efforts are a complement, rather than a substitute, of consumption. All proofs and technical details are relegated in Appendix 2.

2 The economic implications of misbehavior: a general framework

In this section we explore the main economic effects of the consumer's present bias on a behavior decision, rather than on a consumption decision. We adopt the same framework typically used by the literature on sin goods (O'Donoghue and Rabin, 2003, 2006; Haavio and Kotakorpi, 2011). We then focus on the results obtained by this literature on the welfare effects of the incentives to correct a biased consumption decision. Then, we analyze the situation in which the consumer's present-bias affects a behavior, i.e. an effort. Finally, we show that this assumption has stark implications both on the type of incentives, and on their social cost.

2.1 Incentives and welfare effects in case of misconsumption

We borrow from the standard literature of sin taxes (see, e.g., O'Donoghue and Rabin, 2003) and consider initially a simple consumption model with two generic goods, whose quantities are denoted by x and y. For simplicity, we assume that the production of both goods entails constant returns to scale; we denote by c and d the marginal costs of x and y, respectively. The consumer's preferences are expressed by the utility function $V(x, y)$, with $V_x, V_y > 0$,

$V_{xx}, V_{yy} < 0$[3].

Suppose that there is an asynchrony between the benefits and cost of consumption. For example, the cost (or price) of y is paid at time t, while that of x is paid in $t+1$, when the surplus $V(x,y)$ is also achieved.

If there is no time-consistent discount factor, a social planner maximizes the consumer's net surplus $V(x,y) - cx - dy$ by choosing x^* such that $V_x = c$ and y^* such that $V_y = d$.

Our focus is on the case of a myopic consumer, who is biased for the present over the future. In particular, she discounts all future flows by a constant coefficient $\beta < 1$.

When there is some asynchrony in consumption, the consumer's myopia plays a role in that some flows are underrated by the consumer's present self, generating a distortion between the actual and optimal consumption. In our set up, the consumption of y generates a future utility that is discounted excessively by the consumer's present self. For example, y could be the hours of gym attendance, which causes a future benefit in terms of reduced health costs and x is the numeraire.

Denoting with p_x and p_y the prices of the two goods, the consumer chooses the quantities x^c and y^c so as to solve

$$\max_{x,y} \beta \left(V(x,y) - p_x x\right) - p_y y.$$

Then, x^c solves the FOC $V_x = p_x$ while y^c solves $V_y = p_y/\beta$. From the comparison between x^c, y^c and x^*, y^*, we can immediately see that the social planner can realign the consumer's problem to the first best and achieve the optimal consumption by setting

$$\begin{aligned} p_x^* &= c \\ p_y^* &= \beta d, \end{aligned}$$

namely by imposing a distortion of the price of the good whose demand is affected by the consumer's present bias. In this specific example, since $\beta < 1$, we have that $p_y^* < d$: the price is set below cost and the distortion is a subsidy. In fact, the consumer discounts excessively the benefits of the consumption of y and thus underconsumes it. Here, a subsidy is clearly needed to reach the first best consumption level. However, the same argument can be easily applied to the case in which y causes future costs, like in the case of tobacco or alchool

[3] In the literature on sin taxes, the utility function is usually assumed separable. For example, O'Donoghue and Rabin (2003) assume that $V(x,y,Z) = \rho \ln x - \gamma \ln y + f(Z)$, where $\rho, \gamma > 0$ are exogenous parameters and Z is the vector of other goods; Haavio and Kotakorpi (2011) assume that $V(x,y,z) = v(x) - h(y) + z$, where z is a composite good.

consumption, causing overconsumption[4]. In this case, the overconsumption could easily be corrected by the introduction of a tax that raises the price above the marginal cost[5].

If $\beta = 1$ (i.e., the consumer is not present biased), the optimal price is at the marginal cost. Conversely, if $\beta \neq 1$, a tax/credit is applied on the price p_y to correct the distortion of its consumption. In the case consumers are homogeneous in β, the first best can always be achieved.

2.2 Modelling behavior

The above discussion focused on a specific category of consumers' choices, which are related to purchasing activities. In fact, the goods x and y were both acquired at a market price.

However, in many cases our utility depends also on actions that require minimal or no financial outlay and that are not marketed. These behaviors concern the everyday habitual element of an individual's lifestyle, and their cost is mainly related to the effort to avoid wasteful actions or to engage in constructive endeavours. Such actions are beneficial because, by paying attention to our behavior, we can achieve a greater utility given the same levels of consumption. For example, by exercising regularly or attending to our studies we obtain long-run benefits in terms of improved health or better job conditions[6].

The fact that the consumer's choices may concern behavior decisions as well as buying decisions is relevant whenever a less-than-rational consumer enters into the picture.

To see this, suppose that y represents a non contractible behavior rather than a physical good. For example, y is the effort to exercise regularly, which allows to improve our health (i.e., utility) given the same level of consumption x. As an effort, y is not traded on the market and it does not have a market price, but only a cost d (to make our point and keep the parallelism with the standard model presented above, in this section we assume a linear effort cost)[7].

---

[4] A possible way to model this case is by setting an inter-temporal discount factor $\beta > 1$, to account for the fact that the consumer overestimates the future utility of consumption.

[5] Note also that the result remains true regardless of whether the choice of x and y is simultaneous or sequential.

[6] Applications abound. In the energy field, we can reduce the future bill by exerting effort in behaviors such as shutting the windows when air conditioning is on, using a clothes line rather than a tumble drier, putting a full load of washing on rather than a half load (Barr, Gilg and Ford, 2005). Other examples can also be found in the literature on self-control or procrastination (O'Donoghue and Rabin, 1999a, 1999b).

[7] We remark once more that y is a pure effort, net of any price effects; translated in our example, this means that y is the effort that we need to exert, *in addition* to the fact that we pay the subscription to the gym.

In this framework, the quantities (x^*, y^*) are still the solution of the social planner's problem. However, the consumer now solves

$$\max_{x,y} \beta \left(V(x, y) - p_x x\right) - dy. \tag{1}$$

Notice that here only one policy instrument is available, namely p_x. The effort is not contractible, so no incentive can be conditioned to the level of y. As a consequence, it is now impossible to achieve (x^*, y^*), even when consumers are perfectly homogeneous in the present bias, and a welfare loss must necessarily occur[8].

The extent of the welfare loss depends on some features of the utility function $V(x, y)$. To see this, let us find the first order conditions of the consumers' problem (1). We assume without loss of generality that the choice of x and y is simultaneous. The consumer chooses x and y such that

$$V_x(x, y) = p_x \tag{2}$$

$$V_y(x, y) = d/\beta. \tag{3}$$

From (2), the level of p_x affects the choice of x, given y. If $V_{xy} \neq 0$, the choice of x in turn impacts on $V_y(x, y)$ in (3). As a consequence, when $V_{xy} \neq 0$, the price p_x affects not only the choice of x but also that of y. Then, the social planner can exploit p_x to correct y, i.e. the consumer's misbehavior. The downside of this course of action is that a distortion of p_x is also affecting the choice of x from (2). As a consequence, when $V_{xy} \neq 0$, a trade off emerges between correcting the inefficient behavior y and causing an inefficient consumption x.

Conversely, when $V_{xy} = 0$, correcting the behavior through prices by exploiting the cross-elasticities of demand is impossible, so that there is no point in applying a tax/subsidy to the consumption of x and a welfare loss cannot be avoided.

In either case ($V_{xy} \neq 0$ and $V_{xy} = 0$), the first best cannot be achieved, despite the fact that consumers are homogeneous in their myopia.

Of course, from a policymaking perspective, the interesting case is the former, where the two marginal utilities are not independent, as this situation leaves room for the application of incentive schemes. In this case, pricing policies have to be carefully calibrated to trade-off an inefficient behavior and an inefficient consumption.

Standard theory on incentives and moral hazard builds on the assumption that individuals can be incentivized by imposing direct rewards/punishments on the observable outcomes

[8] Note that the inefficiency is originated by the biased proportion of x and y that results from the solution of problem (1). Therefore, it cannot be solved by adopting a different tariff structure for x. In fact, even if a two-parts tariff were applied instead of the per unit price p_x, the fixed component of the tariff would have no impact on the combination of x and y

of their behavior. Overconsumption, for example, can be prevented by a tax, or even by a ban or cap. However, in some situations such schemes can be difficult to implement in practice: the outcome of behavior may not be observable, the direct incentives may be discriminatory, rewards conditional on losing weight or passing an exam may not be sustainable on a large scale and in the long term. In these cases, policymakers can only implement indirect forms of incentives by operating on a substitute (or complement) consumption good. Our analysis contributes to the literature by focusing on these situations. Since such mechanisms are necessarily costly, it becomes important, from a policymaking point of view, to determine their optimal level, by accounting for the trade-off between the cost of the incentive and its effectiveness on behavior.

The welfare effects of a price distortion have been extensively studied ever since Ramsey (1927) raised the question of optimal taxation. In the literature on indirect taxes (Lerner, 1970; Baumol and Bradford, 1970; Diamond and Mirrlee, 1971; Atkinson and Stiglitz, 1972, 1976), the focus is on the optimal tax applied to different goods or sectors. Under some respects, our work draws from this literature, in that we account for the deadweight loss of a higher price. However, we depart from this literature in three important ways.

First, in the literature on indirect taxation the price distortion is necessary due to the imposition of an exogenous constraint to raise some fixed revenues from taxes. Then, the social surplus is affected only through the deadweight loss caused by the departure from marginal cost-pricing. Conversely, in our model, the price distortion is necessary to correct a myopic effort. Then, the surplus is affected not only through the deadweight loss, but also by the endogenous effect of the price on the consumer's behavior, which shifts the demand function.

Second, the literature on indirect taxation has a static perspective, as it considers the problem of allocative inefficiency between different sectors (depending on the elasticities and cross-elasticities of demands) and between different taxation schemes (i.e., indirect vs. direct taxation). Conversely, we look at the dynamic interaction of the consumer's inconsistent decisions over a long time horizon.

Third, considering an effort rather than a marketable good has important welfare implications when we consider a long time span. Intuitively, since the effort is not contractible, incentives cannot target separately efforts exerted at different times. Then, some intertemporal externality may arise, such that short run incentives interact with the long run ones. Conversely, in the case of goods a social planner could impose differentiated prices for any period, so as to address, and possibly resolve, the intertemporal externality (see, e.g., Beshears et al., 2005).

The interaction between long and short run incentives on correcting misbehavior is the focus of this paper. We show that the interaction gives rise to unexpected dynamics, with important policy implications on the design of the incentives on participation and future good behavior. We find that the social planner might even give up correcting the consumer behavior, and this is more likely to happen in both the cases of strong and weak myopia, as in these situations the intertemporal effects are more intense.

2.3 A foreword on alternative regulatory tools

This paper studies how indirect taxes on the consequences of our misbehavior can provide incentives to correct it. Before proceeding with the analysis, a natural question should be addressed: isn't it possible to correct the misbehavior by instruments other than taxes or prices or, more in general, financial incentives?

The answer to the question is definitely positive. As a matter of fact, the government could try to steer the consumers' behavior through a variety of regulatory measures. Some of these measures aim at raising the direct cost of misbehavior, or at reducing the utility it provides. For example, smoking regulations reduce the consumers' utility by forcing them to smoke outdoors. Other measures, such as bans or restrictions, prevent misbehavior at the outset. There are bans that apply only to a subset of situations or individuals. For example, bans on unhealthy food in schools, or bans on alcohol or on gambling for underage. There are also bans targeting the whole population, such as bans on drugs.

While all these measures are certainly successful to some extent, they are also unable to completely fix the problem of consumers' misbehavior for three reasons.

First, they can prevent misbehavior in some situations or for some category of individuals, but they cannot eradicate it in general. Even if schools are forbidden to provide junk food, children can still consume it at their own homes. Banning underage from gambling or alcohol consumption leaves the problem of addiction for people of legal age. Forcing people to smoke outdoors can certainly discourage smoking at public places, but it cannot be enforced in people's houses.

Second, banning an activity does not *per se* prevent people from doing it: drugs are consumed despite their illegal status. The effectiveness of the policy depends on the efficacy of monitoring activities, which in real situations is not perfect. In some cases monitoring is altogether impossible: social planners have no control over the amount of exercise people exert, what they eat or how much they smoke at home.

Third, even if monitoring were perfect, some bans can be circumvented. Underage could access to alcohol by means of friends of legal age. Restricting the time of day when alcohol is available induces a forward-looking consumer to buy in advance.

Indeed, the empirical evidence on the effectiveness of policy measures other than taxes is mixed (Stehr, 2007; Bernheim, Meer and Novarro, 2016; Carpenter and Eisenberg, 2009). In addition, these measures are not without their own social cost. Specifically, bans or other forms of commitment in practice restrict the choice set, causing a disutility to consumers (Gul and Pesendorfer, 2001; Beshears et al., 2005), to the point that consumers' have a low willigness to pay to commit (Laibson, 2015; Augenblick et al., 2014). Sales restrictions work only if they provide a commitment device to time inconsistent consumers; however, they also increase the inconvenience of time consistent consumers, who simply buy in advance. Indeed, Hinnosaar (2016) shows empirically that in terms of consumer welfare, sales restrictions may perform worse than taxes if consumers are heterogeneous.

In conclusion, regulatory measures such as bans or sales restrictions can certainly alleviate the problem of consumers' misbehavior. However, by constraining the consumer's choice, they might entail social costs if consumers display some form of heterogeneity. Even more importantly for our purposes, in most cases they are not able to fully eradicate the problem and some residual amount of misbehavior persists. Specifically, all behaviors that are not observable and non contractible lie beyond the outreach of any direct regulatory intervention of this nature.

This specific component of misbehavior, the one which is not observable and non contractible, is the focus of our work. Translated in terms of our model, the effort variable y is normalized to represent only the specific component of behavior that is non contractible and therefore cannot be addressed by direct regulatory instruments such as bans, caps, floors monitoring activities and incentives applied directly on the contractible component of effort. The only way to discourage non observable misbehavior is through indirect incentive measures, by making consumers voluntarily refrain from it, such as by raising the cost of its consequences. We aim at studying the optimal design of these incentives and the trade-off that should drive the regulatory action.

3 A simple model

In this section we set up a simple three-periods model. The three stages are necessary to study respectively the participation, behavior and consumption decisions. We allow consumers to be time-inconsistent and (possibly) overconfident. Our aim is to find the welfare-maximizing price that provides the right incentives to guide the consumer's behavior. Specifically, we study the role of the price of the consumption good on i) the incentives to participate (i.e. incentive on the long run behavior), ii) the incentives to exert self-restrain effort (i.e. on the short run behavior), and iii) the interaction between the short and long run incentives.

Due to the complex interaction of the forces at play, to make our point in the most straightforward way we will focus on the case in which the efforts and the consumption are substitutes. In Appendix 1 we extend the analysis to the case of complementarity and we show that this assumption will not impact on our conclusions.

A market is characterized by the presence of a single representative consumer. In choosing her level of consumption, the consumer seeks to maximize her utility $V(s)$, that is a function of the level of service s provided by the good that she purchases[9], with $V_s > 0$ and $V_{ss} < 0$. The consumer's level of service is a function of the quantity x consumed and of her self-control effort y to increase the efficiency of her consumption, such that $s = x + y$. The self-control effort allows the consumer to achieve the same level of service s with a lower consumption x. In particular, one additional unit of effort allows to save one unit of quantity consumed. The effort y could entail, for example, exercising regularly to save on healthcare costs, so that s can be interpreted as the level of future health and x as the quantity of healthcare purchased [10]. The self-control effort is not observable and not contractible and it entails an increasing cost $d(y)$. We assume that $d(0) = 0$ (effort entails no fixed cost), $d_y > 0$ (effort is costly), $d_{yy} > 0$ (the cost of effort is convex), and for technical reasons $d_{yyy} \geq 0$ (which is a sufficient condition for the social planner's problem to be concave, see also Laffont and Tirole, 1993). The effort can be exerted only if the consumer paid a participation cost k at an earlier stage (e.g., a sort of 'participation effort'to a wellness program): if the consumer fails to pay k, she is in practice committing to misbehave in the future. This implies that the participation and self-control effort decisions are complementary: to actually achieve some savings, the consumer must invest consistently in good behavior over time. The value of k is perfectly observed by the consumer, while the social planner only knows its distribution. For simplicity, we assume that k is uniformly distributed in the interval $[0, K]$, with $K > 0$[11].

If the consumer committed not to exert self-control effort, the level of service is $s = x$.

[9]For a similar approach based on the consumer's level of service – the comfort– applied to energy efficiency, see also Chu and Sappington (2013).

[10]The model could also be applied with few, if any, adaptations, to the education case (studying assidously to earn a better income) or to the energy sector (coordinating through car pooling to save fuel, switching off the lights whenever we leave a room or using the washing machine at full capacity so as to reduce our energy bills by lowering the consumption of electricity or water).

[11]Importantly, we will show later that the consumers' heterogeneity with respect to k is not the cause of inefficiency. In fact, we will verify that the origin of the deadweight loss is the misbehavior, i.e. the effort bias, which compels the use of indirect incentives.

Conversely, if the consumer paid k, she later has the possibility to exert y and the level of service is given by $s = x + y$.

The production of the good entails a constant marginal cost c, which includes also the social cost imposed on society. The assumption of constant marginal costs is made with no loss of generality on our main results, but significantly streamlines the analysis.

The social planner charges a uniform price p on the consumption x[12]. Given the price, the consumer decides the quantity x to purchase and the self-control effort y; based on these decisions, she achieves the level of service $s = x + y$.

The timing of the game is the following:

1. in $t = 0$ the social planner announces the price p and the consumer privately observes k. Then, the consumer decides whether to commit and, depending on her decision, pays the commitment cost k or not;

2. in $t = 1$ the consumer decides a level of effort $y \geq 0$ if she previously paid k; else, $y = 0$; she also pays the effort cost $d(y)$;

3. in $t = 2$ the consumer decides the quantity consumed x and receives the net surplus $V(s) - px$ from the consumption.

The social planner seeks to maximize the welfare W, defined as the present value of the consumers' utility net of production and effort costs:

$$\max_{x,y} W = \delta\left(V(s) - cx\right) - d(y), \tag{4}$$

where $\delta \leq 1$ is the time-consistent discount factor.

In Section 4 we find the first best policy. In Section 5 we then introduce the assumption of a myopic behavior by the consumer and analyze its impact on the welfare-maximizing policy.

Before proceeding with the analysis, it might be worthwhile to spend a few words on the timing of the model. The reasons to introduce a three-stage game, by separating the participation, effort and consumption decisions are manifold.

First, in many real situations our behavior is conditioned on decisions to which we committed in the past. For example, we can attend to our studies only if we previously enrolled to a university. We can practice swimming only if we sign up to a swimming

[12]In this framework, a fixed fee would merely have redistributive purposes, by returning the proceeds of the production to consumers; we will therefore focus on the optimal per-unit price (the same approach is used in the literature on sin taxes; see, e.g., O'Donoghue and Rabin, 2003 and 2006).

facility. We can exploit some energy-efficiency device only if we previously invested in it. In all these examples, the preliminary actions of investing in higher education or in gym attendance constitute a necessary condition for the possibility to exert the future effort, and in general terms they could be considered an effort in itself. Generalizing this argument, we could say that the benefits of good behavior are often the result of a continued effort that is consistently exerted over a long time span. Typically, keeping a responsible behavior one single time is not sufficient to reduce consumption in the long term.

Second, and more on a technical level, if we had only a two-periods game (for example, by dropping from the model the participation stage), the consumer's myopia would not generate a time-inconsistency problem, but only a bias of the present over the future, analogously to many other kinds of distortion. For instance the behavior generates negative externalities on the environment which the consumer does not take fully into account. Then, a two-periods game would not allow us to study the consequences of the time-inconsistency on the optimal design of incentives.

Third, the decision to participate depends on the expectation about our future behavior. We do not apply to a university or to the gym if we know that we are not going to attend[13]. The expectation about our future behavior is then crucial to determine the long run effort that we exert. As a consequence, the pricing policy should take into account the consumer's perception about her future behavior. Then, introducing a participation stage allows to study the role of the consumer's awareness of her own myopia, and how this awareness affects the level of incentives to correct the present-bias.

Finally, the introduction of an initial stage of the game in which the consumer decides to commit to some future effort allows us to study the effect of the consumer's myopia on the inter-temporal externality generated by her decision to commit in $t = 0$ on the decision to exert effort in $t = 1$. Namely, we can gain some (rather unexpected) insights on the interaction between the incentives on the short run decision to exert effort and the incentives on the long run decision to participate.

4 Benchmark: time-consistent consumers

The social planner's problem is solved backwards. In $t = 2$, given the effort y exerted in $t = 1$ and the price p set in $t = 0$, the quantity x demanded by a consumer is the solution of

$$\max_x V(s) - px.$$

[13] Similarly, we insulate our house only if we expect to 'exploit' the energy-efficiency investment by keeping the windows closed.

The demand x is thus expressed by the FOC $V_s(s) = p$. Then, the consumer achieves a level of service

$$s^* = V_s^{-1}(p). \tag{5}$$

Note that, from (5), a higher price reduces the achieved level of service s (as $V(s)$ is concave in s). In turn, a lower level of service implies a lower gross surplus $V(s)$.

From (5), the consumer's demand for x is

$$x^* = V_s^{-1}(p) - y \tag{6}$$

given $y \geq 0$ exerted in the previous stage. From (6), the quantity demanded is a decreasing function both of the price, and of the impact y of effort on the level of service. In particular, equation (6) shows clearly that y represents the quantity saved due to the effort for a given price p. In fact, given the price p, an increase of y shifts the demand function downward and allows customers to achieve the same level of service $s^* = V_s^{-1}(p)$ with lower quantities consumed x.

In $t = 1$, and provided that the consumer paid k in the previous stage, she chooses the effort y by solving

$$\max_y \delta[V(s^*) - px^*] - d(y),$$

i.e., by using (5) and (6),

$$\max_y \delta[V\left(V_s^{-1}(p)\right) - p\left(V_s^{-1}(p) - y\right)] - d(y). \tag{7}$$

The FOC of (7), $d_y(y) = \delta p$, allows to obtain the first best effort in case the consumer invested in k in the previous period:

$$y^* = d_y^{-1}(\delta p). \tag{8}$$

The marginal benefit of exerting effort is represented by the present value of the savings realized through it. Such value is obtained by multiplying the quantity (i.e., 1) saved by one unit of effort by its present value δp. In the optimum, the marginal benefit of the effort, δp, is equal to its marginal cost $d_y(y)$.

Note that, from (8), the optimal level of effort is increasing in δp, i.e. the price works as an incentive for the consumer to increase the effort. In fact, the effort is a substitute for consumption. Therefore, when the price of consumption increases, it is optimal to increase the quantity-saving effort.

In $t = 0$, the consumer invests k only if the present value of the net surplus obtained from participation is greater than that obtained with no participation, i.e.

$$\delta^2 \left[V(s^*) - p\left(s^* - y^*\right)\right] - \delta d(y^*) - k \geq \delta^2 \left[V(s^*) - ps^*\right]. \tag{9}$$

Condition (9) can be rewritten as

$$k \leq \epsilon^*, \tag{10}$$

where $\epsilon^* = \delta^2 py^* - \delta d(y^*)$ (or, using (8), $\epsilon^* = \delta^2 pd_y^{-1}(\delta p) - \delta d(d_y^{-1}(\delta p))$). Inequality (10) says that the committment not to misbehave is socially efficient only if its benefit $\delta^2 py^*$ in terms of quantities saved is higher than its direct cost k plus its indirect cost $\delta d(y^*)$ due to the first period effort. In particular, the benefit is given by the present value of the savings, and it is equal to the present value of the price $\delta^2 p$ multiplied by the real saving y^*.

From (10), the price affects the participation decision in two ways. First, a higher price increases the nominal value of savings; second, since savings are more valuable, the consumer is enticed to exert more effort (i.e. $y^*(p)$ increases with p).

The interpretation of ϵ^* emerges directly from inequality (10). Recall that the consumer's participation cost k is distributed uniformly in the range $[0, K]$. Since the consumer participates only if $k \leq \epsilon^*$, then ϵ^* represents the degree to which the consumer is willing to commit not to misbehave, and it depends on the cost of participation. Quite importantly for our purposes, ϵ^* can be alternatively interpreted as the customers' base, i.e. as the set of consumer's types who decide to participate because they have a participation cost $k \leq \epsilon^*$.

To determine the social planner's optimal policy we now turn to welfare. The welfare can be expressed by

$$\begin{aligned} W = & \int_0^{\epsilon^*} \left[\delta^2 \left[V(s^*) - c\left(s^* - y^*\right)\right] - \delta d(y^*) - k\right] dk \\ & + \int_{\epsilon^*}^{K} \left[\delta^2 \left(V(s^*) - cs^*\right)\right] dk, \end{aligned} \tag{11}$$

where the firt term is the welfare of consumer's types who participate in $t = 0$, while the second term is the welfare of consumer's type who do not participate. From the FOC of expression (11) w.r.t. p, we find that the first best price corresponds to the marginal cost, as expressed by the following Proposition.

Proposition 1 *With time-consistent consumers, the welfare is maximized at the price $p^* = c$.*

Proof. See Appendix 2. ■

The marginal cost pricing implies that the first best effort is $y^* = d_y^{-1}(\delta c)$ and the consumer participates only if her cost k is lower or equal to the value $\epsilon^* = \delta^2 c d_y^{-1}(\delta c) - \delta d(d_y^{-1}(\delta c))$.

The result of Proposition 1 represents our benchmark for the analysis of the following sections.

5 Time-inconsistent consumers

We assume now that the consumer has present-biased preferences, which the literature (starting from the seminal work of Laibson, 1997, and, more recently, O'Donoghue and Rabin 1999a, 1999b; Della Vigna and Malmendier, 2004) usually represents through the simple functional form

$$U^t(u_t, ..., u_T) = u_t + \beta \sum_{\tau=t+1}^{T} \delta^{\tau-t} u_\tau,$$

where u_τ is her utility in period τ, the parameter δ is the standard long-run discount factor and the parameter β can be interpreted as the coefficient of short-run discounting. A time-consistent consumer has $\beta = 1$, while $\beta < 1$ represents a time-inconsistent preference for immediate gratification and thus denotes short-run impatience. Because of $\beta < 1$, a myopic consumer discounts excessively future flows and this results in self-control problems, originating a bias between what the consumes does, and what she thinks she should do in terms of her own long-term best interest.

In this section we study the time-inconsistent consumer's decisions in terms of effort (Section 5.1) and, going backward, participation (Section 5.2), given a generic price p. Then we use these results to define the welfare-maximizing price (Section 5.3), and we compare it to our benchmark with time-consistent consumers. We denote the results of this section with time-inconsistent consumers by the superscript *inc*.

5.1 Incentives for effort

In $t = 2$, the consumer chooses the level of consumption $x^{inc} = V_s^{-1}(p) - y$, thus achieving the level of service $s^{inc} = V_s^{-1}(p)$. Proceeding backward to the previous period, and provided that in $t = 0$ she participated by paying k, in $t = 1$ the consumer chooses the effort y^{inc} so as to maximize

$$\max_y \quad \beta\delta \left[V(V_s^{-1}(p)) - p(V_s^{-1}(p) - y)\right] - d(y), \tag{12}$$

whose FOC, $d_y(y) = \beta\delta p$, allows to obtain the effort

$$y^{inc} = d_y^{-1}(\beta\delta p). \tag{13}$$

Equation (13) shows that, in the consumer's optimum, the marginal cost of effort $d_y(y)$ is equal to its marginal benefit $\beta\delta p$. By comparing (13) with (8), it is straightforward to see that, given the same price p and provided that the consumer participated in the previous period, the myopic consumer under-exerts effort in $t = 1$, i.e. $y^{inc} < y^*$ for any p. The reason lies in the fact that the benefits of effort are accrued in the future, while its costs are immediate, therefore the benefits are underweighted by a present-biased consumer and the effort is lower than in the first best. Even when the consumer invests k in $t = 0$, she will then under perform in terms of effort in $t = 1$, causing excessive consumption given the investment of the previous stage.

To induce a present-biased consumer to exert a higher level of effort in $t = 1$, it would be sufficient to raise the price. In particular, since $y^{inc} = d_y^{-1}(\beta\delta p)$ and $y^* = d_y^{-1}(\delta c)$, it would be sufficient to impose a price $p = c/\beta$ to achieve the first best level of effort. The price here works as an incentive by increasing the marginal benefit of effort, i.e. the value of the quantity saved by each additional unit of effort.

However, it turns out that the raise in price does not come without its own cost, and the reason must be ascribed to the source of the distortion, as remarked below.

Remark. *If the consumer's bias involves misbehavior rather than misconsumption, the correction of the distortion of y through the price necessarily entails some welfare loss.*

When self-control problems affect consumers to the same degree (i.e., the population of consumer's types is homogeneous in β), raising the price above the marginal (social) cost can fully correct the over-consumption problem (O'Donoghue and Rabin, 2006). This result is well established in the literature on sin taxes on myopic consumption.

However, accounting for consumers' behavior (i.e., having consumers who can commit to a 'good' behavior y^* that allows to save on x) affects dramatically the welfare effects of the policy. In fact, if the consumer's bias concerns only a consumption decision, the price is merely a market instrument which can achieve the first best if all consumers display the same degree of myopia (i.e. the same β)[14].

Conversely, if we account for the possibility of consumers to exert self-restraint, by means for example of a quantity-saving, non-marketed effort, the price is not only a market

[14] In fact, typically the concern of the literature on optimal sin taxes is not when all consumers are equally myopic, but when they are heterogeneous in their myopia, so that helping irrational consumers is detrimental to rational ones. See, e.g., O'Donoghue and Rabin (2003, 2006) and Haavio and Kotakorpi (2011).

instrument (through which the consumers decides the quantity to be consumed) but it is also an incentive for good behavior. In this case, the higher price $p = c/\beta$ makes the consumer choose the first best level y^* of effort, but will later cause underconsumption of x *given* y, so that the consumer's surplus $V(s)$ is reduced. As a consequence, raising the price has negative side-effects on the consumers' surplus, *even when all consumers display myopia to the same degree*. In fact, imposing a price $p = c/\beta$ above the marginal cost (as $\beta < 1$) fully solves the misbehavior problem, but it also causes a welfare loss as $V(s) = V(V_s^{-1}(p))$ for all consumers. A trade-off then emerges between the welfare gain of improving the consumers' behavior and the welfare loss due to the higher price, even if consumers are homogeneous in terms of the time-inconsistency parameter β.

The result that only second-best outcomes can be achieved after a price distortion is well-known by the literature on indirect taxation, which focuses on the optimal price distortion that should be applied when the social planner needs to raise a fixed revenue from taxes. The usual Ramsey distortion that arises in this context may be aggravated by several additional effects, related to the presence of a non-taxable sector (Lerner, 1970), externalities (Sandmo, 1975; Wijkander, 1985; Glazer, 1985) or presence of complements/substitutes (Corlett and Hague, 1953; Kaplow, 2008; Atkinson and Stiglitz, 1976). There are, however, some notable differences with the present analysis. First, the goal of the literature on indirect taxation is the problem of the allocative inefficiency generated by an exogenous constraint (namely, the policymaker's need to raise revenues). Conversely, our attention is directed at defining appropriate incentive schemes to correct a consumer's bias. Second, in our model the price distortion has endogenous effects on welfare. This happens because the demand function depends not only on the price, but also on the level of effort (from $x^{inc} = V_s^{-1}(p) - y$), which in turn is affected by the price.

5.2 Incentives for participation

We now analyze the consumer's choice in $t = 0$. As discussed by Strotz (1955) and Phelps and Pollak (1968), the behavior of consumers with time-inconsistent preferences depends on their beliefs about their own future behavior. Two polar cases have been used in the literature: *sophisticated* agents have rational expectations about their future selves and therefore are fully aware of their self-control problems and correctly predict their future behavior. Conversely, *naive* agents (see also O'Donoghue and Rabin, 2001) do not recognize that they cannot make consistent plans through time and therefore believe their future selves will behave exactly according to their long-run preferences. The consumers' unawareness is motivated by the experimental evidence on overconfidence about positive personal attributes

(Larwood and Whittaker, 1977; Svenson, 1981) and is consistent with field evidence on investment (Madrian and Shea, 2001), task completion (Ariely and Wertenbroch, 2002), and health club attendance (Della Vigna and Malmendier, 2004).

Given that sophisticated and naive consumers differ for their expectations about their future behavior, we denote by $\hat{\beta} \in [\beta, 1]$ the consumer's expectation in $t = 0$ about the discount factor β that her future self will apply in $t = 1$. The higher is $\hat{\beta}$, the higher is the consumers' degree of naivety. In fact, a sophIsticated consumer is fully aware of her time-inconsistency problem, she perfectly anticipates her future behavior and that her short-run discount factor is $\hat{\beta} = \beta$. Conversely, a naive consumer over-estimates her self-restraint and expects her short-run discount factor to be $\hat{\beta} > \beta$, i.e. a naive consumer believes that she will have a more time-consistent behavior than what in reality will be.

The belief $\hat{\beta}$ regarding her own myopia plays a crucial role into the consumer's decision in $t = 0$ about whether to participate. In fact, she will pay k only if savings are high enough, i.e. only if she expects to exert sufficient effort in the future to cover the entry cost k.

In $t = 0$, the consumer expects that in the next period her effort will be $y_0^{inc} = d_y^{-1}(\hat{\beta}\delta p)$. Note that a higher price improves the consumer behavior in $t = 1$ as she will raise her effort. Note also that $y_0^{inc} \neq y^{inc}$: the former is the effort that in $t = 0$ the consumer believes she will exert in $t = 1$, while the latter is the effort that she will actually exert.

Given a generic price p, in $t = 0$ the consumer invests k only if

$$\beta \left[\delta^2 \left(V(s^{inc}) - p \left(s^{inc} - y_0^{inc} \right) \right) - \delta d(y_0^{inc}) \right] - k \geq \beta \left[\delta^2 \left(V(s^{inc}) - p s^{inc} \right) \right]. \tag{14}$$

By straightforward simplifications, condition (14) can be rewritten as

$$k \leq \epsilon^{inc}, \tag{15}$$

where $\epsilon^{inc} = \beta \left(\delta^2 p y_0^{inc} - \delta d(y_0^{inc}) \right)$.

It is worthwhile to analyse the effect of the price on ϵ^{inc}. We have that

$$\frac{\partial \epsilon^{inc}}{\partial p} = \beta \left(\delta^2 y_0^{inc} + \delta^2 p \frac{\partial y_0^{inc}}{\partial p} - \delta d_y(y_0^{inc}) \frac{\partial y_0^{inc}}{\partial p} \right).$$

As $d_y(y_0^{inc}) = \hat{\beta}\delta p$ from the FOC, the previous expression can be rewritten as

$$\frac{\partial \epsilon^{inc}}{\partial p} = \beta \left(\delta^2 y_0^{inc} + \delta^2 p \frac{\partial y_0^{inc}}{\partial p} (1 - \hat{\beta}) \right) > 0.$$

A higher price increases the probability that the consumer participates, by shifting the position ϵ^{inc} of the marginal consumer, i.e. the one whose cost k makes her indifferent between participating or not.

In this framework the benefits of raising the price are twofold. First, the price has a direct positive effect on the effort y^{inc}, which descends from the fact that the price determines the value py of the quantity saved. Second, the price has a positive indirect effect on the consumer's participation in $t = 0$, which descends from the fact that the higher expected effort y_0^{inc} makes it worthwhile sunking the entry cost k.

5.3 The optimal price

Given a price p, the welfare is given by

$$\begin{aligned} W &= \int_0^{\epsilon^{inc}} \left[\delta^2 \left[V(s^{inc}) - c\left(s^{inc} - y^{inc}\right)\right] - \delta d(y^{inc}) - k\right] dk \\ &+ \int_{\epsilon^{inc}}^{K} \left[\delta^2 \left(V(s^{inc}) - cs^{inc}\right)\right] dk. \end{aligned} \tag{16}$$

A quick glance to the welfare function in (16) suggests that the incentives on the behavior in the short run (aimed at increasing y^{inc}) and long run (aimed at increasing ϵ^{inc}) interact. Intuitively, we can exert effort in the short run only if we decided to participate in the long run; moreover, we participate in the long run only if we anticipate that we will exert effort later on. Then, the optimal price needs to account for the intertemporal externality created by each type of incentive.

Before exploring this interaction, we characterize the optimal price.

Proposition 2 *With time-inconsistent consumers, a solution to the welfare maximization problem exists. Moreover, it entails $p^{inc} \geq c$ for all $\beta \in [0, 1]$.*

Proof. See Appendix 2. ■

From Proposition 2, when the consumer is myopic, a price higher than the marginal (social) cost helps to correct the misbehavior by raising the benefits of the good behavior, i.e. by raising the value of the savings allowed by it. A price higher than the marginal social cost is needed both to induce effort, and to provide the incentive to participate.

To better understand the forces at play on determining the level of the optimal price incentives, let us express the FOC of the welfare function in (16) w.r.t. p. We write it below, and for convenience we denote the second and third term by Λ_1 and $\Lambda_2\Lambda_3$:

$$\underbrace{\delta^2 K \frac{p-c}{V_{ss}(s^{inc})}}_{DWL} + \underbrace{\epsilon^{inc}\left(\delta^2 c - \delta d_y(y^{inc})\right)\frac{\partial y^{inc}}{\partial p}}_{\Lambda_1} + \underbrace{\left(\delta^2 c y^{inc} - \delta d(y^{inc}) - \epsilon^{inc}\right)}_{\Lambda_2}\underbrace{\frac{\partial \epsilon^{inc}}{\partial p}}_{\Lambda_3} = 0. \tag{17}$$

Condition (17) presents the marginal social benefits and the marginal social cost of raising the price above c. When $p^{inc} > c$, the first term, $K\delta^2 \frac{p-c}{V_{ss}}$, is negative due to $V_{ss} < 0$;

it represents the marginal social cost caused by the price increase (i.e., the deadweight loss). In fact, a price $p^{inc} > c$ lowers the surplus of all K types of consumers due to a reduction of the level of service by $1/V_{ss}$; such cost occurs in $t = 2$ and is thus discounted by δ^2.

The second and third terms in (17) represent the marginal benefit of increasing the price stemming from the higher effort and participation.

In particular, we see from (17) that the marginal benefit of a price increase above c can be decomposed into three components, respectively related to the variation of effort, the value of participation and the variation of participation. We discuss such components below.

The benefit of higher effort (Λ_1). The first component, $\Lambda_1 = \epsilon^{inc} \left(\delta^2 c - \delta d_y(y^{inc})\right) \frac{\partial y^{inc}}{\partial p}$, constitutes the marginal benefit due to the increase of effort. From the expression of Λ_1, this depends on: i) the extent to which the effort increases due to the higher price (expressed by the factor $\frac{\partial y^{inc}}{\partial p}$), ii) the marginal value of raising effort (expressed by the factor $\delta^2 c - \delta d_y(y^{inc})$) and iii) the probability to exert effort (namely, ϵ^{inc}), as the value of effort can be appropriated only if the consumer participated in $t = 0$, which occurs for the range ϵ^{inc} of different consumer' types.

A higher price leads the current base ϵ^{inc} of consumer's types who participated at stage 0 to increase their effort in $t = 1$ by $\frac{\partial y^{inc}}{\partial p}$, gaining from each additional unit of effort the amount $\delta c y^{inc} - d(y^{inc})$.

The value of participation (Λ_2). The second marginal benefit is related to the term $\Lambda_2 = \delta^2 c y^{inc} - \delta d(y^{inc}) - \epsilon^{inc}$ in (17) and it concerns the value of participation for the marginal consumer, i.e. for the consumer who decides to participate only after the price increase. In general, the social present value of participation for a consumer is given by $\delta^2 c y^{inc} - \delta d(y^{inc}) - k$, i.e. by the present social value of the quantity saved minus the present value of the effort cost and the participation cost. Given that the marginal consumer (indifferent between participating or not) has $k = \epsilon^{inc}$, then Λ_2 represents the social value of participation for the marginal consumer.

The variation of participation (Λ_3). The third marginal benefit is represented by the term $\Lambda_3 = \frac{\partial \epsilon^{inc}}{\partial p}$ in condition (17) and it is related to the effect of the price on the extent to which consumers commit. A higher price shifts the position of the marginal consumer towards the right by $\frac{\partial \epsilon^{inc}}{\partial p}$, increasing the consumer's participation: the price works as an incentive to induce the consumer to participate.

Then, the marginal social benefiit from a price increase, $\Lambda_1 + \Lambda_2 \Lambda_3$, is given by the value generated by the higher effort for the current base of consumer, plus the value of inducing new consumers to participate (and hence exert effort), multiplied by the 'number'of new consumers' types who decide to participate.

The following section addresses the effect of consumer's myopia on these three components, in order to understand how the optimal price depends on the level of the consumer's present-bias.

6 The effect of consumer's myopia on the socially optimal price

The following Corollaries clarify how the optimal price depends on the consumer's myopia.

Corollary 1 *The optimal price is* $p^{inc} = c$ *for* $\beta = 1$ *or* $\beta = 0$*, and* $p^{inc} > c$ *for all* $\beta \in (0, 1)$.

Corollary 2 *The optimal price is maximized for* $\beta \in (0, 1)$.

Proof. See Appendix 2. ■

Interestingly, as highlighted by Corollaries 1 and 2, the optimal price distortion has a maximum for internal levels of β, while it is zero in the two extreme cases of $\beta = 1$ (fully rational) or $\beta = 0$ (fully myopic) consumers.

Corollaries 1 and 2 have stark policy implications. The social planner should give up trying to prevent misbehavior by financial instruments in both the cases of very high and very low present bias, even if all consumers are equally affected by myopia. Financial incentives are best employed when consumers display midway levels of present bias, neither too high nor too low.

Note that the non-monotonicity of the optimal level of incentives is specific for this framework in which the consumer's present bias affects a behavior, rather than a consumption decision. Indeed, in the case of overconsumption of a sin good by a population of identically myopic individuals, the optimal price is higher, the higher is the consumer's myopia and no welfare loss occurs, as the consumption reverts to first best levels.

The fact that in our framework the optimal price distortion is non-monotonic with respect to the degree of myopia is clearly the result of a trade-off between social cost of incentives and their effectiveness on behavior. The following discussion tries to provide a better understanding of such trade-off and of the role played on it by the interaction between short run and long run incentives.

As $s^{inc} = V_s^{-1}(p)$ does not depend on β, then the deadweight loss (the term DWL in (17)) is constant regardless of the level of consumer's myopia, given the price. Then, to understand the role of β on the optimal price, we need to look at the other terms, $\Lambda_1+\Lambda_2\Lambda_3$, which are a function of β through y^{inc} and ϵ^{inc}. We analyze them separately in the discussion that follows.

6.1 β and Λ_1: how myopia affects the design of short run incentives

We analyze now how the three factors in $\Lambda_1 = \epsilon^{inc}\left(\delta^2 c - \delta d_y(y^{inc})\right)\frac{\partial y^{inc}}{\partial p}$ depend on β.

The last factor, $\frac{\partial y^{inc}}{\partial p} = \frac{\beta\delta}{d_{yy}} > 0$, is the marginal increase of y^{inc} w.r.t. the price and it is positive and increasing in β: the higher is the consumer's myopia (i.e., β is close to 0), the less effective is a price increase in inducing short run behavioral change. Myopic consumers are harder to be induced to exert effort because they myopically disregard the savings obtained by higher effort as they will only be obtained in the future.

The second factor is the marginal variation of the value of effort following a price increase, i.e. $\frac{\partial}{\partial p}(\delta c y^{inc} - d(y^{inc})) = \delta^2 c - \delta d_y(y^{inc})$. Due to a raise in price, the consumer exerts more effort y^{inc}. However, the value of the marginal effort depends on the initial level of y^{inc}. In particular, the effort has a linear value $\delta c y$ but a convex cost $d(y)$. Then, $\delta^2 c - \delta d_y(y^{inc})$ is larger when y^{inc} is low, which in turn occurs when myopia is high. This means that improving the effort has a higher marginal value, the higher is the initial misbehavior, i.e. consumers have a high myopia. In other words, the gain of raising the price in terms of better behavior is larger for high levels of myopia.

The factor ϵ^{inc} represents the ex-ante probability to participate, and it is lower, the higher is the level of myopia. Then, the myopia decreases the share of consumer's types affected by the pricing policy on the effort.

In conclusion, when we study how β affects the marginal benefit of the price in terms of higher effort (short run benefits), we find that opposite effects are at play.

On the one hand, the problem of under-exertion of effort is more severe the higher is the degree of myopia. In fact, a higher myopia entails lower effort, i.e. a lower marginal cost of effort due to the convexity of its cost function. Then, the value of any additional effort is higher: increasing effort is socially more valuable the more myopic consumers are.

On the other hand, the higher is the myopia, the less effective is a policy aimed at increasing the effort through price, for two reasons. First, when myopia is high few consumers participated in the first place (i.e. only those consumers with very low participation costs), so that the value of the additional effort can be appropriated only by a small fraction of consumers. In other words, when myopia is high, a policy targeting the effects of misbehavior in the short run (effort) is partially thwarted by the fact that it works only on the small subset of consumers who participate in the long run. Then, any policy directed at providing long run incentives has a positive externality on the impact of policies targeting short run measures. Second, myopic consumers are also more hesitant to respond to the incentives to exert effort, and thus require a higher price distortion to produce any benefit.

Then, when myopia is high, the additional gain in terms of value of effort is large, but the price is scarcely effective to induce behavioral change and the policy can only reach a small fraction of consumers in any case. The opposite is true when myopia is low.

Overall, the divergent effects explain the non-monotonicity of the optimal price in regards to the incentives on effort. We now look at the incentives on participation.

6.2 β and $\Lambda_2\Lambda_3$: how myopia affects the design of long run incentives

We already mentioned that the terms $\Lambda_2 = \delta^2 c y^{inc} - \delta d(y^{inc}) - \epsilon^{inc}$ and $\Lambda_3 = \frac{\partial \epsilon^{inc}}{\partial p}$ in (17) are related to the marginal benefit of the price on participation. We study the two terms separately.

Effect of myopia on the value of participation. The effect of β on Λ_2 is not monotonic. Λ_2 is decreasing in ϵ^{inc} but increasing in y^{inc}. A higher myopia decreases both ϵ^{inc} and y^{inc}: the marginal consumer exerts less effort, but she also has lower entry costs k. In particular, when β is close to 0 (high myopia), both ϵ^{inc} and y^{inc} go to 0 and Λ_2 is close to 0 as well. Conversely, when β is close to 1 (low myopia), y^{inc} and ϵ^{inc} go respectively to y^* and ϵ^*, therefore Λ_2 is close to 0 (recall that $\epsilon^* = \delta^2 c y^* - \delta d(y^*)$). We provide now both the technical and the intutive reason for the non-monotonicity of Λ_2. Technically, the root of the non-monotonicity of Λ_2 is the fact that ϵ^{inc} increases linearly in β, while y^{inc} increases at decreasing rates (due to the convex effort cost function $d(y)$). For sufficiently high levels of myopia (i.e., β close to 0), if β increases, then both y^{inc} and ϵ^{inc} increase; however, the increase of y^{inc} is steeper than that ϵ^{inc}. Then, the marginal value of participation initially increases. However, as β keeps increasing, the increase of y^{inc} slows down while the increase of ϵ^{inc} proceeds at the same constant rate: the same raise in price is able to induce the participation of a consumer who has a linearly higher participation cost and a value from effort which increases, but at lower and lower rates. When sufficiently low levels of myopia are reached, however, participation costs increase at a higher rate than the value of effort, so that the marginal value of participation start decreasing.

More on an intuitive level, when we are strongly present biased, inducing commitment might have little value because we will later exert little effort. On the other hand, when we have a small present bias, the new consumer who is induced to participate has little commitment value because this individual has a participation cost that is very close to its benefit, i.e. her net surplus from participation would be low. To make an example, it is a waste of social resources to convince new people to join a wellness program both when they are very myopic, because they will not attend to it, and when they have low myopia, because only people with high participation cost are not applying yet, and convincing these people to participate would raise low value.

The non-monotonicity of Λ_2 allows an unexpected insight. Out of instinct, we might expect that the more myopic we are, the worse is the time-inconsistency problem and the higher is the social benefit of correcting it by making consumers participate, namely, the higher is the social benefit of participation. Actually, the non-monotonic shape of Λ_2 tells another story. If we are very myopic, the benefit of participating in $t = 0$ is neutralized by the fact that we will refrain from effort in $t = 1$. Importantly, note that it is true whatever is the policy effectiveness in convincing new consumers to participate (i.e., Λ_3): even if a raise in price would be able to significantly expand the consumers' base (because, for example, the consumer naively expects to exert future effort), the value of participation would still be low as nobody would actually exert effort in the short run. From another perspective, the interaction between the short and long run efforts dissipates the benefits of a policy directed at providing long run incentives, both when the present bias is sufficiently strong and when it is weak, and interestingly this happens *despite* the complementarity between efforts. This dissipation makes such a policy not worthwhile, regardless of how difficult it is to induce consumers to participate. Naturally, the same is true also for the opposite direction: a policy aimed at raising effort in the short run may undermine the incentives to participate by a consumer who is only weakly myopic. This is because effort costs increase exponentially, and are thus relatively high for a nearly-rational consumer who already exerts a high (though still suboptimal) effort. Then, the marginal benefit of participation is lower.

Effect of myopia on the variation of participation. Recall the definition of Λ_3,

$$\Lambda_3 = \frac{\partial \epsilon^{inc}}{\partial p} = \beta \left(\delta^2 y_0^{inc} + \delta^2 p \frac{\partial y_0^{inc}}{\partial p} (1 - \hat{\beta}) \right),$$

which shows that Λ_3 is clearly increasing in β. The price works as an incentive to induce the consumer to participate, but its effectiveness decreases for higher levels of myopia. In fact, a consumer with a strong present-bias disregards the benefits of the possibily higher future effort, exactly because they are in the future. The more myopic consumers are, the harder is to convince them to participate.

To summarize, when we study how β affects the marginal benefit of the price in terms of higher participate (long run benefits), we find that opposite effects are at play.

On the one hand, consumers with a high level of myopia are more unwilling to participate, i.e. incentives aimed at inducing participation are less effective. Moreover, they will later exert lower effort, so that a policymaker might be unwilling to implement costly policies directed at providing long run incentives. On the other hand, they also have lower participation costs.

Also in the case of long run incentives, the benefit of implementing a policy targeting the effects of misbehavior in the long run (participation) crucially depends on the effectiveness of short run policies. However, differently from the case of short run incentives, here is no longer clear whether the externalities produced by short run policies on the effectiveness of long run ones are positive of negative, due to the non monotonicity of Λ_2. For high levels of myopia, a short run policy targeting effort causes a positive externality on the effectiveness of a long run policy, because increasing the effort increases the value of participation (due to the steep increase of the effort value $\delta^2 c y^{inc} - \delta d(y^{inc})$ and the lower increase of the participation cost). For low levels of myopia, a short run policy targeting effort causes a negative externality on the effectiveness of long run policies, because increasing the effort decreases the value of participation (as the participation cost increases faster than the effort value $\delta^2 c y^{inc} - \delta d(y^{inc})$): in this case, policies directed at increasing participation are partially thwarted by the low value originated by short run policies.

Despite the complementarity between long and short run effort, the relationship between long and short run policies is non-trivial, and when myopia is low, a policy directed at short run effort might reduce the value of policies targeting participation.

The next section delves more deeply into the impact of the consumer's awareness on the incentives on behavior and participation.

7 The effect of consumer's awareness on the incentives

Similarly to the previous analysis, in this section we study how a change in the consumer's awareness $\hat{\beta}$ affects the terms in (17).

The benefit of higher effort (Λ_1). The extent of this benefit depends on the degree of naivety of the consumer. While the value of the additional effort $(\delta^2 c - \delta d_y(y^{inc})) \frac{\partial y^{inc}}{\partial p}$ does not depend on $\hat{\beta}$ (i.e., in $t = 1$ we will exert the same effort regardless of our naivety), the size of the customer base ϵ^{inc} does. In particular,

$$\frac{\partial \epsilon^{inc}}{\partial \hat{\beta}} = \frac{\beta \delta^3 p^2}{d_{yy}(\hat{\beta}\delta p)}(1 - \hat{\beta}) > 0 : \tag{18}$$

a higher degree of naivety increases the probability that the consumer participates in stage 0.

Condition (18) shows that the more we expect to behave rationally in the future, the less severe is the participation problem, and hence the higher is the consumer's base that can appropriate of the benefits of higher effort. In $t = 0$, a sophisticated consumer (who has a smaller $\hat{\beta}$) anticipates that she will later under-perform in terms of effort. Indeed, a sophisticated consumer not only discounts excessively the future benefit of participation (like a more naive consumer), but she is also aware of her self control problem in the next period (differently from a naive one), and correctly anticipates a lower benefit because of the underperformance in effort. Therefore, the participation problem is less severe the higher is the degree of naivety, and sophistication worsens the inefficiency[15].

In this regard, a price increase causes the consumer to exert the same additional effort, regardless of her being naive or sophisticated. However, a naive consumer is more likely to participate than a sophisticated one, so that the benefits from the additional effort are appropriated with higher probability. It follows that the effort-related marginal benefit from a price increase is higher for a naive than for a sophisticated consumer, and this implies a higher optimal price the higher is $\hat{\beta}$.

The value of participation (Λ_2). The consumer's awareness affects Λ_2 through ϵ^{inc}. Recall that $\frac{\partial \epsilon^{inc}}{\partial \hat{\beta}} > 0$: the more sophisticated the consumer is, the lower is ϵ^{inc}, i.e. the marginal consumer has a lower entry cost k and the actual social benefit from inducing her to participate is higher. As a consequence, the value of participation is higher for a sophisticated consumer than for a naive one. Thus, shifting the marginal consumer provides a higher social value when the consumer is sophisticated rather than naive.

The variation of participation (Λ_3). The effectiveness of the price in inducing participation, $\frac{\partial \epsilon^{inc}}{\partial p}$, depends on her degree of naivety. Whether the relationship between $\hat{\beta}$ and $\frac{\partial \epsilon^{inc}}{\partial p}$ is positive or negative is, however, an empirical matter. Indeed, it results

$$\frac{\partial}{\partial \hat{\beta}} \left(\frac{\partial \epsilon^{inc}}{\partial p} \right) = \frac{\beta \delta^3 p}{d_{yy}(\hat{\beta}\delta p)} \left(1 - \hat{\beta}\right) \left(2 - \hat{\beta}\delta p \frac{d_{yyy}(\hat{\beta}\delta p)}{(d_{yy}(\hat{\beta}\delta p))^2}\right), \tag{19}$$

whose value can be positive or negative, depending on the curvature of the effort cost function. A higher price increases the expected effort y_0^{inc}, and this has both benefits and costs. On the one hand, the expected savings py_0^{inc} are higher (both for the higher price and for

[15]In other terms, if we naively overestimate our willpower to exert effort, we are more likely to sunk the cost of signing up to a gym, applying for a master program or retrofitting our house, even if later we will underperform.

the higher effort). This effect is unambiguously stronger for naive as they have a higher y_0^{inc}. On the other hand, a higher effort implies a higher effort cost. If these effort costs increase rapidly (d_{yyy} is large), naive consumers, who expect to exert more effort than sophisticated ones, have a lower net expected benefit **from the effort**. This means that, when the price increases, they do increase their participation, but they might do so more 'slowly' than sophisticated consumers.

Conversely, if d_{yyy} is small enough, the value of (19) is positive, meaning that the increase of $\hat{\beta}$ increases the rate with which incentives increase participation, i.e. $\partial\epsilon^{inc}/\partial p$. In this case, a higher price moves the position of the marginal consumer to the right, reducing the gap $\epsilon^* - \epsilon^{inc}$ (i.e., the participation inefficiency), and this effect is more pronounced the higher is $\hat{\beta}$. If effort costs are not too steep, an increase of the price is more effective in inducing the consumer to participate, the higher is the consumers' naivety. The higher the effectiveness of a price increase in targeting the participation problem suggests that the optimal price should be higher the more naive the consumers are.

7.1 A synthesis

When we look at the impact of $\hat{\beta}$ on the overall marginal benefit, we find once again that the aforementioned three effects work in opposite directions.

On the one hand, when consumers are naive, they overestimate their willingness to engage in quantity-saving activities later on. Then, they are more likely to sunk the cost that allows them to participate in them. As a consequence, raising the price causes a higher effort by a customer base that is larger the more naive consumers are.

On the other hand, the social gain of engaging an additional naive consumer is lower than for a sophisticated one, as naive consumers, being in general more willing to participate, have higher participation costs.

Finally, a third effect, which can go both ways, is related to the effectiveness of the price on raising the probability of participating. If effort costs do not increase too rapidly, a unit raise of the price increases the probability with which the consumer participates to a larger extent the more naive the consumer is. Conversely, if the rate at which effort costs increase becomes higher and higher, the price is more effective in convincing sophisticated consumers to participate, who anticipate to exert a lower effort than naive ones, and thus expect to have much lower effort costs (i.e., a larger benefit).

The above intuitions suggest that the optimal price might not be monotonic in the degree of consumers' naivety. The following proposition formalizes this results and provides the conditions under which one of the effects prevails.

Proposition 3 *If β is sufficiently low (high), the optimal price is higher (lower) for naive consumers than for sophisticated ones.*

Proof. See Appendix 2. ■

Proposition 3 says that when the consumer has a strong present bias, the optimal price is higher for a naive consumer than for a sophisticated one. Conversely, when the consumer's myopia is low, sophisticated consumers should be given higher price than naives.

When the consumer suffers from a high myopia, she foregoes to participate for two reasons. First, she discounts excessively the value of the future quantities saved in $t = 2$, by applying a discount factor $\beta < 1$; this effect is the same for sophisticated and naives (as they apply the same β to the flows of $t = 2$). Second, the value of the quantities saved is actually low because in $t = 1$ she exerts an effort below the first best. This second effect impacts differently naive and sophisticated, as only sophisticated consumers realize it in $t = 0$. As a consequence, when the consumer's myopia is high, a naive consumer, who overlooks the second effect, is slightly easier to be induced to participate by a price increase, i.e. raising the price has a larger marginal social benefit. This means that the social planner might want to face the welfare loss of a price distortion more, the more naive consumers are.

When the consumer's myopia is low, i.e. she is very nearly rational, the difference between naive and sophisticated consumers in terms of effort is negligible.

For high level of sophistication, the underparticipation problem is very severe. In fact, consumers anticipate their future under-exertion of effort, and this, added to their myopia, further decreases the gain from participation. The participation gap is so severe that even consumers with a low participation cost k fail to invest it, i.e. the participation is very low. Moreover, since sophisticated consumers anticipate their future low effort, they also predict a lower effectiveness of the investment. In this situation, the price is not very successful to convince new consumers to participate (due to its lower effectiveness), but on the other hand the few consumers who can be convinced obtain very large gains, due to their low participation cost. The latter effect has a first order impact, so that when consumers become marginally more naive, the price must be increased to fully exploit it.

In sum, when consumers are sophisticated, it is very difficult but also very rewarding to be able to convince them to participate. It follows that the more naive the consumer is, the easier is to convince her to participate, and thus the higher must be the price. However, when they become very naive, the value of participation for each new consumer becomes lower, and thus the price must be reduced.

7.2 An illustrative example

To gain a better insight on the implications of the consumer's awareness on the optimal price, we conduct here a numerical and graphical analysis.

To this aim, let us assume $d(y) = y^2/2$. Then, in $t = 1$ the consumer is exerting the effort $y^{inc} = \beta\delta p$ from (13), but in $t = 0$ she expects that she will exert $y_0^{inc} = \hat{\beta}\delta p$. Marginal cost of production of x are assumed to be constant and equal to 10.

Figure 1 and 2 show the optimal price incentive (on the vertical axis) as a function of the level of myopia β. The two curves represent the optimal price for two levels of consumer's awareness: a fully naive consumer ($\hat{\beta} = 1$) and a fully sophisticated one ($\hat{\beta} = \beta$). In Figure 1, the elasticity of demand, represented by the parameter V_{ss}, is assumed constant, and its absolute value decreases (i.e., the demand becomes more elastic) from panel (a) ($V_{ss} = -50$) to panel (d) ($V_{ss} = -1$). In Figure 2, we generalize the demand function by assuming a radical –panel (a)– and an exponential –panel (b)– $V(s)$ curve.

By observing the results of the numerical analysis, some insights emerge.

First, the optimal price is higher, the less elastic the demand is. This can be clearly seen by the scale of the optimal price on the vertical axis. The intuition of this result is straighforward: with an inelastic demand, a price increase causes a relatively low deadweight loss. As a consequence, the social planner is more willing to raise the price than when the demand is elastic.

Second, the optimal price is always greater than the marginal cost, and it converges to the marginal cost both when the consumer is very myopic (β goes to 0), and when she is nearly rational (β goes to 1), in line with the results of Proposition 2.

Third, the function of the optimal price is single-peaked, and maximized for $\beta \in (0, 1)$, as postulated in Proposition 2.

Fourth, when the consumer's myopia is sufficiently high, the optimal price is higher when the consumer is naive. Conversely, when the myopia is low, the optimal price is higher for sophisticated consumers than for naive ones.

For a better intuition of the above result, it is convenient to look at Figure 3, where the three marginal benefits, Λ_1, Λ_2 and Λ_3 discussed in the previous sections are presented.

We know that the marginal benefit of increasing the price is $\Lambda = \Lambda_1 + \Lambda_2\Lambda_3$. Each of the three effects is represented in Figure 3 as a function of β, on the horizontal axis: Λ_1 is the curve in red, Λ_2 is the blue curve and Λ_3 is the orange one. Solid lines refer to a fully naive consumer ($\hat{\beta} = 1$), while dashed lines to a sophisticated one ($\hat{\beta} = \beta$).

While Λ_1 is always higher for naive consumers, Λ_2 is higher for sophisticated ones for all β. In fact, Λ_1 reflects the fact that naive consumers are more likely to commit, and hence

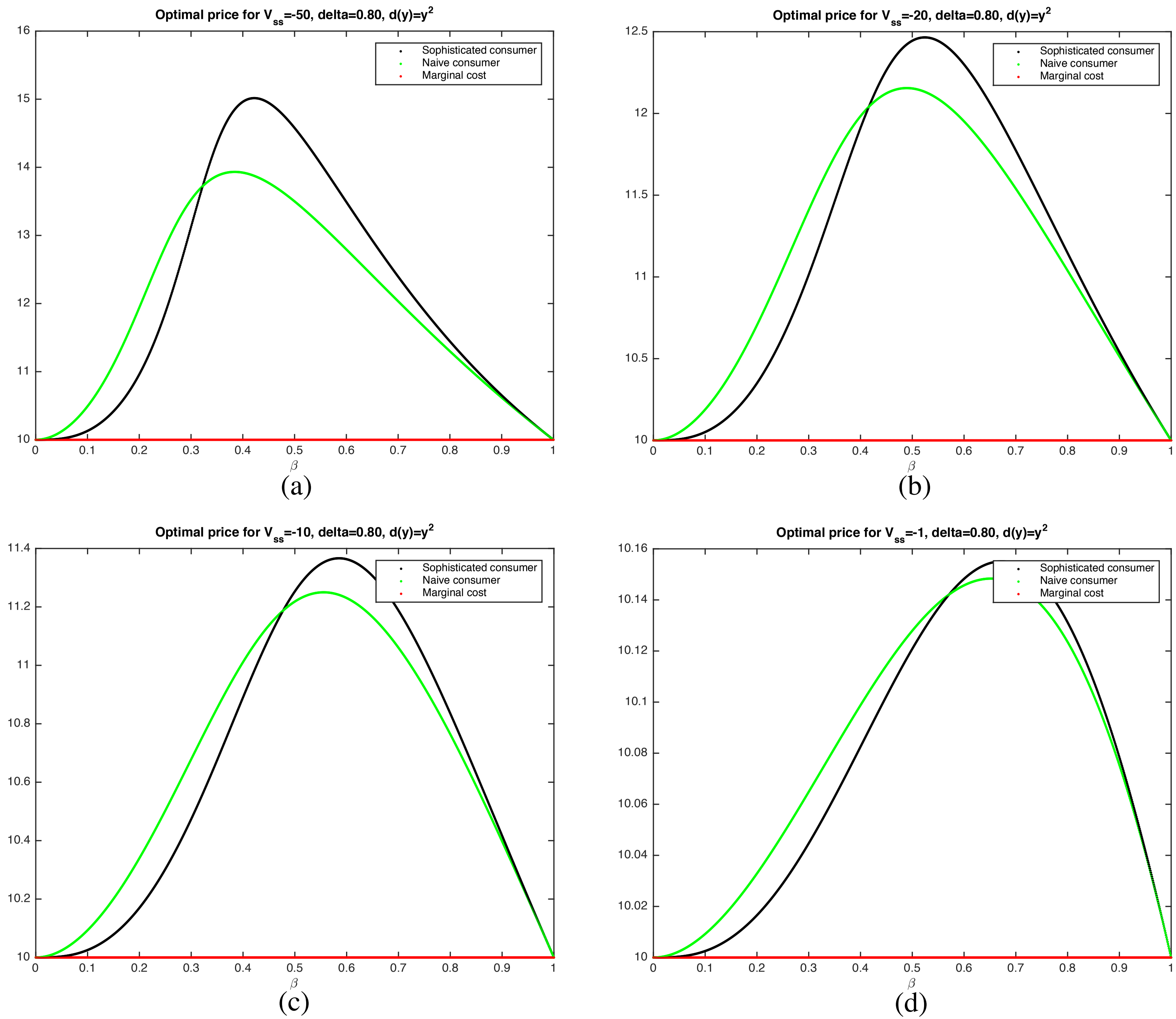
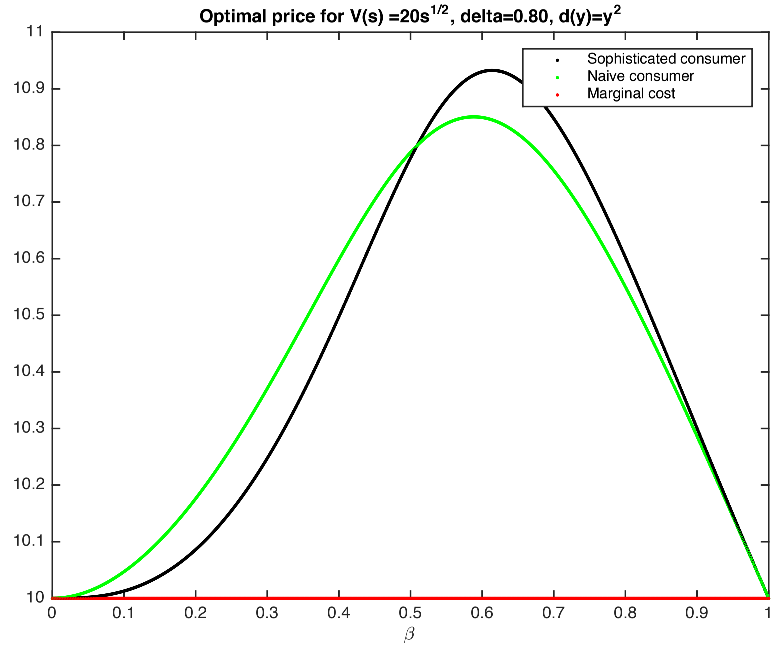
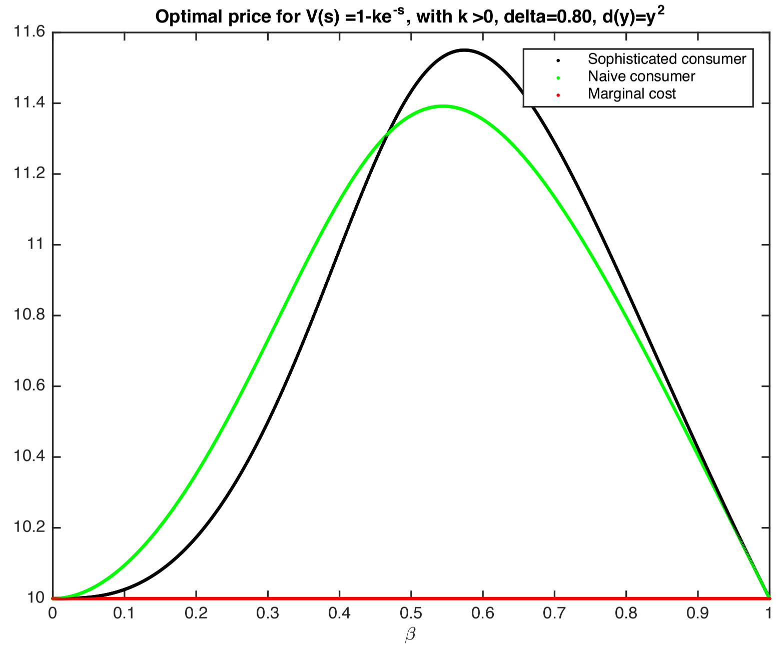


Figure 1: Optimal price for a naive and sophisticated consumer, quadratic utility functions with constant second derivative. Marginal cost equal to 10.



(a)



(b)

Figure 2: Optimal price for a naive and sophisticated consumer, radical (a) and exponential (b) utility functions. Marginal cost equal to 10.

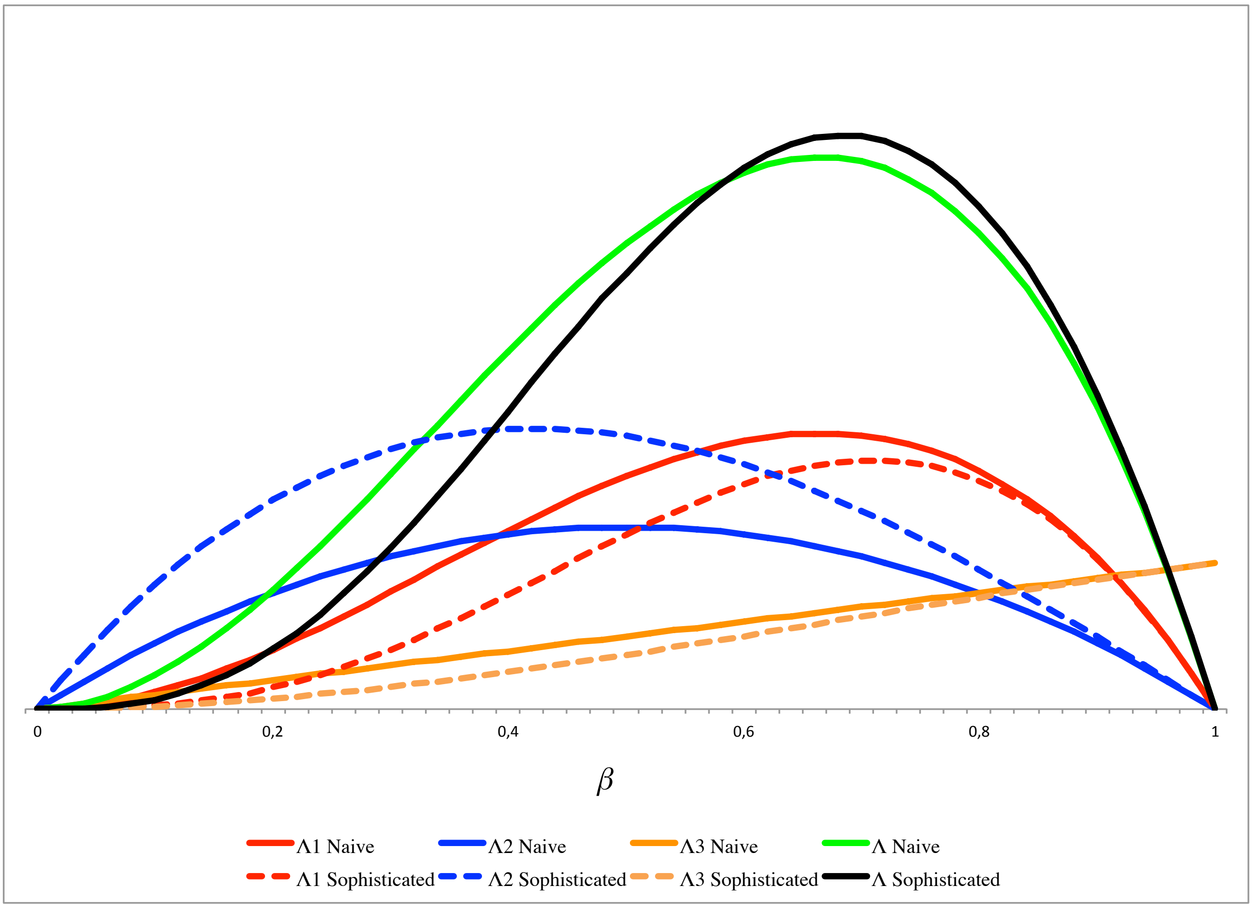


Figure 3: Marginal benefits of a price increase, for a naive and a sophisticated consumer

are more likely to attain the benefits of higher effort. Conversely, Λ_2 reflects the fact that sophisticated consumers have a more severe undercommitment problem, so that inducing them to commit is socially more valuable.

Recall, however, that Λ_2 must be multiplied by Λ_3 (i.e., the value of commitment, Λ_2, can be apppropriated only if the consumer actually commits). In turn, Λ_3 is close to zero for very low levels of β, and then steadily increases. As a result, the advantage that sophisticated consumers have over naive ones in terms of marginal benefits is weakened for low β, and this explains why when myopia is high, the price is consistently higher for naive consumers than for sophisticated ones.

The intuition for this result is the following. If the myopia is sufficiently severe, it becomes very difficult to induce consumers to increase their commitment (Λ_3 is small), because price incentives are only rewarded in the future. Then, it is immaterial that sophisticated consumers would be more valuable if they committed than naive ones, because consumers are simply unwilling to increase their commitment (sophisticated even more so because they anticipate the future under-effort). What really matters is only that naives do commit more, seduced by the illusion that they will exploit their initial investment. Hence, they are more likely to appropriate the benefits of the higher effort, stimulated by the raise in price. For this reason, the price is higher for naive consumers when myopia is high.

If the myopia is low (β is close to 1), Λ_3 is high: the price is very effective in inducing commitment, because we know that we will exploit the sunk cost of commitment thanks to the high effort. Then, the weight of Λ_2 over Λ_1 increases, i.e. the marginal benefit of raising the price is higher for sophisticated consumers. In fact, if it is easy to increase commitment, then the value of commitment becomes crucially important, and the value of commitment is higher for sophisticated consumers, who suffer from a more severe undercommitment inefficiency.

8 Conclusion

The literature on consumer's present bias is typically concerned with the effects of people's myopic decisions on consumption. We contribute to this literature by modelling myopic behavior decisions which involve a distortion in an effort, rather than in the demand for a good. This analysis provides an interpretation to the empirical puzzle on the low effectiveness of financial incentives in the long run.

We show that a large incentive provided to correct a severe consumer's misbehavior might be suboptimal if we don't account for the degree of myopia and awareness of the individual. Specifically, we find that financial incentives should be larger for intermediate levels of myopia, while they shrink to zero both for very high or very low levels of myopia.

The non-monotonic profile of the socially optimal level of financial incentives arises from the trade-off between the value of incentives and their effectiveness. A very myopic consumer obtains a larger benefit from correcting her behavior, but she is also more difficult to convince, because she myopically disregards the future benefits of effort. Conversely, a consumer with very low myopia has limited advantages from improving her behavior, but incentives are most effective because she recognizes its importance.

Moreover, we find that in a long run perspective this trade-off is complicated by the interplay of the levels of effort exerted at different stages. In particular, if the consumer's myopia is low, an increase of effort in later periods influences negatively the effectiveness of incentives on behavior at earlier stages, such as participation decisions.

Consumers' awareness of their own myopia also influences the size of financial incentives. One the one hand, raising the incentive is less effective to induce sophisticated consumers to improve their behavior. In fact, sophisticated consumers anticipate that they will misbehave in the future, and thus are not going to exploit their present effort, so that there is no point in exerting it. On the other hand, the misbehavior problem is more severe for sophisticated consumers, so that improving their behavior is more valuable from a social point of view.

This paper will hopefully provide useful insights for policymakers and for future empirical reasearch testing the effect of incentives aimed at correcting consumers' behavior.

References

[1] Ashram, N., Karlan, D. and Wesley, Y., 2006. Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines. *Quarterly Journal of Economics*, 121(2), 635-672.

[2] Ariely, D., and K. Wertenbroch, 2002. Procrastination, Deadlines, and Performance: Self-Control by Precommitment. *Psychological Science*, XIII: 219-224.

[3] Atkinson, A.B. and Stiglitz, J.E., 1972. The structure of indirect taxation and economic efficiency. *Journal of Public Economics*, 1(1): 97-119.

[4] Atkinson, A.B. and Stiglitz, J.E., 1976. The Design of Tax Structure: Direct versus Indirect Taxation, *Journal of Public Economics*, 6: 55-75.

[5] Augenblick, N., M. Niederle, and C. Sprenger, 2013. Working Over Time: Dynamic Inconsistency in Real Effort Tasks. NBER Working Paper No. 18734.

[6] Baron, D. P., Myerson, R.B., 1982. Regulating a Monopolist with Unknown Costs. *Econometrica*, 50(4), 911-930.

[7] Barr, S., Gilg, A.W., and Ford, N., 2005. The household energy gap: examining the divide between habitual- and purchase-related conservation behaviors. *Energy Policy*, 33(11), 1425-1444.

[8] Baumol, W.J. and Bradford, D.F., 1970. Optimal Departures From Marginal Cost Pricing. *The American Economic Review*, 60(3) 265-283

[9] Bernheim, B.D., J. Meer, and N.K. Novarro, 2016. Do Consumers Exploit Precommitment Opportunities? Evidence from Natural Experiments Involving Liquor Consumption. *American Economic Journal: Economic Policy*, forthcoming.

[10] Beshears J., Choi J., Laibson D. and Madrian B., 2005. Early Decisions: A Regulatory Framework. *Swedish Economic Policy Review*, 12(2): 41-60.

[11] Carpenter, C. S., and D. Eisenberg, 2009. Effects of Sunday Sales Restrictions on Overall and Day-Specific Alcohol Consumption: Evidence From Canada. *Journal of Studies on Alcohol and Drugs*, 70(1): 126133.

[12] Chu, L. Y., Sappington, D.E.M., 2013. Motivating Energy Suppliers to Promote Energy Conservation. *Journal of Regulatory Economics*, 43(3), 229-247.

[13] Commission for Energy Regulation, 2011. Electricity Smart Metering Customer Behaviour Trials (CBT) Findings Report. CER11080a.

[14] Deci, E.L., Koestner, R., and Ryan, R.M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6): 627-668.

[15] Diamond, P. and Mirrlees, J., 1971. Optimal Taxation and Public Production II: Tax Rules. *The American Economic Review*, 61(3): 26178

[16] Dixit, A., 1970. On the Optimum Structure of Commodity Taxes. *The American Economic Review*, 60(3): 295-301

[17] Della Vigna, S., and Malmendier, U., 2004. Contract Design and Self-Control: Theory and Evidence. *Quarterly Journal of Economics*, CXIX (2), 353-402.

[18] Fuller, E., 2011. Smoking, drinking and drug use among young people in England in 2010. NHS Information Centre for Health and Social Care.

[19] Giné, X., Karlan, D., Zinman, J., 2010. Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation. *American Economic Journal: Applied Economics*, 2(4):213-35

[20] Gul, F. and Pesendorfer, W., 2001. Temptation and Self-Control. *Econometrica*, 69: 1403-1435.

[21] Haavio, M. and Kotakorpi, K., 2011. The political economy of sin taxes. *European Economic Review*, 55(4): 575-594.

[22] Hinnosaar, M., 2016. Time inconsistency and alcohol sales restrictions. *European Economic Review*, 87: 108-131,

[23] Kaiser Family Foundation Health Research and Educational Trust. Employer Health Benefits: 2013 Annual Survey.

[24] Kaur, S., Kremer, M., and Mullainathan S., 2015. Self Control at Work. *Journal of Political Economy*. 123 (6): 1227-1277.

[25] Kreps, D.M., 1997. Intrinsic Motivation and Extrinsic Incentives. *The American Economic Review*. 87 (2): 359-364.

[26] Laffont, J.J., and Tirole J., 1993. A Theory of Incentives in Procurement and Regulation *MIT Press*.

[27] Laibson, D.I., 2015. Why Don't Present-Biased Agents Make Commitments? *The American Economic Review*, 105(5): 267-72.

[28] Laibson, D.I., 1997. Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics*, CXII, 443-477.

[29] Larwood, L., and W. Whittaker, 1977. Managerial Myopia: Self-Serving Biases in Organizational Planning. *Journal of Applied Psychology*, 62(2): 194-198.

[30] Lerner, A.P., 1970. On Optimal Taxes with an Untaxable Sector. *The American Economic Review*, 60:284-94.

[31] Madrian, B.C., and D. Shea, 2001. The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior. *Quarterly Journal of Economics*, CXVI, 1149-1187.

[32] Mattke, S., H. Liu, J. Caloyeras, C.Y. Huang, K.R. Van Busum, D. Khodyakov and V. Shier, 2013. Workplace Wellness Programs Study: Final Report. Santa Monica, CA: RAND Corporation.

[33] Mattke, S., Kapinos, K., Caloyeras, J.P., Taylor, E.A., Batorsky, B., Liu, H., Van Busum, K.R. and Newberry, S., 2015. Workplace Wellness Programs. Services Offered, Participation, and Incentives. *Rand Health Quaterly*. 5(2):7.

[34] O'Donoghue, T.D., and M. Rabin, 1999a. Doing It Now or Later. *The American Economic Review*, LXXXIX, 103-124.

[35] O'Donoghue, T.D., and M. Rabin, 1999b. Incentives for Procrastinators. *Quarterly Journal of Economics*, CXIV, 769-816.

[36] O'Donoghue, T.D., and M. Rabin, 2001. Choice and Procrastination. *Quarterly Journal of Economics*, CXVI, 121-160.

[37] O'Donoghue, T.D., and M. Rabin, 2003. Studying Optimal Paternalism, illustrated by a Model of Sin Taxes. *The American Economic Review*, 93, 186-191.

[38] O'Donoghue, T.D., and M. Rabin, 2006. Optimal Sin Taxes. *Journal of Public Economics*, 90, 1825-1849.

[39] Patel MS, Asch DA, Troxel AB, Fletcher M, Osman-Koss R, Brady J, Wesby L, Hilbert V, Zhu J, Wang W, Volpp KG, 2016. Premium-Based Financial Incentives Did Not Promote Workplace Weight Loss In A 2013-15 Study. *Health Affairs* (Millwood), 35(1):71-9.

[40] Patient Protection and Affordable Care Act, 2010. Pub 111148, 124 Stat 119

[41] Persson, R.; Cleal, B.; Bihal, T.; Hansen, S. M.; Jakobsen, M. .; Villadsen, E.; Andersen, L.L., 2013. Why Do People With Suboptimal Health Avoid Health Promotion at Work? *American Journal of Health Behavior*, 37(1): 43-55(13)

[42] Phelps, E.S. and Pollak, R.A., 1968. On Second-Best National Saving and Game-Equilibrium Growth. *Review of Economic Studies*, 35(2), 185-199.

[43] Ramsey, F., 1927. A Contribution to the Theory of Taxation. *Economic Journal*, 37: 47-61.

[44] Robroek, S.J., van Lenthe, F.J., van Empelen, P. and Burdorf, A., 2009. Determinants of participation in worksite health promotion programmes: a systematic review. *International Journal of Behavioural Nutrition and Physical Activity*. 6:26.

[45] Stehr, M., 2007. The Effect of Sunday Sales Bans and Excise Taxes on Drinking and Cross-Border Shopping for Alcoholic Beverages. *National Tax Journal*, 60(1), 85.

[46] Strotz, R.H., 1955. Myopia and Inconsistency in Dynamic Utility Maximization. *Review of Economic Studies*, XXIII, 165-180.

[47] Svenson, O., 1981. Are We All Less Risky and More Skillful than Our Fellow Drivers? *Acta Psychologica*, XLVII, 143-148

Appendix 1: Complementary effort

In some cases, the consumption of a good is possible only in the face of an upfront effort. This is true, for example, any time the consumption must be preceded by investing some effort in learning activities. In these cases, the effort of learning is a necessary, indispensable asset, whose absence makes consumption impossible and thus prevents to obtain the related surplus.

To analyze this situation in the most streamlined way, we slightly adapt our basic model by assuming that the utility $V(\min\{x, y\})$ depends on the minimum between the quantity x consumed of some good and effort y exerted in $t = 1$. To model the complementarity between the consumption x and the effort, we also assume that the consumer has access to the consumption only if she invested the fixed amount k in $t = 0$, which constitutes the initial effort cost. The value of k is perfectly observed by the consumer, while the social planner only knows that it is uniformly distributed in the interval $[0, K]$, with $K > 0$.

If no effort is exerted in $t = 0$ or $t = 1$, the consumption is zero and so is the utility, which is normalized to $V(\min\{x, 0\}) = 0$.

If, on the other hand, in $t = 0$ the consumer invested in effort, in $t = 2$ she will choose the quantity so as to maximize $V(x) - px$, given a price p. Therefore, the demand for the good is given by $x = V_x^{-1}(p)$.

A myopic consumer will invest in effort if and only if

$$\beta\delta \left[V(V_x^{-1}(p)) - pV_x^{-1}(p)\right] - k \geq 0$$

Let us denote with $\epsilon^{inc} = \beta\delta \left(V(V_x^{-1}(p)) - pV_x^{-1}(p)\right)$ the marginal consumer. Note that

$$\frac{\partial \epsilon^{inc}}{\partial p} = \beta\delta \left(\frac{V_x}{V_{xx}} - \frac{p}{V_{xx}} - V_x^{-1}(p)\right) = -\beta\delta V_x^{-1}(p) < 0 \tag{20}$$

as in the optimum $V_x = p$. From (20), we find that whenever the effort is a complement, rather than a substitute, of the consumption, the price is negatively correlated with the position of the marginal consumer. A reduction of the price improves the consumer's behavior. Indeed, the effort in this case works as a sort of commitment to consume a certain amount of good in the future. Reducing the price of the latter, and thus our future expenditure, raises our willigness to commit.

The social planner maximizes

$$W = \delta \left(V(V_x^{-1}(p)) - cV_x^{-1}(p) \right) \epsilon^{inc} - \int_0^{\epsilon^{inc}} k dk, \tag{21}$$

The optimal price is obtained from the FOC of (21), which allows to obtain the result expressed by the following Proposition.

Proposition 4 *With time-inconsistent consumers, if the effort is complementary to the consumption, the optimal price is $p^{inc} = c$ for $\beta = \{0, 1\}$, and $p^{inc} < c$ for all $\beta \in (0, 1)$.*

Proof. See the Appendix 2. ■

If consumers are fully rational (i.e., $\beta = 1$), the optimal price is at the marginal cost and the marginal consumer has $k = \delta \left(V(V_x^{-1}(c)) - cV_x^{-1}(c) \right)$.

However, myopic consumers discount excessively the net surplus obtained from the investment in effort. As a consequence, if marginal cost pricing were applied, underinvestment occurs. In fact, when $p = c$, $\epsilon^{inc} < \epsilon^* = \delta \left(V(V_x^{-1}(c)) - cV_x^{-1}(c) \right)$.

To fix the underinvestment problem, the surplus that consumers are able to achieve thanks to the effort must be raised, and this is obtained by decreasing the price below the marginal cost.

Appendix 2: Proofs

Proof of Proposition 1. By substituting (5) and (8) in (11), the latter can be rewritten as

$$\begin{aligned} W = & \int_0^{\epsilon^*} \left[\delta^2 \left[V(V_s^{-1}(p)) - c \left(V_s^{-1}(p) - d_y^{-1}(\delta p) \right) \right] - \delta d(d_y^{-1}(\delta p)) - k \right] dk + \\ & + \int_{\epsilon^*}^{K} \left[\delta^2 \left(V(V_s^{-1}(p)) - cV_s^{-1}(p) \right) \right] dk, \end{aligned}$$

i.e.

$$W = K\delta^2 \left[V(V_s^{-1}(p)) - cV_s^{-1}(p) \right] + \delta^2 c d_y^{-1}(\delta p) \epsilon^* - \delta d(d_y^{-1}(\delta p)) \epsilon^* - \frac{(\epsilon^*)^2}{2}$$

To find the optimal price, we find the FOC of W w.r.t. p:

$$K\delta^2\frac{V_s(s^*)-c}{V_{ss}(s^*)}+\delta^2 c\left(\frac{\delta}{d_{yy}(y^*)}\epsilon^*+d_y^{-1}(\delta p)\frac{\partial\epsilon^*}{\partial p}\right)-\delta\left(\frac{\delta d_y(y^*)}{d_{yy}(y^*)}\epsilon^*+d(d_y^{-1}(\delta p))\frac{\partial\epsilon^*}{\partial p}\right)-\epsilon^*\frac{\partial\epsilon^*}{\partial p}=0.$$

Given that $V_s(s^*)=p$ and $d_y(y^*)=\delta p$, the previous expression becomes

$$K\delta^2\frac{p-c}{V_{ss}(s^*)}+\delta^2 c\left(\frac{\delta}{d_{yy}(y^*)}\epsilon^*+d_y^{-1}(\delta p)\frac{\partial\epsilon^*}{\partial p}\right)-\delta\left(\frac{\delta^2 p}{d_{yy}(y^*)}\epsilon^*+d(d_y^{-1}(\delta p))\frac{\partial\epsilon^*}{\partial p}\right)-\epsilon^*\frac{\partial\epsilon^*}{\partial p}=0.$$

By collecting terms, we obtain

$$K\delta^2\frac{p-c}{V_{ss}(s^*)}-\frac{\delta^3\epsilon^*}{d_{yy}(y^*)}(p-c)+\frac{\partial\epsilon^*}{\partial p}\left(\delta^2 c d_y^{-1}(\delta p)-\delta d(d_y^{-1}(\delta p))-\epsilon^*\right)=0.$$

Since $\epsilon^*=\delta^2 p d_y^{-1}(\delta p)-\delta d(d_y^{-1}(\delta p))$, the last term can be further simplified and the expression becomes

$$K\delta^2\frac{p-c}{V_{ss}(s^*)}-\frac{\delta^3\epsilon^*}{d_{yy}(y^*)}(p-c)-\frac{\partial\epsilon^*}{\partial p}\delta^2 d_y^{-1}(\delta p)\,(p-c)=0.$$

which is verified for $p=c$.

■

Proof of Proposition 2 and Corollaries 1 and 2.

We prove the Proposition and Corollaries through two Lemmas, concerning respectively the characterization (Lemma 1) and the existence (Lemma 2) of a solution of the social planner's problem.

Lemma 1 *If a solution exists, it entails $p^{inc}=c$ for $\beta=1$ or $\beta=0$, and $p^{inc}>c$ for all $\beta\in(0,1)$.*

Proof. Expression (16) can be rewritten as

$$W \;=\; K\delta^2\left[V(V_s^{-1}(p))-cV_s^{-1}(p)\right]+\delta^2 c d_y^{-1}(\beta\delta p)\epsilon^{inc}-\delta d(d_y^{-1}(\beta\delta p))\epsilon^{inc}-\frac{(\epsilon^{inc})^2}{2}$$

To find the optimal price, we find the FOC of W w.r.t. p:

$$\begin{aligned}&K\delta^2\frac{V_s(s^{inc})-c}{V_{ss}(s^{inc})}+\delta^2 c\left(\frac{\beta\delta}{d_{yy}(y^{inc})}\epsilon^{inc}+d_y^{-1}(\beta\delta p)\frac{\partial\epsilon^{inc}}{\partial p}\right)+\\&-\delta\left(\frac{\beta\delta d_y(y^{inc})}{d_{yy}(y^{inc})}\epsilon^{inc}+d(d_y^{-1}(\beta\delta p))\frac{\partial\epsilon^{inc}}{\partial p}\right)-\epsilon^{inc}\frac{\partial\epsilon^{inc}}{\partial p}=0.\end{aligned}$$

Given that $V_s(s^{inc}) = p$ and $d_y(y^{inc}) = \beta\delta p$, the previous expression becomes

$$K\delta^2 \frac{p-c}{V_{ss}(s^{inc})} + \delta^2 c \left(\frac{\beta\delta}{d_{yy}(y^{inc})} \epsilon^{inc} + d_y^{-1}(\beta\delta p) \frac{\partial \epsilon^{inc}}{\partial p} \right) +$$
$$-\delta \left(\frac{\beta^2\delta^2 p}{d_{yy}(y^{inc})} \epsilon^{inc} + d(d_y^{-1}(\beta\delta p)) \frac{\partial \epsilon^{inc}}{\partial p} \right) - \epsilon^{inc} \frac{\partial \epsilon^{inc}}{\partial p} = 0.$$

By collecting terms, we obtain

$$K\delta^2 \frac{p-c}{V_{ss}(s^{inc})} + \frac{\beta\delta^3\epsilon^{inc}}{d_{yy}(y^{inc})}(c-\beta p) + \frac{\partial \epsilon^{inc}}{\partial p} \left(\delta^2 c d_y^{-1}(\beta\delta p) - \delta d(d_y^{-1}(\beta\delta p)) - \epsilon^{inc} \right) = 0, \quad (22)$$

where

$$\epsilon^{inc} = \beta \left(\delta^2 p d_y^{-1}(\hat{\beta}\delta p) - \delta d(d_y^{-1}(\hat{\beta}\delta p)) \right)$$
$$\frac{\partial \epsilon^{inc}}{\partial p} = \beta\delta^2 \left(d_y^{-1}(\hat{\beta}\delta p) + p \frac{\hat{\beta}\delta}{d_{yy}(y_0^{inc})}(1-\hat{\beta}) \right).$$

The solution of condition (22) allows to find p^{inc}. The proofs proceeds in two steps. We first prove that $p^{inc} = c$ for $\beta = 0$ or $\beta = 1$. Secondly, we prove that $p^{inc} \leq c$ cannot be a solution of the social planner's problem for $\beta \in (0,1)$.

When $\beta = 0$, $\epsilon^{inc} = \frac{\partial \epsilon^{inc}}{\partial p} = 0$. Then, (22) becomes $K\delta^2 \frac{p-c}{V_{ss}(s^{inc})} = 0$, which is verified for $p^{inc} = c$.

When $\beta = 1$, $\hat{\beta} = 1$ as well; then, $p^{inc} = c$ from Proposition 1.

Let us prove now that $p^{inc} > c$ for $\beta \in (0,1)$. Suppose, by contradiction, that $p^{inc} \leq c$. Then, the sum of the first two terms of condition (22) would be strictly positive. Since $\frac{\epsilon^{inc}}{\partial p}$ is also positive, the welfare maximization problem has an internal solution when the last term in the parenthesis, that we denote by $L(\beta, \hat{\beta})$, is strictly negative, i.e.

$$L(\beta, \hat{\beta}) = \delta c d_y^{-1}(\beta\delta p) - d(d_y^{-1}(\beta\delta p)) - \beta \left(\delta p d_y^{-1}(\hat{\beta}\delta p) - d(d_y^{-1}(\hat{\beta}\delta p)) \right) < 0.$$

We now study the function $L(\beta, \hat{\beta})$ and show that, even when it is minimized, it is still positive for $p \leq c$, thus contradicting the hypothesis that $p^{inc} \leq c$ is an equilibrium when $\beta \in (0,1)$.

To this aim, note that the function $L(\beta, \hat{\beta})$ is always decreasing in $\hat{\beta}$. This means that it is always mimimized when $\hat{\beta} = 1$, for all p and β.

Moreover, when $\hat{\beta} = 1$, we have that

$$\frac{\partial L(\beta,1)}{\partial \beta} = \frac{\delta^2 cp}{d_{yy}(\beta\delta p)} - \frac{\beta\delta^2 p^2}{d_{yy}(\beta\delta p)} - \left(\delta p d_y^{-1}(\delta p) - d(d_y^{-1}(\delta p)) \right)$$
$$\frac{\partial^2 L(\beta,1)}{\partial \beta^2} = \delta^2 p \frac{-p d_{yy}(\beta\delta p) - (c-\beta p)\delta p d_{yyy}(\beta\delta p)}{(d_{yy}(\beta\delta p))^2}.$$

Since by hypothesis $p \leq c$, then $c - \beta p > 0$; moreover, $d_{yy} > 0, d_{yyy} \geq 0$. Hence, $\frac{\partial^2 L}{\partial \beta^2} < 0$, implying that $L(\beta, 1)$ is a strictly concave function in the support of β. As a consequence, it is minimized for either $\beta = 0$ or $\beta = 1$. In $\beta = 0$, $L(0,1) = 0$. Conversely, in $\beta = 1$, $L(1,1) = \delta(c-p)d_y^{-1}(\delta p)$, which is positive or null for $p \leq c$. It follows that when $p \leq c$, $\min L(\beta, \hat{\beta}) \geq 0$ for all $\beta \in (0,1)$. As a consequence, when $p \leq c$, $L(\beta, \hat{\beta}) \geq 0$ for all $\beta, \hat{\beta} \in (0,1) \times [\beta, 1)$.

Then, the l.h.s. of (22) is strictly positive for $p \leq c$, which implies that there does not exist a solution with $p^{inc} \leq c$ for all $\beta \in (0,1)$. ■

Lemma 1 only shows that the optimal price must be higher than the marginal cost. However, in principle the problem might not be closed. The following Lemma ensures that the problem is closed and thus a solution always exists.

Lemma 2 *A solution to the welfare maximization problem always exists.*

Proof. When $\beta = 0$ or $\beta = 1$, a solution exists and it is equal to $p^{inc} = c$ from Lemma 1. Then, we need only to show that a solution exists when $\beta \in (0,1)$.

First of all, observe that all functions in (22) are continuous in p. Moreover, in $p = c$, the l.h.s. of (22) is strictly positive when $\beta \in (0,1)$ by Lemma 1. Should the l.h.s. of (22) be strictly negative for some $p > c$ when $\beta \in (0,1)$, then there exists some value of p such that the FOC in (22) is equal to zero by continuity, i.e a solution to the welfare maximization exists.

Take the price $p = c/\beta$. Then, the sum of the first two terms of (22) is strictly negative. Since $\frac{\partial \epsilon^{inc}}{\partial p}$ is always positive, then we need only to prove that the third term, i.e., $L(\beta, \hat{\beta})$, is negative when $p = c/\beta$.

Let us focus on $L(\beta, \hat{\beta})$. We want to find the values of $\beta, \hat{\beta}$ such that $L(\beta, \hat{\beta})$ is maximized, and show that for these values L is still negative when p is high enough. Note that $L(\beta, \hat{\beta})$ is decreasing in $\hat{\beta}$. This implies that $L(\beta, \hat{\beta})$ is maximized when $\hat{\beta} = \beta$. By susbstituting $\hat{\beta} = \beta$ and $p = c/\beta$ in the expression of $L(\beta, \hat{\beta})$, it becomes

$$L(\beta, \beta) = \delta c d_y^{-1}(\delta c) - d(d_y^{-1}(\delta c)) - \beta \left(\delta \frac{c}{\beta} d_y^{-1}(\delta c) - d(d_y^{-1}(\delta c)) \right)$$

i.e.

$$L(\beta, \beta) = -(1-\beta) d(d_y^{-1}(\delta c)) < 0$$

for all $\beta \in (0,1)$.

Hence the l.h.s. of (22) is strictly negative for $p = c/\beta > c$.

Since the l.h.s. is strictly positive for $p = c$, it is strictly negative for $p = c/\beta > c$ and is it constituted by continuous functions, then there exists at least one price $p \in (c, c/\beta)$ such that the l.h.s. is equal to zero by continuity, i.e. a solution to the welfare maximization problem exists.

■ ■

Proof of Proposition 3. From the FOC condition, the term

$$\Lambda = \left[\frac{\delta^3\beta c}{d_{yy}} - \frac{\delta^3\beta^2 p}{d_{yy}}\right]\epsilon^{inc} + \left[\delta^2 c d_y^{-1}(\beta\delta p) - \delta d(d_y^{-1}(\beta\delta p)) - \epsilon^{inc}\right]\frac{\partial\epsilon^{inc}}{\partial p}$$

represents the marginal social benefit of increasing the price and it is a function of $\hat{\beta}$. Then, the higher is the marginal benefit Λ, the larger is the optimal price distortion.

In order to assess how the marginal benefit depends on $\hat{\beta}$, we need to determine the sign of $\frac{\partial\Lambda}{\partial\hat{\beta}}$. If $\frac{\partial\Lambda}{\partial\hat{\beta}} > 0$, an increase of the consumers' naivety $\hat{\beta}$ increases the marginal benefit Λ of increasing the price, implying that the optimal price must increase. Conversely, if $\frac{\partial\Lambda}{\partial\hat{\beta}} < 0$, an increase of $\hat{\beta}$ makes the optimal price decrease. We compute the derivative of Λ w.r.t. $\hat{\beta}$:

$$\frac{\partial\Lambda}{\partial\hat{\beta}} = \frac{\beta\delta^3}{d_{yy}}(c-\beta p)\frac{\partial\epsilon^{inc}}{\partial\hat{\beta}} + \frac{\partial^2\epsilon^{inc}}{\partial p\partial\hat{\beta}}\left(\delta^2 c d_y^{-1}(\beta\delta p) - \delta d(d_y^{-1}(\beta\delta p)) - \epsilon^{inc}\right) - \frac{\partial\epsilon^{inc}}{\partial p}\frac{\partial\epsilon^{inc}}{\partial\hat{\beta}}, \quad (23)$$

where

$$\begin{aligned}
\epsilon^{inc} &= \beta\left(\delta^2 p d_y^{-1}(\hat{\beta}\delta p) - \delta d(d_y^{-1}(\hat{\beta}\delta p))\right)\\
\frac{\partial\epsilon^{inc}}{\partial\hat{\beta}} &= \frac{\beta\delta^3 p^2}{d_{yy}}(1-\hat{\beta})\\
\frac{\partial^2\epsilon^{inc}}{\partial p\partial\hat{\beta}} &= \frac{\beta\delta^3 2p}{d_{yy}}(1-\hat{\beta})\\
\frac{\partial\epsilon^{inc}}{\partial p} &= \beta\delta^2\left(d_y^{-1}(\hat{\beta}\delta p) + p\frac{\hat{\beta}\delta}{d_{yy}}(1-\hat{\beta})\right)
\end{aligned}$$

By substituting these expressions in (23), we obtain

$$\begin{aligned}
\frac{\partial\Lambda}{\partial\hat{\beta}} &= \frac{\beta\delta^3}{d_{yy}}(c-\beta p)\frac{\beta\delta^3 p^2}{d_{yy}}(1-\hat{\beta}) + \\
&+ \frac{\beta\delta^3 2p}{d_{yy}}(1-\hat{\beta})\left(\delta^2 c d_y^{-1}(\beta\delta p) - \delta d(d_y^{-1}(\beta\delta p)) - \beta\left(\delta^2 p d_y^{-1}(\hat{\beta}\delta p) - \delta d(d_y^{-1}(\hat{\beta}\delta p))\right)\right) + \\
&- \beta\delta^2\left(d_y^{-1}(\hat{\beta}\delta p) + p\frac{\hat{\beta}\delta}{d_{yy}}(1-\hat{\beta})\right)\frac{\beta\delta^3 p^2}{d_{yy}}(1-\hat{\beta}),
\end{aligned}$$

i.e.

$$
\begin{aligned}
\frac{\partial \Lambda}{\partial \hat{\beta}} &= \frac{\beta\delta^3 p(1-\hat{\beta})}{d_{yy}} \cdot \left[(c-\beta p)\frac{\beta\delta^3 p}{d_{yy}} + \right. \\
&+ 2\left(\delta^2 c d_y^{-1}(\beta\delta p) - \delta d(d_y^{-1}(\beta\delta p)) - \beta\left(\delta^2 p d_y^{-1}(\hat{\beta}\delta p) - \delta d(d_y^{-1}(\hat{\beta}\delta p))\right)\right) + \\
&- \left. \beta\delta^2 p \left(d_y^{-1}(\hat{\beta}\delta p) + \frac{\hat{\beta}\delta p(1-\hat{\beta})}{d_{yy}} \right) \right] = \\
&= \frac{\beta\delta^3 p(1-\hat{\beta})}{d_{yy}} \cdot \left[\frac{\beta\delta^3 p}{d_{yy}} \left(c - \beta p - p\hat{\beta}(1-\hat{\beta}) \right) - \beta\delta^2 p d_y^{-1}(\hat{\beta}\delta p) + \right. \\
&\left. +2\left(\delta^2 c d_y^{-1}(\beta\delta p) - \delta d(d_y^{-1}(\beta\delta p)) + \beta\delta d(d_y^{-1}(\hat{\beta}\delta p)) - \beta\delta^2 p d_y^{-1}(\hat{\beta}\delta p)\right)\right].
\end{aligned}
$$

Since the first factor, $\frac{\beta\delta^3 p(1-\hat{\beta})}{d_{yy}}$, is always positive, the sign of the derivative is given by the term in the square parenthesis, that is a function of $\beta, \hat{\beta}$ and that we denote $M(\beta, \hat{\beta})$.

When β goes to 0, we have

$$
\left.\frac{\partial \Lambda}{\partial \hat{\beta}}\right|_{\beta=0} = \frac{\beta\delta^3 p(1-\hat{\beta})}{d_{yy}} \frac{\beta\delta^3 p}{d_{yy}} (c - p\hat{\beta}(1-\hat{\beta})) = 0^+,
$$

Hence, when β goes to 0, $\frac{\partial \Lambda}{\partial \hat{\beta}} > 0$: the marginal benefit of raising the price (and thus the optimal price) is higher for naive consumers than for sophisticated ones.

When β goes to 1, $\hat{\beta}$ goes to 1 as well, as $\hat{\beta} \geq \beta$. Then,

$$
\left.\frac{\partial \Lambda}{\partial \hat{\beta}}\right|_{\beta=1} = \frac{\beta\delta^3 p(1-\hat{\beta})}{d_{yy}} \cdot \left[-\delta^2 p d_y^{-1}(\delta p) \right] = 0^-.
$$

Hence, when β goes to 1, $\frac{\partial \Lambda}{\partial \hat{\beta}} < 0$: the optimal price is lower for naive consumers than for sophisticated ones.

■

Proof of Proposition 4. The welfare can be rewritten as

$$
W = \beta\delta^2 \left(V(V_x^{-1}(p)) - cV_x^{-1}(p) \right) \left(V(V_x^{-1}(p)) - pV_x^{-1}(p) \right) - \frac{1}{2} \left[\beta\delta \left(V(V_x^{-1}(p)) - pV_x^{-1}(p) \right) \right]^2
$$

We compute the FOC w.r.t. p:

$$
\begin{aligned}
&-\beta\delta^2 \left(\frac{V_x}{V_{xx}} - \frac{c}{V_{xx}} \right) \left(V(V_x^{-1}(p)) - pV_x^{-1}(p) \right) + \\
&-\beta\delta^2 \left(V(V_x^{-1}(p)) - cV_x^{-1}(p) \right) \left(\frac{V_x}{V_{xx}} - \frac{p}{V_{xx}} - V_x^{-1}(p) \right) + \\
&+\frac{\beta^2\delta^2}{2} 2 \left(V(V_x^{-1}(p)) - pV_x^{-1}(p) \right) \left(\frac{V_x}{V_{xx}} - \frac{p}{V_{xx}} - V_x^{-1}(p) \right) = 0.
\end{aligned}
$$

Since $V_x = p$ from the demand function, the previous condition can be simplified into

$$-\frac{p-c}{V_{xx}}\left(V(V_x^{-1}(p)) - pV_x^{-1}(p)\right) +$$
$$+V_x^{-1}(p)\left(V(V_x^{-1}(p)) - cV_x^{-1}(p)\right) +$$
$$-\beta V_x^{-1}(p)\left(V(V_x^{-1}(p)) - pV_x^{-1}(p)\right) = 0.$$

Moreover, given that $\epsilon^{inc} = \beta\delta\left(V(V_x^{-1}(p)) - pV_x^{-1}(p)\right)$, we substitute $V(V_x^{-1}(p)) - pV_x^{-1}(p) = \frac{\epsilon^{inc}}{\beta\delta}$ into the FOC and obtain

$$-\frac{p-c}{V_{xx}}\frac{\epsilon^{inc}}{\beta\delta} + V_x^{-1}(p)\left(V(V_x^{-1}(p)) - cV_x^{-1}(p) - \frac{\epsilon^{inc}}{\delta}\right) = 0. \tag{24}$$

When $\beta = 1$, then $\epsilon^{inc} = \delta\left(V(V_x^{-1}(p)) - pV_x^{-1}(p)\right)$. In this case, (24) becomes

$$-\frac{p-c}{V_{xx}}\left(V(V_x^{-1}(p)) - pV_x^{-1}(p)\right) + V_x^{-1}(p)\left[V(V_x^{-1}(p)) - cV_x^{-1}(p) - \left(V(V_x^{-1}(p)) - pV_x^{-1}(p)\right)\right] = 0$$

i.e.

$$-\frac{p-c}{V_{xx}}\left(V(V_x^{-1}(p)) - pV_x^{-1}(p)\right) + [V_x^{-1}(p)]^2(p-c) = 0$$

Since $V(V_x^{-1}(p)) - pV_x^{-1}(p) > 0$ (as it constitutes the future surplus obtained from the consumption), it must be $p^{inc} = c$.

When $\beta = 0$, (24) can be rewritten as

$$-\frac{p-c}{V_{xx}}\left(V(V_x^{-1}(p)) - pV_x^{-1}(p)\right) +$$
$$+V_x^{-1}(p)\left(V(V_x^{-1}(p)) - cV_x^{-1}(p)\right) = 0,$$

which is verified for $p^{inc} = c$.

Consider now the case in which $0 < \beta < 1$. We want to prove that $p^{inc} < c$. Suppose, by contradiction, that $p^{inc} \geq c$.

Then, from (24), the term $-\frac{p-c}{V_{xx}}\frac{\epsilon^{inc}}{\beta\delta}$ would be positive (as $p \geq c$ and $V_{xx} < 0$). Then, to have a solution the second term $V(V_x^{-1}(p)) - cV_x^{-1}(p) - \frac{\epsilon^{inc}}{\delta}$ should be negative, i.e.

$$V(V_x^{-1}(p)) - cV_x^{-1}(p) \leq \beta\left[V(V_x^{-1}(p)) - pV_x^{-1}(p)\right].$$

This implies that

$$\beta \geq \frac{V(V_x^{-1}(p)) - cV_x^{-1}(p)}{V(V_x^{-1}(p)) - pV_x^{-1}(p)}.$$

Note that, being $p \geq c$, the r.h.s. of the previous inequality is greater than 1, which contradicts the assumption that $\beta < 1$.

Hence, $p^{inc} < c$ for all $\beta \in (0, 1)$. ■