Bureau d'économie théorique et appliquée (BETA)
UMR 7522

Documents de travail

« Environmental Incentives: Nudge or Tax? »

Auteurs

Benjamin Ouvrard, Sandrine Spaeter

Document de Travail n° 2016 – 23

Avril 2016

Faculté des sciences économiques et de gestion
Pôle européen de gestion et d'économie (PEGE)
61 avenue de la Forêt Noire
F-67085 Strasbourg Cedex

Secrétariat du BETA
Géraldine Del Fabbro
Tél. : (33) 03 68 85 20 69
Fax : (33) 03 68 85 20 70
g.delfabbro @unistra.fr
www.beta-umr7522.fr

Environmental Incentives: Nudge or Tax?*

Benjamin OUVRARD†and Sandrine SPAETER‡

Abstract

We consider a model where individuals can voluntarily contribute to improve the quality of the environment. They differ with regard to their confidence in the announcement made by the regulator about the risk of pollution, modelized in a RDEU model, and to their environmental sensitivity. We compare the efficiency of a tax in increasing individual contributions with the advantages of a nudge based on the announcement of the social optimum to each individual. Under some conditions, a nudge performs better than a tax, in particular, because the individual reaction depends directly on sensitivity, while only indirectly with a tax. Moreover, a nudge does not require information about private contributions, contrary to a tax based on the contributions that are not provided compared

*We would like to thank Stefan Ambec, Anthony Heyes, Mathieu Lefebvre, François Pannequin and Bruno Rodrigues for their helpful comments. We are also grateful to the participants at the "Journée Incertitude et Décision Publique" (Nanterre, 2014), the BETA seminar (Strasbourg, 2015), the EAERE Summer School (Venice, 2015), the WRIEC (Munich, 2015), the EEA conference (Manheim, 2015) and the FAERE annual conference (Toulouse, 2015).

†, BETA, CNRS and University of Strasbourg; E-mail: bouvrard@unistra.fr, Tel: +33 (0) 368 852 091

‡, BETA, CNRS and University of Strasbourg; E-mail: spaeter@unistra.fr, Tel: +33 (0) 368 852 076

to the social optimum. Lastly, its implementation is much cheaper. Yet, some drawbacks are discussed and simulations illustrate our results.

Key Words : incentives; nudge; environmental sensitivity; probability distorsion; tax.

JEL Codes : Q50, D8.

1 Introduction

In Environmental Economics, there has been a long tradition of market-based incentives to regulate environmental pollution (Cropper and Oates (1992), Helfand et al. (2003)). Regulations through prices (monetary constraints) and regulations through quantities (volume constraints) are commonly used. In both cases, the objective is to change consumers' behavior to pollute less. However, these tools have some drawbacks. Considering taxes, they are sometimes difficult to implement from a political point of view (Thalmann (2004), Gaunt et al. (2007)). Lobbies and/or political parties struggle against their implementation. Recently, the French government intended to implement an environmental tax on heavy trucks. However, because of the public pressure it faced, in October 2014 the government postponed this tax *sine die* . Experimental evidence (Kallbekken et al. (2011)) also highlights the lack of public support to Pigouvian taxes. One reason is that individuals do not perceive the difference between a Pigouvian tax (intended to correct an externality) and a Ramsey one (intended to raise revenue). Once implemented, the governments need to consider the costs due to the collection of the tax. Finally, because agents are generally opposed to the implementation of taxes, there exists a "social cost" of taxation. Turning to subsidies, one main issue comes from the financial limits governments have to deal with. Moreover, inefficiencies can occur given that firms on a given market may be interested in the subsidy only. Tradable permits can restore an optimum, and present also the advantage to let firms bargain between them. However, transaction costs (Stavins (1995), Jaraité-Kažukauské and Kažukauskas (2015)) and the initial allocation of permits (Sartzetakis (2004), Hahn and Stavins (2011)) may be responsible for inefficiences. The research of an alternative, or a complementary tool is thus justified.

One direction can be found in Thaler and Sunstein's (2009) book. They argue in favor of *nudges* that are ways of improving consumers' behavior by alerting them on their behaviour and/or on the neighbors' behaviors, through the disclosure of some specific information or the use of simple techniques such as default options[1]. In this paper, we focus on the disclosure of information. A nudge must be simple to implement, costless and should not constrain individuals[2].

This new tool has been recently studied both by psychologists and economists. Nolan et al. (2008) use descriptive norms[3] to make individuals increase their conservation behavior, that is to say to use less energy. They show that those who received an information about the (similar) neighbor who consumes the less reduced their electricity consumption, compared to the control group. Goldstein et al. (2008), still using descriptive norms, obtain a towel reuse rate of 49.3% in hotels. In particular, their descriptive norm mentioned that most guests in the room re-used their towel. Schultz et al. (2007) emphasize the boomerang effect that may be created by the use of descriptive norms: those discovering that they contribute more than similar individuals

[1]Default options are options that are pre-selected for individuals. For instance, some banks do no longer send paper reports concerning bank accounts (default option). However the customers can formally ask to receive them again. Egebark and Ekström (2016) propose the use of default options to reduce paper use. In particular, they obtain that default options induced a reduction of paper use of 15% compared to the baseline.

[2]This means that the set of options that individuals initially have must not be reduced when introducing a nudge.

[3]Descriptive norms are based on the disclosure of information on the usual behavior of a majority of individuals in a given situation.

might decrease their contribution. Thus they propose to mix injunctive norms[4] with descriptive norms to counterbalance this possible boomerang effect.

Natural field experiments[5] have also been conducted to study the impact of nudges on energy conservation. In Allcott (2011), the company Opower sent periodic reports to households mentioning their own level of consumption, the average level of consumption of similar households, and the one of the most efficient neighbor. Tips to reduce electricity consumption were also included. A mean reduction of 2,1% was observed among households: such a reduction would have corresponded to an increase of the prices of electricity of 11% to 20% in the short run. Similar results were found by Ayres et al. (2013) in the context of electricity consumption and natural gas. Ferraro and Price (2013) focused on water use. In their study, they compared three different treatments: technical advice letter (TAL); TAL and an appeal to prosocial preferences (APP); and TAL, APP and a social comparison. They found that the last combination was inducing the highest reduction in water use (4,8%) for an average household, compared to a control group household. Finally, Costa and Kahn (2013) show that individuals do not react identically to a nudge. In their study, they show that political liberals are much more likely to reduce their electricity consumption when receiving personal reports, than political conservatives. This raises the question of knowing what are the conditions on individuals' preferences that induce a positive

[4] Injunctive norms refer to "rules or beliefs as to what constitutes morally approved and disapproved conduct" (Cialdini (1990), p.1015). The use of emoticons is an example of an injunctive norm: a smiley (sad emoticon) indicates that the behavior is (not) approved.

[5] Harrison and List (2004) define a natural field experiment as an experiment in which "the environment is one where the subjects naturally undertake [the] tasks and where the subjects do not know that they are in an experiment".

impact of a nudge on their behavior. Banerjee et al. (2014) propose a lab experiment to study the impact of information disclosure on the performance of a monetary incentive mechanism to induce landowners to spatially coordinate. They show that subjects receiving additional information on the behavior of their indirect neighbors are more likely to coordinate on the Pareto optimal outcome, compared to those receiving only information on the behavior of their direct neighbors.

Although nudges are considered as a rather new incentive tool, politicians already focus on. This can be observed in the reports that have been written on health prevention (see the Behavioural Insights Team's reports (2010, 2013) in Great Britain), and on environmental protection (the report by the Centre d'Analyse Stratégique - CAS (2011), the Behavioural Insights Team's report (2011), or the report of the OECD (2012)). The European Commission's report (2016) details the different nudges that have been tested in different contexts. In particular, normative messages have been used to encourage taxpayers to use online services (in France), or to pay on time (in Great Britain). In France, Thoyer et al. (2015) consider nudges as a tool for the regulation of pesticides' use in the agricultural sector. In particular, they show that such an instrument may induce the creation of new social norms, that may be adopted by landowners.

Although quoted papers display encouraging results in the field of energy conservation, they miss a theoretical model that provides some theoretical predictions as a support to policies. In this paper we propose such a theoretical model. It will permit us to compare the efficiency of a nudge to a tax. We also provide some simulations in order to illustrate our theoretical results.

In the context of energy conservation, a reduction of consumption by individuals who are not constrained can be seen as a voluntary contribution to the environmental good. The existing literature on voluntary contributions is quite important (see the surveys by Ledyard (1995) and Chaudhuri (2011)). By focusing on random pollution, Etner et al. (2007) consider the environmental quality as a public good. They show that optimistic individuals contribute less than pessimistic ones. Etner et al. (2009) also show the importance of initial wealth and of the initial level of environmental quality on the voluntary contributions. Salanié and Treich (2009) study the regulator's point of view. They consider a model in which citizens may hold different beliefs (optimism or pessimism) from the regulator's ones. In particular, they show that a paternalistic regulator (who takes into account a social utility rather than the sum of citizen's preferences) may over-regulate compared to a populist one (who takes into account the citizens' beliefs).

In this paper, we consider only optimistic individuals in the sense of the RDEU model (Quiggin, 1982), those who tend, *a priori*, to undercontribute. More precisely, individuals have a more or less high degree of confidence in the announcement made by the regulator about the risk of pollution. This degree of confidence is captured by the distortion function. We also define a disutility function of pollution that depends on the individual sensitivity to the environment, an intrinsic characteristic of her preferences. This sensitivity can be health vulnerability, psychological feelings, personal convictions about environmental considerations, etc. By doing so, we want to capture a diversity of behaviors. Indeed, there is evidence in the literature of consumers' heterogeneity. For instance, Kotchen and Moore (2008) propose a model in which they consider conservationists (who incur guilt when consuming conventional electricity) and nonconservationists. Testing their predictions with an empirical study, they find that conservationists consume 10% less electricity than nonconservationists. Conservationists are also more likely to participate in green-electricity programs. In another study, Kahn (2007) find that Californian environmentalists use more often "green" transportation (hybrid cars for instance) than the average consumer. Thus, it appears that considering an individual heterogeneity through different levels of environmental sensitivity is fair.

We believe that there is not necessarily a positive correlation between the degree of probability distortion of the individual and her sensitivity to environmental considerations. For instance, a given individual may strongly distort the probabilities because she believes that the regulator overestimates the risk of environmental pollution, although she is sensitive to environmental quality in the meantime. Finally, contributions to the environmental public good depend on two dimensions: the distorsion probability and a qualitative index of individual environmental sensitivity. In this context, we show that a nudge may be more efficient, under some conditions, than a tax in inducing more contributions. This is good news knowing that implementing a nudge may also mobilize less ressources than a tax policy, as discussed in the paper. We explain that a nudge, contrary to a tax, is able to discriminate between individuals with different intrinsic characteristics. More precisely, two different individuals (regarding their intrinsic characteristics) may contribute the same for environmental quality, before and after taxation. We show that this is no longer the case with the implementation of a nudge because individuals' reaction depends on environmental sensitivity: two individuals differing with respect to their environmental sensitivity react differently, even if they were contributing the same for environmental quality in the absence of incentives. Some illustrative simulations are also provided.

Our main contribution is based on the identification of theoretical predictions about the impact to be expected from the use of a nudge, and the comparison with a more standard tax policy. Indeed, in our knowledge we propose the first model that modelizes the nudge and individuals' reaction to it. A second contribution deals with the empirical testability of our results. Indeed, both risk perception and environmental sensitivity can be elicited either by experiments or by questionnaires, as explained in the paper.

In section 2, first we present the private optima without any incentive regulatory policy. Then we introduce a tax based on individual contributions. In section 3, we define the nudge and we build the individual reaction to the nudge. We evaluate the private optima and compare the impact of a nudge with the impact of the tax policy on individual contributions. In Section 4, simulations are provided. Section 5 concludes the paper.

2 Theoretical predictions when monetary incentives are used

In this section, we present a model in which individuals can voluntarily, and financially, contribute to improve the quality of the public good (environment). First, there is no outside incentives to contribute. Second a standard tax policy is introduced. In both models, individuals are more or less optimistic regard the announce made by the regulator about the level of the risk of pollution. They are also more or less (psychologically and/or physically) sensitive to environmental quality.

2.1 The benchmark model with probability distortions

We start introducing the model and the different assumptions that we consider. Then we provide the private optima and compare them to the social ones in terms of voluntary contributions.

2.1.1 Assumptions

We consider an economy with a fixed population. Individuals face an aggregate level of pollution (public bad) emitted by human activity. They can voluntarily contribute to make decrease this level of pollution. The current random level of pollution, $\tilde{P}$, is given by

$$\tilde{P} = \tilde{e}Y - b(A) \tag{1}$$

where $\tilde{e}Y$ is the pollution coming from the current production Y. $\tilde{e}$ is a random variable the values of which belong to the interval $[\underline{e};\overline{e}]$, with $\underline{e} \geq 0$ and $\overline{e} \leq 1$. For the sake of simplicity, we assume that $\tilde{e}$ is uniformly distributed, $F(.)$ being the distribution function and $f(.)$ its density.

A is the sum of individual contributions, $A = \sum\limits_{i} a_{k,j}$, with each individual i choosing her level of contribution $a_{k,j}$ in a non-cooperative way. Function $b(A)$, with $b'(A) > 0$ and $b''(A) < 0$, represents the public benefit of pollution reduction coming from the individual contributions.

Each individual incurs a disutility from pollution formalized by the disutility function $d(\tilde{P}, s^j)$, which is increasing and convex in $\tilde{P}$: $0 < d_P < +\infty$ and $d_{PP} \geq 0$. The disutility of a given level of pollution P may differ from one individual to another one because of differences in their individual sensitivity to the environmental good, captured by the qualitative variable s^j. For each individual, s^j can take one of two possible values: $s^j \in \{s^l; s^h\}$ with $s^h > s^l$. s^h represents an individual highly sensitive to the environment, and s^l an individual less sensitive. This sensitivity can be health vulnerability, psychological feelings, personal convictions about environmental considerations, etc. Hence, two agents facing the same pollution P do not bear the same disutility of this pollution: a higher sensitive agent presents a higher marginal disutility of pollution: $d_{Ps^j} > 0$. Finally, we also assume that individual i's sensitivity has only a first order impact on the individual disutility of pollution: $d_{PPs^j} = 0$.

Individuals have also heterogeneous perceptions of the risk of pollution. To simplify, we consider two types of individuals regarding risk perception, both being optimistic about the risk of pollution announced by the regulator (or experts): the optimistic individuals always overevaluate the probability of having to bear a level of pollution lower than a given threshold, whatever this threshold. Individuals differ according to their type, $\theta^k = \{\theta^O; \theta^o\}$, with $\theta^O < \theta^o \leq 1$. Type θ^O is highly optimistic, while Type θ^o is less optimistic regard the distribution of pollution announced by the regulator. The threshold 1 represents an individual who takes the information disclosed by the regulator as given and who does not transform it. To formalize the heterogeneity in risk perception, let us denote by $H(F(e), \theta^k)$ the probability transformation function of individuals of Type θ^k.[6] Function H satisfies the following properties:

$$\begin{array}{ll} (i) & H(F(\overline{e}), \theta^k) = F(\overline{e}) = 1, \quad \forall \theta^k \\ (ii) & H(F(\underline{e}), \theta^k) = F(\underline{e}) = 0, \quad \forall \theta^k \\ (iii) & \frac{dH}{de} = \frac{\partial H}{\partial F}.F'(e) > 0 \quad \text{and} \quad \frac{d^2H}{de^2} < 0, \quad \forall e \in]\underline{e}, \overline{e}[\end{array} \tag{2}$$

[6]This formulation is derived from a RDEU model (Quiggin (1982)).

Type θ^O being more optimistic than Type θ^o, the former overevaluates more the probability of having to bear a level of pollution that is lower than a given level. Thus we have

$$H(F(e), \theta^O) \geq H(F(e), \theta^o) \geq F(e) \text{ with } \theta^O < \theta^o \leq 1 \text{ and } \forall e,$$

with at least one strict inequality for each relation[7]. This property corresponds to a first order stochastic dominance: $F(.)$ dominates $H(.)$.

In this configuration, the difference to be made between the risk perception and the environmental sensitivity is essential. On one hand, one can consider risk perception as an indicator on how much individuals are confident[8] in the informations they receive from the regulator about the distribution $F(e)$. This is captured by the function $H(.,.)$.

[7]An example of function $H(.)$ that satisfies the previous properties is:

$$H[F(e), \theta^k] = \left(\frac{\bar{e}}{\underline{e}}\right)^{1-\theta^k} \times F(e)$$

with $\underline{e} > 0$, and θ^k the degree of distortion. The higher the level of θ^k, the lower the distortion of the objective distribution $F(e)$ by agent k. Other admissible functional forms exist in the literature as the linear in log odds (in Wu and Gonzalez (1999)) but initially proposed by Lattimore et al. (1992)) with the curvature parameter equal to 1, and the elevation parameter larger than 1. Similarly, Prelec (1998) proposed an admissible function if the curvature parameter is equal to 1, and the anti-index pessimism is between 0 and 1.

[8]Sinclair-Desgagné and Gozlan (2003) also consider a model in which a stakeholder (which can be an activist organization) is more or less confident with a report made by a polluter (a firm for instance) on its level of pollution. We differ from their paper because the individuals in our model do not have the possibility to perform tests to verify the information they receive. Moreover, in our model the regulator provides some information to help individuals choosing their level of contribution, while in Sinclair-Desgagné and Gozlan (2003), information is provided by the regulator and the stakeholder decides to "accept" or "boycott" the polluters' activity.

The degree of confidence of a given individual may differ depending on who is providing the information (governments, NGO, experts, etc).[9] On the other hand, sensitivity to the environment is different from risk perception (confidence in our model): the qualitative variable s_j is an intrinsic characteristic of the individual preferences and it does not depend on a third party. It characterizes the impact pollution has on the psychological and/or physical welfare of the individual. Contrary to risk perception and in some manner, it cannot be "manipulated". Sensitivity to the environment is something that individuals *live*, while risk perception is linked to something that individuals *interpret*.

Thus each individual is characterized by a subjective type of risk perception θ^k and a level of environmental sensitivity s^j. Individual i can then be called individual (k, j) and four profiles may exist: $(O, h), (O, l), (o, h)$ and (o, l). For instance, individual (O, h) presents a high environmental sensitivity but a low confidence in what is announced by the regulator (she is highly optimistic).

Lastly, each individual receives the same fixed wage w, which is shared between their private consumption $c_{k,j}$ and their contributions $a_{k,j}$ to the environmental quality. The individual utility of consumption is $u(.)$ with $u'(.) > 0$ and $u''(.) \leq 0$. Without any external incentives, the total expected utility of individual i or (k, j) is thus:

$$U_{k,j} = \int_{\underline{e}}^{\overline{e}} \left(u(c_{k,j}) - d(P, s^j)\right) dH\left(F(e), \theta^k\right)$$

with $w = c_{k,j} + a_{k,j}$. Thus her program writes:

$$\max_{a_{k,j}} \int_{\underline{e}}^{\overline{e}} \left(u(w - a_{k,j}) - d(P, s^j)\right) dH\left(F(e), \theta^k\right) \tag{3}$$

[9] Siegrist and Cvetkovich (2000) and Slovic (2013) confirm that some correlation exists between the individuals' level of trust and their risk perception.

$$s.t. \quad a_{k,j} \geq 0 \tag{4}$$

This program is different from the regulator's one as discussed below. Contrary to the regulator, the individual does not take into account the benefit of her contribution to the public good on the welfare of the other individuals.[10]

As explained below, beliefs and environmental sensitivity are two different dimensions of the individual's characteristics.[11] Hence regulatory tools, ideally, should be built on both dimensions. Nevertheless some characteristics are not easily observable by a regulator, in particular, environmental sensitivity. In what follows, we show that the regulator does not need to know who is who in order to implement the socially optimal contribution when individuals do not distort the information received on the risk of pollution. Unfortunately, this is no longer the case when individuals distort the announced distribution of the risk of pollution, namely when they are optimistic in our setting.

2.1.2 Private optima of optimistic individuals

Let us, first, focus on the private optimum of each individual without any external incentive. They depend on her type θ^k and her environmental sensitivity s^j. Each

[10] Recall that a utilitarian and perfectly informed regulator maximizes the social welfare characterized by $\sum_j \sum_k U_{k,j}$.

[11] The difference between optimism and environmental sensitivity is well established in the psychological literature. Individuals are said to be optimistic when they "expect things to go their way, and generally believe that good rather than bad things will happen to them" (Scheier and Carver (1985, p.219)). See also Scheier and Carver (1992). Environmental sensitivity is defined as "an empathetic perspective toward the environment" (Hungerford and Volk (1990, p.11). It corresponds to feelings or attitudes that individuals express toward the environment.

individual (k,j) considers Program (3)-(4). The first order condition for a private interior solution $a^p_{k,j}$ is given by:

$$-u'(w - a^p_{k,j}) + b'(A) \int_{\underline{e}}^{\overline{e}} d_P(P, s^j) dH\left(F(e), \theta^k\right) = 0 \qquad (5)$$

The second order condition is satisfied.[12]

Recall that the disutility from pollution is more important for an individual highly sensitive to the environment than for a less sensitive individual. In equation (5), the expected marginal benefit of the individual contribution is affected by risk distortion, not the marginal cost. Assuming that the regulator can observe[13] individuals' confidence toward the announcement of the risk of pollution, and thus θ^k, we obtain Proposition 1.

Proposition 1 *Let us consider optimistic individuals with preferences characterized in a RDEU model.*

(i) All individuals shall contribute the same at the social optimum, whatever their sensitivity and risk perception.

(ii) All individuals contribute less than the social optimum.

(iii) The most optimistic individuals are not systematically the lowest contributors. Precisely, we have: $a^p_{O,l} < a^p_{o,l} < a^p_{o,h}$ *and* $a^p_{O,l} < a^p_{O,h} < a^p_{o,h}$ *but nothing else can be said without additional assumptions.*

---

[12] Indeed we have

$$u''(w - a^p_{k,j}) + b''(A) \int_{\underline{e}}^{\overline{e}} d_P(P, s^j) dH(F(e), \theta^k) - (b'(A))^2 \int_{\underline{e}}^{\overline{e}} d_{PP}(P, s^j) dH(F(e), \theta^k) < 0$$

[13] The Eurobarometer in Europe, or the General Social Survey in the United States, are surveys measuring individuals' confidence.

From Point (i), individuals shall contribute to the public good the same part of their wealth despite their heterogeneity in individual sensitivity and risk perception. Indeed, recall that the regulator is a utilitarist one and that individuals differ through their preferences, not their initial wealth. The regulator maximizes $\sum_j \sum_k U_{k,j}$ so that the preferences of one individual are embedded in the socially optimal contributions of the others: the effect of each individual's contributions is internalized by everyone. Obviously, this is because environmental quality is a public good.

When focusing on private optimal levels, optimism explains underinvestment in the public good, whatever the environmental sensitivity. Recall that we consider incomplete information about environmental sensitivity regard the regulator, not a total absence of information. The regulator knows how many individuals are highly (slightly) sensitive (one for each case in our simple model). The result of Point (ii) would no longer hold if no information at all were available for the regulator. In such a setting, a highly sensitivity could counterbalance optimism such that some individuals could privately overcontribute compared to the optimum which would be computed by the regulator with a mean level of sensitivity. We do not consider this case in this paper.

Point (iii) illustrates the trade-off between optimism and environmental sensitivity: all types (k, j) cannot be ranked regard their individual contributions. In particular, an individual who does not trust the regulator at all when announcing the risk of pollution (type O in our simple model) can still invest more in the public good than more trustful individuals if her environmental sensitivity is high.

In what follows, we introduce a tax in the model in order to obtain the theoretical predictions on individual contributions that will be compared to those obtained with a nudge in Section 3.

2.2 Implementation of a tax

Consider that individual contributions are observable by the regulator. He implements a tax on the contributions that are not provided by comparison with the social optimum denoted as $a^*_{k,j}$. Let us denote as $t(a_{k,j})$ an exogenous and continuous tax function[14], which satisfies $t'(a_{k,j}) \leq 0$, $t(a_{k,j}) = 0 \quad \forall a_{k,j} \geq a^*_{k,j}$. The new program writes:

$$\max_{a_{k,j}} \int_{\underline{e}}^{\overline{e}} \left(u(w - a_{k,j} - t(a_{k,j})) - d(P, s^j)\right) dH\left(F(e), \theta^k\right) \tag{6}$$

$$s.t. \quad a_{k,j} \geq 0$$

The first-order condition for an individual private interior solution $a^t_{k,j}$ is:

$$-(1 + t'(a^t_{k,j})).u'(w - a^t_{k,j} - t(a^t_{k,j})) + b'(A) \int_{\underline{e}}^{\overline{e}} d_P(P, s^j) dH\left(F(e), \theta^k\right) = 0 \tag{7}$$

The second order condition is always satisfied if the second best tax function is linear or convex in the non provided contributions.

---

[14]A simple example is

$$t(a_{k,j}) = \begin{cases} \tau(a^*_{k,j} - a_{k,j}) & \text{if } a_{k,j} < a^*_{k,j} \\ 0 & \text{if } a_{k,j} \geq a^*_{k,j} \end{cases}$$

With $0 < \tau < 1$. We derive the optimal structure of the second best tax in the Appendix.

Proposition 2 *Assume that the regulator does not know who is who (incomplete information). He implements an increasing second best tax on the socially optimal contributions that are not provided by the individuals. Then, the tax policy is not able to discriminate among the different environmental sensitivities. In particular, for individuals (o,l) and (O,h):*

i) If $a^P_{o,l} = a^P_{O,h}$ then $a^t_{o,l} \gtreqless a^t_{O,h}$ iff $d_{PPP} \lesseqgtr 0$.

ii) If $a^P_{o,l} > a^P_{O,h}$ then $a^t_{o,l} > a^t_{O,h}$ iff $d_{PPP} \leq 0$.

iii) If $a^P_{o,l} < a^P_{O,h}$ then $a^t_{o,l} > a^t_{O,h}$ iff $d_{PPP} \geq 0$.

Under complete information, a first best tax function exists and the regulator is able to push each individual to choose the socially optimal level of contributions. This first best tax function depends, for each individual, on her environmental sensitivity and on her risk perception. Unfortunately, it is no longer possible to implement it whenever information becomes incomplete. We derive the optimal second best tax function in the Appendix, when the regulator knows the different types (k,j) in the economy, but does not know who is who. This tax function is increasing in the non provided contributions. It depends only on the individual contributions.

Although the individual takes also into account her marginal disutility of pollution when deciding $a^t_{k,j}$. This is formalized by the role played by d_{PPP} in Proposition 2. Recall that $d(P,s^j)$ is the disutility of aggregate pollution P borne by the individuals with sensitivity j (whatever their risk perception). Now, by using the notation $v^j(\tilde{P}) = -d(\tilde{P},s^j)$, we are facing their utility function of pollution with $v^j(0) = v_{max}$. The appropriate properties become $v^j_P < 0$, $v^j_{PP} < 0$.

Following Ebert et al. (2016) and Ebert and van de Kuilen (2015), our individuals prefer less than more regard pollution ($v^j_P < 0$), and they are risk averse ($v^j_{PP} < 0$).

Indeed, from Eeckhoudt and Schlesinger (2006), $v^j_{PP} < 0$ means that the individuals prefer to combine good with bad, and dislike increases in pollution *risk*. Hence, they are risk averse over pain just like they would be risk averse over money.

Now, focusing on the sign of d_{PPP} gives interesting information about how the individual manages downside risk, in other terms an additional risk that would appear in bad states (in the high state of pollution $\overline{e}$ in our setting). Indeed, following Ebert et al. (2016) and Kimball (1990), an individual with preferences that satisfy $d_{PPP} \leq 0$ is said to be non prudent (or imprudent), in a sense of preferring a non avoidable additional risk in a low state, here in state $\underline{e}$. On the contrary, a prudent individual ($d_{PPP} > 0$) would prefer to bear the additional risk in the state of high pollution ($\overline{e}$) in our setting. Actually, as shown in Ebert et al. (2016), an individual preferring more money to less money, and being risk averse and prudent when facing monetary risks, is also the one who prefers less pollution (pain) to more pollution, and who is risk averse and non prudent when facing pollution (pain) risks. Here, $d_{PPP} < 0$ seems to be the fair assumption.

Finally, knowing that risk attitudes (d_{PP}, d_{PPP}) may be ellicited through experimentation, interviews or analyses of revealed preferences, it is empirically possible to anticipate some individuals' reactions to the tax in our simple model. Unfortunately, knowing $d_{PPP}(P, s^j)$ does not permit it to make depend the tax function on environmental sensitivity. Hence two agents having chosen the same level of contribution a, but with different sensitivities, bear the same amount of tax $t(a)$. From Condition (7) the tax only affects the utility of consumption, which is certain and identical for two individuals contributing the same level before taxation, whatever their environmental sensitivity. The second best tax does not depend on environmental sensitivity, while the reaction to it does.

In the following section, we show how the introduction of a nudge may partly solve this lack of refinement.

3 Implementation of a nudge

First, we model the reaction of the individuals to a nudge. Second, we calculate the private optima of contribution. Third, we compare the results with those obtained with the tax.

3.1 Modelling the reaction to the nudge

The nudge considered in this model is an action of information disclosure. Once the regulator makes her announcement denoted as $\hat{a}$, we claim that individuals may adjust their contribution differently depending on their own environmental sensitivity s^j. Let $\eta(\hat{a}|s^j)$ model the impact of the nudge $\hat{a}$ on the individual welfare. Knowing that the individual will compare (or not) the announcement to her own behavior, it is fair to assume that

$$\eta(\hat{a}|s^j) \equiv g(a_{k,j} - \hat{a}|s^j)$$

with $g(0) = 0$, $g_{a_{k,j}} \leq 0$ and $g_{a_{k,j}a_{k,j}}(.) \geq 0$. These assumptions are in line with what is observed in random field experiments (Allcott, 2011; Ayres et al., 2013; Ferraro and Price, 2013).[15] With this design of $g(.|s^j)$ we simply make the assumption

[15] Recall that in our model, individual contributions are never higher than the social ones. Thus $g(a_{k,j} - \widehat{a}|s^j)$ is always positive if the nudge is the the announcement of the social individual contribution a^*. However, if the nudge is based on the mean contribution, then $g(a_{k,j} - \widehat{a}|s^j)$ becomes

that individuals consider the distance between their contribution and the announcement. This is different from Figuières et al. (2013). Indeed we take into account the individual sensitivity to the environment, which is an intrinsic characteristic. Finally, we assume that the higher the sensitivity to the environment, the higher the marginal reaction to the regulator's announcement: $g_{a_{k,j}s^j}(a_{k,j} - \hat{a}|s^j) < 0$.

It is important to notice that, contrary to the implementation of the preceding tax policy, implementing a nudge does not require observable contributions. In particular, the nudge can be the social optimum to be reached, a mean contribution, or the maximum individual contribution whenever this information is available.

Finally, the topic we are concerned with can be related to Bernheim (1994)'s one. In his model, individuals make a balance between their intrinsic preferences and their *status*[16]. In particular, individuals want to be perceived as a good type (to have a high *status*). Akerlof (1997)'s work is also of interest. He considers two types of models. In the first one, individuals are concerned with their status. They seek for a different status compared to the others' one. In his second model, he considers conformist behavior. Individuals thus want to mimic other individuals' behavior, and incur a loss of utility if they depart from others' status.

In the next subsection, we assume that the regulator announces the social optimal contribution.

negative for individuals who contributed more than the mean level. In such a case, the boomerang effect, as discussed in the introduction, is also captured by $g_{a_{k,j}} \leq 0$.

[16]Defined as "popularity, esteem, or respect" (p. 843). Individuals are esteemed if they act in a way that is well perceived by the others.

3.2 Private optima

The regulator announces the socially optimal contribution $a^*_{k,j}$ to individual i, $i = (O,l),(O,h)$, $(o,l),(o,h)$. Each individual considers the following program:

$$\max_{a_{k,j}} \int_{\underline{e}}^{\overline{e}} \left(u(w - a_{k,j}) - d(P, s^j)\right) dH\left(F(e), \theta^k\right) - g(a_{k,j} - a^*_{k,j}|s^j) \tag{8}$$

$$s.t. \quad a_{k,j} \geq 0$$

The first order condition for a private interior solution $a^n_{k,j}$ is:

$$-u'(w - a^n_{k,j}) + b'(A) \int_{\underline{e}}^{\overline{e}} d_P(P, s^j) dH\left(F(e), \theta^k\right) - g_{a_{k,j}}(a^n_{k,j} - a^*_{k,j}|s^j) = 0 \tag{9}$$

The second order condition is satisfied.[17] We obtain the following results,

Proposition 3 *Consider a regulator who discloses the information about the social optimum to each individual.*

(i) All individual contributions increase following the introduction of the nudge.

(ii) The nudge permits to discriminate among individuals having different sensitivity although sensitivity is private information. In particular, if $a^p_{O,h} = a^p_{o,l}$ *and* $a^t_{O,h} = a^t_{o,l}$*, we also have* $a^n_{O,h} > a^n_{o,l}$.

From Proposition 3, a nudge may also induce higher contributions, such as a tax based on individual contributions. The first main difference is that a nudge does not ask for information about individual contributions contrary to an incentive tax policy.

[17]Indeed:

$$u''(w - a^n_{k,j}) + b''(A) \int_{\underline{e}}^{\overline{e}} d_P(P, s^j) h\left(F(e), \theta^k\right) f(e) de$$
$$-(b'(A))^2 \int_{\underline{e}}^{\overline{e}} d_{PP}(P, s^j) h\left(F(e), \theta^k\right) f(e) de - g_{a_{k,j}a_{k,j}}(a^n_{k,j} - a^*_{k,j}|s^j) < 0$$

Second, by comparing Point (ii) of Proposition 3 with Proposition 2 we conclude that a nudge can solve the issue raised by the fact that a tax cannot not be built on individual sensitivities.

However, it is important to notice that we obtain these results by ignoring the announcement issue for the nudge. Indeed, the reaction to the nudge could also depend on the confidence in the announcement. Thus, it could be possible that the difference in sensitivity compensates for the divergence in confidence. In that case, we could have that both individuals react the same and thus contribute the same.

Let us go further in the comparison between the tax and the nudge. Thanks to an integration by part, the first order condition (9) obtained with a nudge can be written:

$$\begin{aligned} u'(w - a^n_{k,j}) &= b'(A)\left(d_P(\overline{P}, s^j) - Y\int_{\underline{e}}^{\overline{e}} d_{PP}(P, s^j)dH\left(F(e), \theta^k\right)\right) \\ &\quad - g_{a_{k,j}}(a^n_{k,j} - a^*_{k,j}|s^j) \end{aligned} \tag{10}$$

with $\overline{P} = \overline{e}Y - b(A)$. Besides, thanks to an integration by part, the first order condition (7) under a tax policy can be written

$$-(1+t'(a^t_{k,j})).u'(w-a^t_{k,j}-t(a^t_{k,j}))+b'(A)\left(d_P(\overline{P}, s^j) - Y\int_{\underline{e}}^{\overline{e}} d_{PP}(P, s^j)dH\left(F(e), \theta^k\right)\right) = 0,$$

Or

$$\begin{aligned} u'(w - a^t_{k,j}) &= b'(A)\left(d_P(\overline{P}, s^j) - Y\int_{\underline{e}}^{\overline{e}} d_{PP}(P, s^j)H\left(F(e), \theta^k\right)de\right) \\ &\quad -t'(a^t_{k,j}).u'(w - a^t_{k,j} - t(a^t_{k,j})) + u'(w - a^t_{k,j}) - u'(w - a^t_{k,j} - t(a^t_{k,j})) \end{aligned} \tag{11}$$

In both conditions (10) and (11), the left-hand-side term and the first expression in the right-hand-side term are identical for a given level of individual contributions. Now, notice that in (11) the remaining terms (second line) do not depend on the environmental sensitivity s^j while the remaining term in (10) does so. The tax does not explicitely incorporate the subjective and personal characteristic s^j. On the contrary, the reaction $g(.)$ of individual i to the nudge depends on her environmental sensitivity: a nudge policy can explicitely build over this individual and personal characteristic although the regulator does not know who is who.

Furthermore, a nudge is an announcement and it does not require to collect money from a regulatory agency, contrary to a tax. Thus such a "soft" instrument is cheaper to implement than a traditionnal tax policy. Lastly, a nudge policy seems to be much more socially acceptable than a tax policy.

However, nudging is also concerned by some drawbacks. First, a nudge does not create additionnal financial credits contrary to a tax policy. There is no opportunity to benefit from some potential double dividend when considering a nudge. Second, a boomerang effect may be observed, while it does not exist with a tax since contributions beyond the social optimal are not taxed. Third, the question of manipulation[18] holds when disseminating information (Hausman and Welch (2010), Vallgarda (2012), Wilkinson (2013)). For Wilkinson (2013) and Hausman and Welch (2010), whether individuals' *autonomy*[19] of choices is maintained or not is the main feature that distin-

[18] According to Wilkinson (2013), an action is manipulative if: i) individuals' decision-making process is *perverted*, ii) this perversion is done by an *intentional actor*, and iii) the *autonomy* of choices is violated.

[19] Defined as the "control an individual has over his or her own evaluations and choices" (Hausman and Welch (2010, p.128)).

guishes paternalistic actions from non-paternalistic ones. Indeed, if individual A influences individual B in her decision, then individual B's autonomy of choices has been restricted, whatever the intention (good or bad motives). Hausman and Welch (2010) give two opposite examples of nudges. For instance, *rational persuasion* does respect individuals' autonomy because it relies on the presentation of evidence.[20] However, *shaping* may be seen as a manipulation because it may impact individuals' autonomy of choices by influencing them.[21] Vallgarda (2012) argues that nudges are meant to influence the automatic behavior[22]. In that case, it may be difficult for individuals not to act in the way they are influenced. If we compare with the implementation of a corrective tax, it is possible to pay the tax and to continue behaving as before the introduction of the tax. In our model, the nudge we consider is the disclosure of the social optimum. Thus, manipulation is not at stake for it fits with the presentation fo an *evidence* or an objective fact. Besides individuals' autonomy is not infringed.

Section 4 provides some simulations that illustrate our results.

4 Parameterized examples

In this section, we illustrate our results thanks to some simple simulations. We consider two individuals, with different environmental sensitivity and risk perception.

[20] A physician uses rational persuasion when asking a patient to stop smoking because she is developing a lung cancer.

[21] Shaping is a method similar to conditioning. It consists in decomposing the targeted task in subtasks. Each time a sub-task leading to the targeted task is complete, the individual is complimented or rewarded.

[22] Defined as the behavior "in which people engage without making conscious choices" (Vallgarda (2012, p.201)).

4.1 Benchmarck model

The utility of consumption for each agent (k, j) writes $u(w_{k,j}) = 100(w - a_{k,j}) - 10(w - a_{k,j})^2$ with $w_{k,j} = w - a_{k,j}$. Let us denote as a_1 and a_2 the contributions of individual 1 and 2 respectively without any external incentive. The public benefit of pollution reduction coming from the individual contributions is $b(A) = \sqrt{a_1 + a_2}$. The level of current pollution is $P = Y\tilde{e} - \sqrt{a_1 + a_2}$. The random variable of pollution $\tilde{e}$ follows a uniform distribution $F(e)$, with $\underline{e} = 0.05$ and $\bar{e} = 1$. The disutility coming from pollution for Agent (k, j) writes $d(P, s^j) = P^2 + 8e^{s^j}P$. We set $Y = 64$ and $w = 5$.[23] We also set $s^h = 1$ for an individual highly sensitive to the environment, and $s^l = 0$ for an individual less sensitive to the environment. We consider the following probability transformation function:

$$H[F(e), \theta^k] = \left(\frac{\bar{e}}{\underline{e}}\right)^{1-\theta^k} \times F(e)$$

Let us denote as a^* the social optimum. After computation of the social program, we obtain $a^* = 1.73$ for any agent as expressed by Point i) of Proposition 1.The different cases we consider are described below.

Case 1. Consider the individuals (O, l) and (o, h). We set $\theta^O = 0.6$ for the highly optimistic individual, and $\theta^o = 0.8$ for the less optimistic one. We obtain $a^p_{O,l} = 0.91$ and $a^p_{o,h} = 1.28$. Both contribute less than the social optimum computed with full information, $a^* = 1.73$, as predicted in Point ii) of Proposition 1.

Case 2. Consider the individuals (O, h), and (o, l). We keep the same values of θ and s^j as in Case 1. We obtain $a^p_{O,h} = 1.14$ and $a^p_{o,l} = 1.05$. Again, individuals

[23] With this parameters, the level of pollution is always positive.

contribute less than the social optimum ($a^* = 1.73$). We also obtain Point iii) of Proposition 1 with individual (O, h) contributing more than individual (o, l): the difference in environmental sensitivity more than compensates for the difference in risk perception.

4.2 Incentives

We now consider the implementation of a tax on the contributions that are not provided by comparison with the social optimum. Then we focus on a nudge based on the announcement of a^*. We focus on Point ii) of Propositions 2 and 3.

Case 3. Consider individuals (O, h) characterized by $\theta = 0.2$ and $s^h = 1$, and (o, l) characterized by $\theta = 0.45$ and $s^l = 0$. In the absence of outside incentives, we obtain $a^p_{O,h} = 0.83$ and $a^p_{o,l} = 0.88$, which is another possible situation displayed by Point iii) of Proposition 1. Thus comparing with Case 2, we obtain that it is not possible to rank *ex ante* the contributions of all agents with respect to their environmental sensitivity and to their risk perception.

Consider a linear tax function $t(a_{k,j}) = 0.25(a^* - a_{k,j})$, where $a^* = 1.35$. Both individuals increase their level of contribution: $a^t_{O,h} = 0.93$ and $a^t_{o,l} = 1.02$.

Instead of the tax, consider now a nudge which consists of announcing the individual social optimum $a^* = 1.35$ to each of both agents. The reaction to the nudge is formalized thanks to the fonction $g(a_{k,j} - a^*|s^j) = -\dfrac{(a_{k,j} - a^*)}{(a_{k,j} + 1)} \times (2 + e^{s^j})$. We find that $a^n_{O,h} = a^n_{o,l} = 0.94$. Finally, eventhough agents had different contributions, it is possible to make them contribute a same level as suggested by the social optimum. Knowing that the reaction ot the nudge is an intrinsic characteristic which depends on environmental sensitivity, it could be possible to even obtain the social optimum for highly sensitive (and too optimistic) individuals. With these illustrations, we do not oppose tax and nudge. We want to show that the nudge can perform as well as a tax. A combination of both could be even more welfare improving (see Ouvrard (2016)).

5 Conclusion

In this paper we have considered individuals who can voluntarily contribute to improve the quality of environment. They differ in their risk perception of pollution and in their environmental sensitivity. Risk perception relies on the confidence the individual has in the annoucement relative to the risk of pollution made by the regulator (a government for instance): an optimistic individual, in our setting, believes that the risk of pollution is lower than the one announced by the regulator. Thus she distorts the distribution of risk by underevaluating the risk of high pollution. Environmental sensitivity is an intrinsic characteristic of the individual that refers to either physical or psychological vulnerability to pollution.

Our first result is that risk perception alone cannot explain the difference between the level of contributions among different individuals contrary to what is studied in Etner et al. (2007) and in Salanié and Treich (2009). Environmental sensitivity must be taken into account.

We have compared the impact of a tax with the impact of a nudge on individual contributions. We have shown that a nudge may perform better than a tax under some fair assumptions: the reaction to a nudge depends directly on environmental sensitivity contrary to a tax. More precisely, a tax is an *unidimensional* instrument in the sense it depends only on the distance between individuals' level of contribution and the social optimum. On the contrary, the reaction to a nudge, which also depend on this distance, is more or less important depending on individuals' sensitivity to the environment. This is the second main result of this paper. We illustrated it with some simulations.

Nudging has also the advantage of being less demanding regard individual information. Computing the optimal tax rate requires that the regulator perfectly knows each individual environmental sensitivities and risk perception. Moreover, if the regulator wants to tax the distance to the social optimum, it is necessary to identify each individual contribution.

However, nudging may not have a persistent effect through time. This must still be experimentally analyzed, together with the use of a mixed instrument that would combine tax and nudge. As evoqued above, the questions of manipulation and ethic remain outstanding issues.

APPENDIX

Proof of Proposition 1

i) When the environmental sensitivity and the risk perception of all agents are known by the regulator, he chooses the level $a^*_{k,j}$ for Agent (k,j) that satisfies the following social program:

$$\max_{a_{k,j}} \int_{\underline{e}}^{\overline{e}} \left(u(w-a_{k,j}) - d(P,s^j)\right) dH\left(F(e),\theta^k\right) + \sum_{-i} \int_{\underline{e}}^{\overline{e}} \left(u(w-a_{-i}) - d(P,s^j_{-i})\right) dH\left(F(e),\theta^k_{-i}\right)$$

with a_{-i} the level of contribution of any other individual, with s^j_{-i} her environmental sensivity, θ^k_{-i} her type, and $-i \in I = \{(o,l);(o,h);(O,l);(O,h)\}$. The first order condition for an interior solution $a^*_{k,j}$ is

$$\begin{aligned} 0 &= -u'(w - a^*_{k,j}) + b'(A) \int_{\underline{e}}^{\overline{e}} d_P(P, s^j) dH\left(F(e), \theta^k\right) \\ &\quad + \sum_{-i} b'(A) \int_{\underline{e}}^{\overline{e}} d_P(P, s^j_{-i}) dH\left(F(e), \theta^k_{-i}\right) \end{aligned} \tag{12}$$

This expression can also be written as

$$0 = -u'(w - a^*_{k,j}) + \sum_{i} b'(A) \int_{\underline{e}}^{\overline{e}} d_P(P, s^j_i) dH\left(F(e), \theta^k_i\right)$$

It is the same for any agent so that their individual, socially optimal, contributions are identical to each other. This is Point i).

After an integration by parts of the second term of (12), it becomes

$$\begin{aligned} 0 = -u'(w - a^*_{k,j}) + b'(A) \left(d_P(\overline{P}, s^j) - Y \int_{\underline{e}}^{\overline{e}} d_{PP}(P, s^j) H\left(F(e), \theta^k\right) de \right) \\ + \sum_{-i} b'(A) \int_{\underline{e}}^{\overline{e}} d_P(P, s^j_{-i}) dH\left(F(e), \theta^k_{-i}\right) \end{aligned} \tag{13}$$

with $\bar{P} = \bar{e}Y - b(A)$. The third term corresponds to the public good effect. The private first order condition for an agent who distorts probabilities is given by (5). An integration by parts of its second term gives:

$$-u'(w - a^P_{k,j}) + b'(A) \left(d_P(\overline{P}, s^j) - Y \int_{\underline{e}}^{\overline{e}} d_{PP}(P, s^j) H(F(e), \theta^k) de \right) = 0 \tag{14}$$

Substracting expression (13) from (14), both being evaluated at the social level $a^*_{k,j}$, gives after simplification

$$-\sum_{-i} b'(A) \int_{\underline{e}}^{\overline{e}} d_P(P, s^j_{-i}) dH\left(F(e), \theta^k_{-i}\right)$$

which is negative. And finally $a^P_{k,j} < a^*_{k,j} \quad \forall j, \forall k$. This is Point ii).

For Point iii), considering (14) respectively for individuals (O,l) and (o,l) and substracting the former from the latter, both being evaluated at $a^p_{o,l}$, gives after simplification

$$-b'(A)Y \int_{\underline{e}}^{\overline{e}} d_{PP}(P)(H(F(e),\theta^o) - H(F(e),\theta^O))de,$$

which is positive since $H(F(e),\theta^o) < H(F(e),\theta^O)$, and $d_{PPs_j} = 0$ by assumption. Finally $a^p_{O,l} < a^p_{o,l}$.

Besides, substracting the first order condition of individual (o,h) from the first order condition of individual (o,l), both being evaluated at $a^p_{o,l}$, gives after simplification:

$$b'(A)\left(d_P(\overline{P},s^l) - d_P(\overline{P},s^h)\right)$$

which is negative since $d_{Ps^j}(\overline{P},s^j) > 0$ by assumption. Moreover, $d_P(\overline{P},s^l) - d_P(\overline{P},s^h)$ is constant because $d_{PPs_j} = 0$ by assumption. And finally $a^p_{o,l} < a^p_{o,h}$. The proof of $a^p_{O,l} < a^p_{O,h} < a^p_{o,h}$ is straightforward when following the steps of the previous proof. Finally, still using the same steps for $a^p_{o,l}$ and $a^p_{O,h}$, we are not able to rank them against each other. Proposition 1 is demonstrated. ♦

Proof of Proposition 2

i) Assume that $a^p_{o,l} = a^p_{O,h}$. According to point iii) of Proposition 1, this is possible if

$$b'(A)\left(d_P(\overline{P},s^l) - d_P(\overline{P},s^h) - Y\int_{\underline{e}}^{\overline{e}} d_{PP}(P)(H(F(e),\theta^o) - H(F(e),\theta^O))de\right) = 0$$

that is to say

$$d_P(\overline{P},s^l) - d_P(\overline{P},s^h) = Y\int_{\underline{e}}^{\overline{e}} d_{PP}(P)(H(F(e),\theta^o) - H(F(e),\theta^O))de$$

The first order condition of Agent (k, j) in the tax setting in given by

$$-(1+t'(a^t_{k,j})).u'(w-a^t_{k,j}-t(a^t_{k,j}))+b'(A)\left(d_P(\overline{P}, s^j) - Y\int\limits_{\underline{e}}^{\overline{e}} d_{PP}(P, s^j)dH\left(F(e), \theta^k\right)\right) = 0 \tag{15}$$

Substracting this first order condition for individual (O, h) from the first order condition for individual (o, l), both evaluated at $a^t_{o,l}$, gives

$$b'(A)\left(d_P(\overline{P}, s^l) - d_P(\overline{P}, s^h) - Y\int\limits_{\underline{e}}^{\overline{e}} d_{PP}(P)(H(F(e), \theta^o) - H(F(e), \theta^O))de\right)$$

By assumption $d_{PPs_j} = 0$, so that $d_P(P, s^j)$ is additive in P and s^j. Thus, we have that $d_P(\overline{P}, s^l) - d_P(\overline{P}, s^h)$ is constant in $a_{k,j}$ whatever (k, j).

If $d_{PPP} = 0$, then $Y\int\limits_{\underline{e}}^{\overline{e}} d_{PP}(P)(H(F(e), \theta^o) - H(F(e), \theta^O))de$ is constant in $a_{k,j}$. Thus

$$d_P(\overline{P}, s^l) - d_P(\overline{P}, s^h) = Y\int\limits_{\underline{e}}^{\overline{e}} d_{PP}(P)(H(F(e), \theta^o) - H(F(e), \theta^O))de$$

and $a^t_{o,l} = a^t_{O,h}$.

If $d_{PPP} < 0$, then the negative term $Y\int\limits_{\underline{e}}^{\overline{e}} d_{PP}(P)(H(F(e), \theta^o) - H(F(e), \theta^O))de$ is lower under the implementation of a tax than in the absence of outside incentives since individual contributions have increased in optimum. Thus, we have

$$0 > d_P(\overline{P}, s^l) - d_P(\overline{P}, s^h) > Y\int\limits_{\underline{e}}^{\overline{e}} d_{PP}(P)(H(F(e), \theta^o) - H(F(e), \theta^O))de$$

and $a^t_{o,l} > a^t_{O,h}$.

If $d_{PPP} > 0$, the result is reversed: $a^t_{o,l} < a^t_{O,h}$.

For Points ii) and iii), the proofs follow the same steps as in the proof of Point i).

♦

Proof of Proposition 3

i) Knowing that the individual, socially optimal, level of contribution is identical from on agent to another one (see Point i) of Proposition 1), we use the notation a^* to denote it. The first order condition for individual (k, j) under a nudge system is given by (10). Substracting (14) from (10), both being evaluated at $a^p_{k,j}$, gives after simplification

$$-g_{a_{k,j}}(a^P_{k,j} - a^*|s^j)$$

which is positive, by assumption. Finally $a^n_{k,j} > a^p_{k,j}$.

ii) Assume that $a^p_{O,h} = a^p_{o,l} = a^P$. From the previous proof, the difference between the first order condition of the initial setting and the one with the nudge, both measured at a^P, is $-g_{a_{k,j}}(a^P - a^*|s^h)$ for individual (O, h), and $-g_{a_{k,j}}(a^P - a^*|s^l)$ for individual (o, l).

We have by assumption that $g_{a_{k,j}s_j} < 0$ so that $-g_{a_{k,j}}(a_{k,j} - a^*|s^l) < -g_{a_{k,j}}(a_{k;j} - a^*|s^h)\ \forall a_{k;j}$. Finally, $a^n_{O,h} > a^n_{o,l}$. Proposition 3 is demonstrated. ♦

Optimal tax design

When the regulator knows individuals' environmental sensitivities and has some information on individual risk perception, but does not know who is who (as in Proposition 2), his program under taxation is:

$$\max_{a_{k,j}} \sum_i u(w - a_{k,j} - t^{SB}(a_{k,j})) - \sum_i \int_{\underline{e}}^{\overline{e}} \big(d(P, s^j)\big)\, dH(F(e), \theta^k)$$

with t^{SB} the second best tax function, with $i \in I = \{(o, l); (o, h); (O, l); (O, h)\}$.

The social first order condition for individual (k,j):

$$-(1+t'^{SB}(a^t_{k,j})).u'(w-a^t_{k,j}-t^{SB}(a^t_{k,j}))+\sum_i b'(A)\int_{\underline{e}}^{\overline{e}} d_P(P,s^j)dH(F(e),\theta^k)=0 \quad (16)$$

Rearranging, we obtain

$$t'^{SB}(a^t_{k,j})=\frac{\sum_i b'(A)\int_{\underline{e}}^{\overline{e}} d_P(P,s^j)dH(F(e),\theta^k)}{u'(w-a^t_{k,j}-t^{SB}(a^t_{k,j}))}-1 \quad (17)$$

which is negative from equation (16).

References

[1] Akerlof, G.A., 1997. "Social Distance and Social Decisions." *Econometrica*, 65: 1005–1027.

[2] Allcott, H., 2011. "Social Norms and Energy Conservation." *Journal of Public Economics*, 95: 1082–1095.

[3] Ayres, I., S. Raseman and A. Shih, 2013. "Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage." *Journal of Law, Economics and Organization*, 29: 992–1022.

[4] Banerjee, S., F.P. de Vries, N. Hanley and D.P. van Soest, 2014. "The Impact of Information Provision on Agglomeration Bonus Performance: An Experimental Study on Local Networks." *American Journal of Agricultural Economics*, 96: 1009–1029.

[5] Bernheim, B.D., 1994. "A Theory of Conformity." *Journal of Political Economy*, 102: 841–877.

[6] Cabinet Office Behavioural Insights Team, 2010. "Applying behavioural insight to health".

[7] Cabinet Office Behavioural Insights Team, 2011. "Behaviour change and energy use".

[8] Cabinet Office Behavioural Insights Team, 2013. "Applying Behavioural Insights to Organ Donation: preliminary results from a randomised controlled trial".

[9] Centre d'Analyse Stratégique (CAS), 2011. "Green nudges: new incentives for ecological behaviour. " Note d'Analyse 216.

[10] Chaudhuri, A., 2011. "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature." *Experimental Economics*, 14: 47–83.

[11] Cialdini, R.B., R.R. Reno and C.A. Kallgren, 1990. "A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places." *Journal of Personality and Social Psychology*, 58: 1015–1026.

[12] Costa, D.L. and M.E. Kahn, 2013. "Energy conservation "nudges" and environmentalist ideology: evidence from a randomized electricity field experiment." *Journal of the European Economic Association*, 11: 680–702.

[13] Cropper, M.L. and W.E. Oates, 1992. "Environmental Economics: A Survey." *Journal of Economic Literature*, 30: 675–740.

[14] Ebert, S. and G. van de Kuilen, 2015. "Measuring Multivariate Risk Preferences." Working paper.

[15] Ebert, S., L. Eeckhoudt L. and S. Spaeter, 2016. "Risk preferences when less is more."Mimeo.

[16] Eeckhoudt, L. and H. Schlesinger, 2006. "Putting Risk in Its Proper Place." *American Economic Review*, 96: 280–289.

[17] Egebark, J. and M. Ekström, 2016. "Can indifference make the world greener?" *Journal of Environmental Economics and Management*, 76: 1–13.

[18] Etner, J., M. Jeleva and P.A. Jouvet, 2007. "Risk Perceptions, Voluntary Contributions and Environmental Policy." *Research in Economics*, 61: 130–139.

[19] Etner, J., M. Jeleva and P.A. Jouvet, 2009. "Pessimism or Optimism: a Justification to Voluntary Contributions Toward Environmental Quality." *Australian Economic Papers*, 48: 308–319.

[20] Ferraro, P. and M.K. Price, 2013. "Using Non-Pecuniary Strategies to Influence Behavior: Evidence from a Large Scale Field Experiment." *The Review of Economics and Statistics*, 95: 64–73.

[21] Figuières, C., D. Masclet and M. Willinger, 2013. "Weak Moral Motivation Leads to the Decline of Voluntary Contributions." *Journal of Public Economic Theory*, 15: 745–772.

[22] Gaunt, M., T. Rye and S. Allen, 2007. "Public Acceptability of Road User Charging: The Case of Edinburgh and the 2005 Referendum." *Transport Reviews*, 27: 85–102.

[23] Goldstein, N.J., R.B. Cialdini and V. Griskevicius, 2008. "A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels." *Journal of Consumer Research*, 35: 472–482.

[24] Hahn, R.W. and R.N. Stavins, 2011. "The Effect of Allowance Allocations on Cap-and-Trade System Performance." *The Journal of Law and Economics*, 54: S267–S294.

[25] Harrison, G.W. and J.A. List, 2004. "Field Experiments." *Journal of Economic Literature*, 42: 1009–1055.

[26] Hausman, D.M and B. Welch, 2010. "To Nudge or Not to Nudge." *The Journal of Political Philosophy*, 18: 123–136.

[27] Helfand, G.E, P. Berck and T. Maull, 2003. "The theory of pollution policy." In K.-G. Mäler and J.R. Vincent (eds.). *Handbook of Environmental Economics*, 1, Elsevier Science.

[28] Hungerford, H.R. and T.L. Volk, 1990. "Changing Learner Behavior Through Environmental Education." *The Journal of Environmental Education*, 21: 8–21.

[29] Jaraité-Kažukauské, J. and A. Kažukauskas, 2015. "Do transaction costs influence firm trading behaviour in the european emissions trading system?" *Environmental and Resource Economics*, 62: 583–613.

[30] Kahn, M.E., 2007. "Do greens drive Hummers or hybrids? Environmental ideology as a determinant of consumer choice." *Journal of Environmental Economics and Management*, 54: 129–145.

[31] Kallbekken, S., S. Kroll and T.L. Cherry, 2011. "Do you not like Pigou, or do you not understand him? Tax aversion and revenue recycling in the lab." *Journal of Environmental Economics and Management*, 62: 53–64.

[32] Kimball, M.S., 1990. "Precautionary Saving in the Small and in the Large." *Econometrica*, 58: 53–73.

[33] Kotchen, M.J. and M.R. Moore, 2008. "Conservation: From Voluntary Restraint to a Voluntary Price Premium." *Environmental and Resource Economics*, 40: 195–215.

[34] Lattimore, P.K., J.R. Baker and A.D. Witte, 1992. "The influence of probability on risky choice." *Journal of Economic Behavior and Organization*, 17: 377–400.

[35] Ledyard, J., 1995. "Public Goods: A Survey of Experimental Research." In J. H. Kagel and A. E. Roth (eds.). *The Handbook of Experimental Economics*, 111–194, Princeton, NJ: Princeton University Press.

[36] Nolan, J., W. Schultz, R.B. Cialdini, N. Goldstein and V. Griskevicius, 2008. "Normative Influence is Underdetected." *Personality and Social Psychology Bulletin*, 34: 913–923.

[37] Organisation for Economic Cooperation and Development (OECD), 2012. "Behavioural Economics and Environmental Policy Design." Project Description, Empirical Policy Analysis Unit.

[38] Ouvrard, B., 2016. "Persistence of the effects of monetary and non-monetary incentives: an experimental approach." Mimeo.

[39] Prelec, D., 1998. "The Probability Weighting Function." *Econometrica*, 66: 497–527.

[40] Quiggin, J., 1982. "A Theory of Anticipated Utility" *Journal of Economic Behavior and Organization*, 3: 323–343.

[41] Salanié F. and N. Treich, 2009. "Regulation in Happyville." *The Economic Journal*, 119: 665–679.

[42] Sartzetakis, E.S., 2004. "On the efficiency of competitive markets for emission permits." *Environmental and Resource Economics*, 27: 1–19.

[43] Scheier, M.F. and C.S. Carver, 1985. "Optimism, Coping, and Health: Assessment and Implications of Generalized Outcome Expectancies." *Health Psychology*, 4: 219–247.

[44] Scheier, M.F. and C.S. Carver, 1992. "Effects of Optimism on Psychological and Physical Well-Being: Theoretical Overview and Empirical Update." *Cognitive Therapy and Research*, 16: 201–228.

[45] Schultz, W., J. Nolan, R.B. Cialdini, N. Goldstein and V. Griskevicius, 2007. "The Constructive, Destructive, and Reconstructive Power of Social Norms." *Psychological Science*, 18: 429–434.

[46] Siegrist, M. and G. Cvetkovich, 2000. "Perception of Hazards: The Role of Social Trust and Knowledge.", *Risk Analysis*, 20: 713-719.

[47] Sinclair-Desgagné, B. and E. Gozlan, 2003. "A theory of environmental risk disclosure.", *Journal of Environmental Economics and Management*, 45: 377–393.

[48] Slovic, P., 2013. *The feeling of risk. New Perspectives on Risk Perception.* Earthscan Publications.

[49] Stavins, R.N., 1995. "Transaction Costs and Tradeable Permits." *Journal of Environmental Economics and Management*, 29: 133–148.

[50] Thalmann, P., 2004. "The public acceptance of green taxes: 2 million voters express their opinion." *Public Choice*, 119: 179–217.

[51] Thaler, Richard H. and Sustein Cass R., 2009. *Nudge: Improving Decisions about Health, Wealth, and Happiness.* Penguin Books.

[52] The European Commission, 2016. "Behavioural Insights Applied to Policy." The European Commission, URL: http://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/behavioural-insights-applied-policy-european-report-2016?search

[53] Thoyer, S., R. Préget, L. Kuhfuss, P. Le Coënt, F. Gautier-Pelissier, J. Subervie, L. Ibanez, M. Désolé, M. Tidball, 2015."Comportement et usage des pesticides: pour des contrats environnementaux innovants (Coud'pouce)", rapport final, programme Pesticides, APR 2011, 53 p.

[54] Vallgarda, S., 2012. "Nudge - A new and better way to improve health?" *Health Policy*, 104: 200–203.

[55] Wilkinson, T.M, 2013. "Nudging and Manipulation." *Political Studies*, 61: 341–355.

[56] Wu, G. and R. Gonzalez, 1999. "On the shape of the Probability Weighting Function." *Cognitive Psychology*, 38: 129–166.