

RESEARCH ARTICLE

Environmental incentives facing private information

Franz Wirl 

Faculty of Business, Economics and Statistics, University of Vienna, Wien, Austria
Corresponding author. E-mail: franz.wirl@univie.ac.at

(Submitted 20 January 2020; revised 12 October 2020; accepted 1 February 2021; first published online 22 March 2021)

Abstract

Environmental incentives are characterized by two distinct features: (1) a benefit-cost trade-off; and (2) private information about the trade-off. This suggests a degree of freedom of where to attach the private information, either to the benefit or the costs, as long as these choices imply the same behavior absent incentives (*‘observation equivalent’*). However, we show that different *observation equivalent* specifications can lead to different incentives. This is demonstrated for two cases: rainforest protection and contributions to a public good. Therefore, the choice of a private information parameter must be justified against observation equivalent alternatives.

Keywords: conservation; incentives; observation-equivalence; principal-agent model

JEL classification: D86; J33

1. Introduction

Given the many environmental issues and their increasing importance, initiatives exist and are proposed globally, regionally, between nations, nationally and to individuals. The first objective of this paper is to investigate how private information affects environmental incentives in different fields. Private information is pervasive in many if not most cases in which incentives are offered to address environmental objectives. However, the consequences of private information on the design of incentives and policies are ignored to the detriment not only of economic efficiency but also of the environment. Examples range from joint implementation (compare Wirl *et al.*, 1998), to the related clean development mechanism, to energy conservation programs (as a topical example, consider the European Union which issues white certificates to utilities in order to trigger conservation, i.e., ‘the butcher is asked to sell fish’ to quote Ruff (1988); see Wirl (2015)). And these are only examples related to climate change.

The second, formal but presumably more important, contribution is to draw attention to the following possibility, which is overlooked in most if not all related applications: different specifications of an agent’s trade-off, which are *observation-equivalent* absent

incentives, can lead to different incentives under private information. This is particularly important for environmental issues, because they are characterized by trade-offs between economic costs and environmental benefits. This is in contrast to labor and procurement incentives in which the agents face only costs and thus no trade-offs.

2. Framework

2.1. Set up of the objectives

An environmental action (x) affects the weakly concave payoff of a principal, $V(x)$. The action x can be either a good one – e.g., abatement of emissions, protection of the rainforest, wildlife or energy conservation – or a bad one – e.g., hunting, logging and emitting. The principal (a ‘she’) is interested in reducing the environmental harm, i.e., increasing ‘good’ but decreasing ‘bad’ actions that are carried out by an agent (‘he’). The agent has, however, a different objective (W) consisting of benefit, $b(x)$ increasing and weakly concave, and cost, $c(x)$ increasing and convex, and has private information, either about the level of his cost (t) or his benefit (θ).

Although one could define all actions with respect to their environmental benefit by introducing ($-x$) for harmful actions, we suggest allowing for x being beneficial or harmful for the environment depending on the problem. However, to include the many different applications within we suggest the following Assumptions:

- 1 The private information parameters scale the cost and benefit levels with the means $Et = E\theta = 1$ so that the functions $b(x)$ and $c(x)$ correspond to the average types, $t = 1 = \theta$.
- 2 Although the principal does not know either t or θ , she knows the cumulative probability distribution $F(t)$, with the support $t \in [t, \bar{t}]$. The implied hazard rate is assumed to be increasing (as in the entire principal-agent literature). Analogously for $G(\theta)$, $\theta \in [\underline{\theta}, \bar{\theta}]$. As a consequence, the principal can only maximize her expected payoff by offering a subsidy s , free of deadweight loss, for the action x , i.e.,

$$\max_{\{x(\omega), s(\omega), \omega \in \Omega\}} \int_{\Omega} (V(x(\omega)) - s(\omega)) d\Phi, \quad \omega = t \text{ or } \theta, \quad \Phi = F \text{ or } G. \quad (1)$$

- 3 The principal prefers high types. Assuming a beneficial action (i.e., $V' > 0$), this implies for the agent’s payoffs,

$$W^t(x, t) = b(x) - \frac{c(x)}{t}, \quad (2)$$

$$W^\theta(x, \theta) = \theta b(x) - c(x), \quad (3)$$

depending on which parameter is the agent’s private information. Therefore, both, $W^t_{xt} > 0$ and $W^\theta_{x\theta} > 0$ satisfy Mirrlees’ single crossing property. If the action is harmful, i.e., $V' < 0$, then, t and θ switch sides (in order to retain the ordering that high types are more efficient for the principal),

$$W^t(x, t) = b(x) - tc(x), \quad (4)$$

$$W^\theta(x, \theta) = \frac{b(x)}{\theta} - c(x), \quad (5)$$

and the single crossing property holds for this case too, but this time for the opposite inequalities, $W_{xt}^t < 0$ and $W_{x\theta}^\theta < 0$.

- 4 Given the scaling in point 3 above, both private information parameters are identically distributed,

$$F(t) = G(\theta) \forall t = \theta.$$

- 5 The uniform distribution is assumed, which must, due to the assumption in point 1, be symmetric around $Et = 1$ with the spread $\pm\delta$,

$$F(t) = \frac{t - \underline{t}}{\bar{t} - \underline{t}}, \quad t \in [1 - \delta, 1 + \delta] \quad f(t) := \dot{F}(t) = \frac{1}{\bar{t} - \underline{t}} = \frac{1}{2\delta},$$

so that the hazard rate (h) is increasing,

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{1}{1 + \delta - t}. \tag{6}$$

The dots denote total derivatives with respect to the private information parameter, either t or θ .

None of the above assumptions is crucial. Obviously, normalization does not affect the results. Assumption 2 is standard. Assumption 3 ensures the same directions for the mechanisms (from ‘the left to the right’), i.e., moving from inefficient low types to more efficient high types irrespective of whether the action is a good or bad one. Accounting for deadweight losses is straightforward within the chosen framework (solving optimal control models) but is omitted in order to focus on how private information and how different but *observation equivalent* setups affect incentives. The linear (and reciprocal) scaling of either benefits or costs is not restrictive, because all results hold for other monotonic scales (e.g., logarithmic, reciprocal expressing types in terms of their inefficiency, etc.) too. Of course, the assumption of the uniform distribution simplifies all calculations but is not crucial either.

Proposition 1. *Given this setup and no incentives, both formulations of the agent’s payoff, W^t and W^θ , are observation equivalent with respect to attaching the private information parameter either to the cost (t) or the benefit (θ). More precisely, $x_0(t) = x_0(\theta)$ for all $t = \theta$.*

Proof: Assume an environmentally benign action (i.e., either objective equations (2) or (3)) and the optimal action of agent of type t

$$x_0(t) = \arg \max_x b(x) - \frac{c(x)}{t},$$

then the following necessary and sufficient optimality condition must hold,

$$tb' = c'.$$

This condition characterizes also the choice of the type θ ,

$$x_0(\theta) = \arg \max_x \theta b(x) - c(x) \implies \theta b' = c'.$$

Since $\theta = t \implies x_0(t) = x_0(\theta)$. Analogous for an action that harms the principal’s environmental objective V and thus the objectives equations (4) and (5). □

Remark 1. *The distribution F (given the identity between F and G , the analysis can be restricted to a single distribution) allows for two interpretations for equation (1). First, maximizing the expected value of the principal's payoff from incentivizing a single agent. This applies to one of the examples below (protecting a rainforest). Second, the principal maximizes her payoff over an aggregate of agents with individually unknown types; this interpretation is appropriate for the other example of voluntary contributions of many to a public good.*

2.2. Examples

Two concrete problems, one with negative (logging rainforests) and the other with positive (contributing to a public good) actions, are used. Given the focus of the paper, how private information and how different but *observation equivalent* setups affect incentives, the examples – the protection of a rainforest is based on Harstad and Mideksa (2017) and the contribution to a public good affected by warm glow (Andreoni, 1990) is based on Wichman (2016) – are taken from the literature but modified in order to account for private information.

2.2.1. Contributions to a public good

Private contributions can be important on at least two grounds. NGOs are important lobbyists for the environment and they depend on private donations but also on public money, at least in Europe – and that to a quite significant extent. A government may use private donations as a gauge for its population's willingness to pay for a particular project and therefore how much to back it up or to motivate the population to contribute (e.g., by offering tax deductions). Donations are collected and given to victims of earthquakes, floods, other natural disasters (at home and abroad) and of wars and conflicts, and recently to refugees, etc.

Wichman (2016) considers voluntary contributions to a public good (x) that deliver the benefit of a warm glow (Andreoni, 1990). Ma and Burton (2016) find evidence of the warm glow for Australian electricity consumers and the response of many to the refugee crisis in 2015 is a confirmation of the importance of warm glow against own economic interest. The principal's benefit is linear in the (total) contributions, and the principal values the donation of one dollar more than the most empathic agent,

$$V = vx, \quad v > \bar{\theta} = \bar{t} > 1. \quad (7)$$

Therefore, the action x is a good one. Following Remark 1, we use the plural, agents, in the case of voluntary contributions to a public good. Incentives are in the form of subsidies, tax breaks and the promise to double one's contribution. However, Benabou and Tirole (2006) warn that monetary incentives can be counterproductive if they lower the value of the signal of showing one's green or social attitude.

Transforming the setup in Wichman (2016) into a principal-agent problem requires abstracting from free riding. An agent's benefit from his contribution is to enjoy the warm glow,

$$b(x) = wx, \quad w = 1,$$

linear for simplicity and normalized (to avoid unnecessary parameters). In an environmental context, x may reflect a costly but environmentally friendly action, such as buying an electric car, installing photovoltaic panels, and donating to wildlife protection. Such actions do not only deliver a warm glow but could offer other benefits such as fuel saving

and signalling an investor’s green attitude. That is, b includes all benefits and not only the warm glow. However, this benefit is not free due to the associated loss of hedonistic (logarithmic) utility given his wage income l (identical across types),

$$c(x) = \beta [\ln(l) - \ln(l - x)].$$

Assuming that the cost parameter t is an agent’s private information and dropping the constant term ($\beta \ln(l)$), an agent has the following objective:

$$W(x, t) = x + \frac{\beta \ln(l - x)}{t}, \tag{8}$$

in which $(1/t)$ measures the loss of hedonistic utility so that the principal prefers higher (i.e., less greedy) types. The superscript used in equation (2) is dropped because the argument (here t) is given explicitly; this applies also to the case of θ below.

Alternatively, the private information parameter may be attached to the benefit, i.e., to the warm glow. Then an agent has the objective:

$$W(x, \theta) = \theta x + \beta \ln(l - x), \tag{9}$$

and the application of proposition 1 yields the claimed observation equivalence,

$$x_0(t) = \max \{0, l - \beta/t\} = x_0(\theta) = \max \{0, l - \beta/\theta\}, \tag{10}$$

after accounting for the non-negativity constraint, $x \geq 0$, because contributions cannot turn negative. It seems a priori unclear which of the two specifications, equations (8) or (9), one should favor.

2.2.2. Rainforest

The protection of tropical forests mitigates climate change (as a sink for negative emissions) and maintains biodiversity. Incentives are only successful if (financial) benefits from conservation measures are larger than costs for relevant decision-makers (Kremen *et al.*, 2000). Incentives include support to eco-tourism, the compensation of willing landowners for not developing their land and for losses from wildlife depredation.¹

Harstad and Mideksa (2017) investigate how donors can protect natural habitats. A principal is concerned about conservation, say Norway’s government about a tropical rainforest of original size R and an agent who enjoys the benefits from logging (x) and from the remaining parts of the rainforest ($R - x$). Although different conservation projects (such as wetlands or wildlife) fit the framework, the reference will be to a rainforest. Norway’s benefit from protecting a particular patch of a rainforest (V) is presumably linear in the protected size given the many options including those in other countries,

$$V = v(R - x), \quad v > 0. \tag{11}$$

Clearly, the agent’s action, logging, is a negative one in the sense of harming the principal’s interest ($V' < 0$). The agent’s benefit from logging is linear if he is a price taker

¹ According to the The Economist (2013), a study by the University of Chicago’s Field Museum of Natural History estimates that lions on average kill livestock worth \$270 a year but bring in \$17,000 from tourists. Compensating owners for livestock lost to lions may have reduced locals’ incentive to look after their herds.

(p denotes the profit per unit),

$$b(x) = px, \quad p > 0,$$

but logging is also costly if the agent appreciates the remaining rainforest. Assuming a standard (logarithmic) utility function (Harstad and Mideksa (2017) assume linear relations and ignore private information), the recipient country's convex cost from logging is

$$c(x) = \ln(R) - \ln(R - x),$$

due to the loss associated with a shrinking rainforest. Combining and dropping the constant term ($\ln(R)$) yields the agent's objective in terms of his private information about the costs,

$$W(x, t) = px + t \ln(R - x). \tag{12}$$

The cost parameter t enters now linearly so that the principal again prefers high types since logging is an action that harms the principal. The superscript t (and θ below) is again dropped because the argument t is made explicit in W .

Alternatively, assume that the principal knows the agent's valuation of the rainforest but not his need for cash from timber exports, which leads to the agent's objective:

$$W(x, \theta) = \frac{1}{\theta}(px) + \ln(R - x), \tag{13}$$

and a high value of θ indicates less need for cash. Proposition 1 implies observation equivalence,

$$x_0(t) = R - \frac{t}{p}, \quad x_0(\theta) = R - \frac{\theta}{p}, \tag{14}$$

and logging x_0 remains positive as long as $t < pR$.

The crucial difference from the other example is that it seems easier to choose the private information parameter, namely the developing country's evaluation of its rainforests, i.e., the cost parameter t .

3. Optimal incentive contracts

3.1. Contributing to a public good

3.1.1. Cost is private information

$$U(\hat{t}, t) := W(x(\hat{t}), t) + s(\hat{t}) \tag{15}$$

defines the net payoff of an agent of type t pretending the type \hat{t} using $W = W^t$ from equation (2) but dropping the superscript since the argument t is explicitly indicated. The revelation principle allows us to restrict the analysis to incentives that 'force' the agent to tell the truth. This adds the inequality,

$$U(\hat{t}, t) \leq U(t, t) =: U(t) \quad \forall t, \hat{t} \in [\underline{t}, \bar{t}], \tag{16}$$

to the principal's optimization problem equation (1).

Assuming that the cost parameter t , which scales the agent’s consumption benefit, is private information, the principal’s optimal offer can be derived by solving the following optimal control problem with the constraints equations (18) and (19):

$$\max_{\{x(t) \geq 0\}} \int_{\underline{t}}^{\bar{t}} \left[vx(t) + x(t) + \frac{\beta \ln(l - x(t))}{t} - U(t) \right] dF(t), \tag{17}$$

$$\dot{U}(t) = W_t = -\frac{\beta \ln(l - x(t))}{t^2}, \tag{18}$$

$$U(t) \geq U_0(t) := \max \left\{ \frac{\beta \ln(l)}{t}, l - \frac{\beta}{t} + \frac{\beta \ln\left(\frac{\beta}{t}\right)}{t} \right\} \quad \forall t \in [\underline{t}, \bar{t}]. \tag{19}$$

The objective equation (17) results from subtracting the subsidies, $s = U - W$ due to the definition in equation (15) using equation (2) for W , from the principal’s gross payoff V from equation (7). The control constraint, $x(t) \geq 0$, stipulates that only non-negative contributions are possible. The differential equation (18) results from applying the envelope theorem to compute the total derivative of $U(t)$ from equation (16) and is called the *incentive compatibility constraint*; recall that dots denote total derivatives with respect to the private information parameter. Although equation (18) is only derived from the first order condition that truth telling is optimal, we verify that it is also sufficient. The last constraint equation (19) is called the *participation* (or *individual rationality*) *constraint* and it requires that the principal’s offer must exceed the agent’s outside option, the reservation price U_0 , which is given by substituting x_0 from equation (10) into equation (19); the two cases in equation (19) reflect that an agent may contribute, $x_0 > 0$, or not, $x_0 = 0$. While most principal-agent models assume a constant (and normalized to 0) reservation price, U_0 here is type dependent even if $x_0 = 0$ because $\dot{U}_0 = -(\beta \ln(l - x_0)/t^2) \neq 0$ and possibly declining as well as $U(t)$.

Setting up the Hamiltonian (H , λ denoting the costate and dropping the argument t) in general terms,

$$H = [V + W - U]f + \lambda W_t,$$

the first order optimality conditions for interior solutions are

$$(V_x + W_x)f + \lambda W_{tx} = 0, \tag{20}$$

$$\dot{\lambda} = f, \quad \lambda(\bar{t}) = 0. \tag{21}$$

Integrating equation (21) yields $\lambda = F - 1$, and substituting it into equation (20) yields

$$V_x + W_x = \frac{W_{xt}}{h}. \tag{22}$$

Equating the left hand side in equation (22) to zero determines the first best,

$$x^1 := \arg \max_x V + W, \tag{23}$$

from maximizing the joint objective of principal and agent (and ignoring possible constraints, such as non-negativity). The right hand side is positive (given a positive action,

as in the case of private contributions to public good, $W_{xt} = (c/t^2) > 0$) and determines the agency cost that leads to an allocation below the first best except at $t = \bar{t}$, at which point the right hand side in equation (22) vanishes due to $(1/h(\bar{t})) = 0$. Thus, $x(\bar{t}) = x^1(\bar{t})$, which is known as *no distortion at the top*. The solution of equation (22), denoted x^r , is called the *relaxed program* (see Fudenberg and Tirole, 1992), thus the use of the superscript r . It is optimal only if it satisfies all constraints: of the control (non-negativity), of participation equation (19) and of monotonicity, here $\dot{x} \geq 0$, since the principal wants higher contributions from the agent.

Applying the condition equation (22) to the objectives equations (7) and (8),

$$v + 1 - \frac{\beta}{t(l-x)} = \frac{\beta(\bar{t}-t)}{t^2(l-x)}, \tag{24}$$

allows us to compute the interior (= relaxed program) solution,

$$x^r(t) = l - \frac{\beta\bar{t}}{t^2(1+v)}. \tag{25}$$

Therefore, positive contributions, $x^r(t) > 0$, result only for the types

$$t > t^{\min} := \sqrt{\frac{\beta\bar{t}}{l(1+v)}}. \tag{26}$$

And once the relaxed program is positive, it always exceeds the agents' own contributions, but remains below the first best (except at the top),

$$x_0 \leq x^r \leq x^1 = \max \left\{ 0, l - \frac{\beta}{t(1+v)} \right\}.$$

The type t^{\min} determines the marginal type, i.e., only the types $t \geq t^{\min}$ participate. The reason is that U cuts U_0 at $t = t^{\min}$ from below,

$$\dot{U}(t) = -\frac{\beta \ln(l - x^r(t))}{t^2} > \dot{U}_0(t) = -\frac{\beta \ln(l)}{t^2}, \quad t > t^{\min} \text{ and close to } t^{\min},$$

so that the participation constraint equation (19) is met first locally and then for all $t > t^{\min}$ due to $x^r > x_0$.

Proposition 2. *The optimal mechanism consists of the relaxed program $x^r(t)$ from equation (25) coupled with the boundary solution (x_0 and thus no contribution),*

$$x(t) = \begin{matrix} 0 & \leq \\ & \text{if } t \in \max \{ \underline{t}, t^{\min} \} \\ l - \frac{\beta\bar{t}}{t^2(1+v)} & < \end{matrix}, \tag{27}$$

and the subsidies assuming $t^{\min} > \underline{t}$,

$$s(t) = \begin{matrix} 0 & \leq \\ & \text{if } t \in t^{\min} \\ \frac{2\sqrt{\beta l(1+v)}/t}{\sqrt{\bar{t}}} + \frac{\beta\bar{t}}{t^2(1+v)} - l - \frac{2\beta}{t} & > \end{matrix}. \tag{28}$$

Integrating the incentive compatibility constraint equation (18) yields the payoff of an agent of type t ,

$$U(t) = U_0(t^{\min}) + \int_{t^{\min}}^t W_z(x^r(z), z) dz$$

$$= \frac{\beta(\ln((\beta\bar{t}/(1 + \nu)t^2)) - 2)}{t} + \frac{2\sqrt{\beta(l(1 + \nu))}}{\sqrt{t}},$$

and substituting,

$$W(x^r(t), t) = l + \frac{\beta \ln((\beta\bar{t}/(1 + \nu)t^2))}{t} - \frac{\beta\bar{t}}{t^2(1 + \nu)},$$

determines the subsidies equation (28) if $t^{\min} > \underline{t}$, because $s = U - W$.

Figure 1 shows the contract in a schematic form although it is based on the numerical example used in figure 2 for a comparison with the alternative representation equation (9) of the agents' payoffs. The panel at the top shows how the contributions depend on the type: $x(t)$ is increasing and concave. Absent incentives, only a minority contributes and only a little (x_0). This may justify the incentive because it increases participation as well as contributions; for completeness, the first best (x^1) is included. The panel on the bottom left side shows the reservation price (U_0) and the agents' payoffs, if incentivized (U) but only for the participating types, $t \geq t^{\min}$; note that both are declining but U less than U_0 for $t \geq t^{\min}$ so that the participation constraint is met (as claimed). Replacing in equation (28) by solving equation (24) for

$$t = \sqrt{\frac{\beta\bar{t}}{(1 + \nu)(l - x)}},$$

yields subsidies depending on the contributions (i.e., $s(x)$, which is suppressed due to its cumbersome expression), increasing and convex.

3.1.2. Benefit – that is, the warm glow – is private information

Assuming that the agents' payoffs are given by equation (9) and following the above outlined procedure, the optimal contract can be derived by solving the following optimal control problem:

$$\max_{\{x(\theta) \geq 0\}} \int_{\underline{\theta}}^{\bar{\theta}} [vx(\theta) + \theta x(\theta) + \beta \ln(l - x(\theta)) - U(\theta)] dF, \tag{29}$$

$$\dot{U}(\theta) = W_\theta = x(\theta) > 0, \tag{30}$$

$$U(\theta) \geq U_0(\theta) := \max \left\{ \beta \ln(l), l\theta - \beta + \beta \ln \left(\frac{\beta}{\theta} \right) \right\} \forall \theta \in [\underline{\theta}, \bar{\theta}]. \tag{31}$$

The maximization of the expected value of all contributions over and above the subsidies, i.e., the objective equation (29), is constrained by incentive compatibility equation (30) and individual rationality equation (31) in which the reservation price $U_0(\theta)$ is again type dependent but this time increasing, $\dot{U}_0 = x_0 > 0$, except at the boundary $x_0 = 0$.

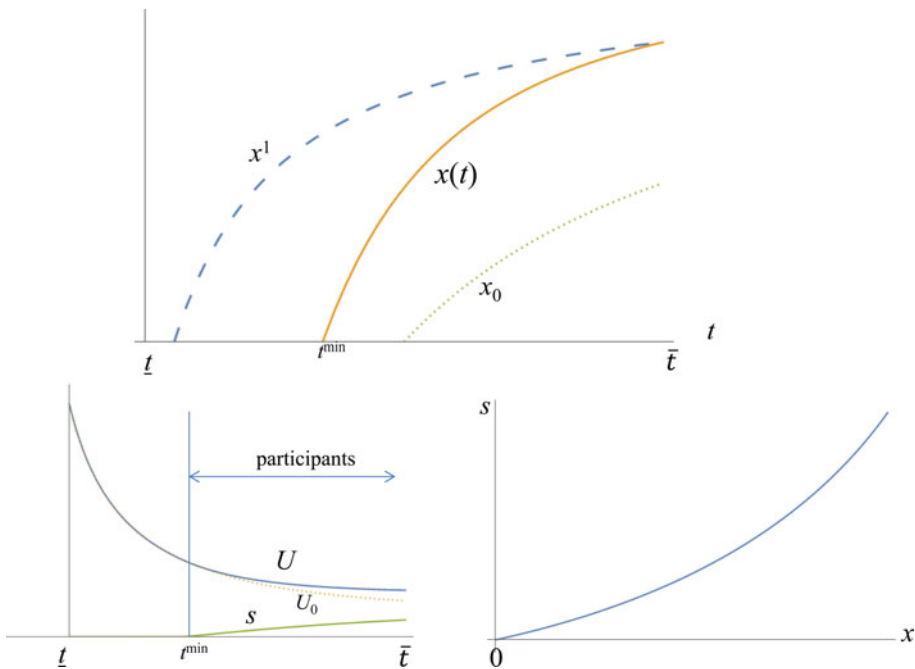


Figure 1. Optimal incentives in order to increase private contributions to a public good if the cost parameter t is private information.

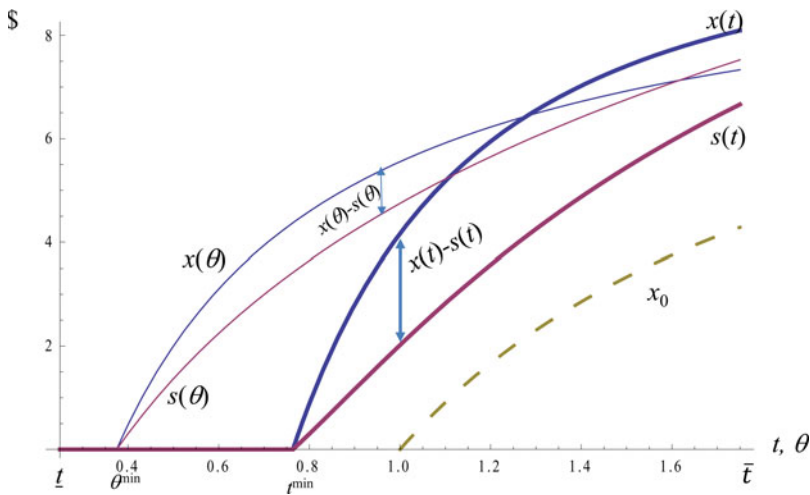


Figure 2. Comparing the contributions to a public good and the necessary subsidies if either the cost parameter (t , bold, measuring the consumption benefit) or the benefit parameter (θ about the warm glow) is private information, $v = 2$, $\beta = 10$, $l = 10$.

Expressing the contracting problems as optimal control problems reveals some differences in spite of the equivalence out of contract: first, between the integrands, objective equation (29) versus equation (17); second, different ‘dynamic’ constraints resulting from the incentive compatibility constraints, equation (30) versus equation (18); and third, different path constraints because of the different reservation prices, equation (31) versus equation (19).

The derivation is as in section 3.1.1. Therefore, we can start with the computation of the relaxed program (using again the superscript r) from (22) using the specification equation (8),

$$v + \theta - \frac{\beta}{l - x} = \bar{\theta} - \theta \implies x^r(\theta) = l - \frac{\beta}{v - (\bar{\theta} - 2\theta)} \leq x^1 = l - \frac{\beta}{v + \theta}. \quad (32)$$

The contributions are of course again below the first best (except at the top, $\theta = \bar{\theta}$) and cannot exceed the wage (l). Hence, $v + 2\theta > \bar{\theta}$ for which $v > \bar{\theta}$ is sufficient, i.e., the principal always values the contribution more than the agent with the highest warm glow (actually, $v > \bar{\theta} - \theta$ is sufficient). Given this sufficient condition, the relaxed program exceeds the unincentivized contributions but must in addition be positive,

$$x^r(\theta) = l - \frac{\beta}{v - (\bar{\theta} - 2\theta)} \geq 0 \iff \theta \geq \theta^{\min} := \frac{\bar{\theta} - v}{2} + \frac{\beta}{2l}, \quad (33)$$

so that only higher types, $\theta \geq \theta^{\min}$, contribute at least if incentivized (assuming $\theta^{\min} > \theta$). And indeed, the intersection of the relaxed program with the abscissa at θ^{\min} determines the marginal type since $\dot{U}(\theta) = x^r > \dot{U}_0(\theta) = x_0$ for all $\theta \geq \theta^{\min}$ ensures that the participation constraint holds for all $\theta \geq \theta^{\min}$. Therefore, the subsidy policy enlarges the set of contributing types since $x_0(\theta) > 0 \iff \theta > \beta/l > \theta^{\min}$, compare equation (9).

Proposition 3. *Assuming $\theta^{\min} > \theta$, the optimal contract asks the types $\theta \geq \theta^{\min}$ from equation (33) to contribute $x^r(\theta)$, increasing with respect to the type θ , from equation (32) and pays the subsidies,*

$$s(\theta) = \left\{ \begin{array}{l} 0 \\ \frac{1}{2} \left[l(v - \bar{\theta}) - \frac{\beta}{v + 2\theta - \bar{\theta}} + \beta \ln \frac{l(v + 2\theta - \bar{\theta})}{\beta} \right] \end{array} \right. \text{ for } \theta \begin{array}{l} < \\ \geq \end{array} \left. \right\} \theta^{\min}, \quad (34)$$

which are increasing and convex with respect to types and contributions,

$$\sigma(x) = (v - \bar{\theta})x + \beta \ln \frac{l}{l - x} \text{ if } x > 0. \quad (35)$$

This mechanism includes more types than its counterpart equation (2), i.e.,

$$t^{\min} > \theta^{\min} \iff l \geq \frac{\beta(1 + v - 2\bar{t})}{(v - \bar{t})(1 + v)} \text{ since } \theta = t. \quad (36)$$

Inequality equation (36) depends on the wage level and the following corollary gives a sufficient condition for $t^{\min} > \theta^{\min}$ that renders the numerator in equation (36), $(1 - \bar{t}) + (v - \bar{t})$, negative.

Corollary 1. $t^{\min} > \theta^{\min}$ if the range of the agents' types, $\delta = \bar{t} - 1$, exceeds the difference between the principal's and the agents' maximal valuations of their contributions, $v - \bar{t}$.

Figure 2 compares the agents' voluntary contributions without incentives (x_0) and with incentives depending on which parameter, t or θ , is treated as private information and includes the necessary subsidies. The first and crucial observation is that the incentives differ across the two representations of the agents' objectives equations (8) and (9) although both descriptions imply the same actions in the absence of incentives (*observation equivalent*). Incentives based on private information about the warm glow (θ) include more participants but the induced contributions over and above the necessary subsidy are very small. Indeed, subsidies can even exceed an agent's contribution (e.g., for all types $\theta > 1.62$ in figure 2). It is hard to recommend such incentive schemes in particular if one accounts in addition for deadweight costs of public funds. However, the principal still benefits because she values the contribution above the agent's warm glow. If the consumption benefit is private information (t), participation is more selective. Although it does not yield higher (except for the high types and less average) contributions, it comes much cheaper in terms of subsidies. Hence, the increment raised is substantially larger. Nevertheless, the above conclusion about θ extends to t : it is also a poor instrument and this is without accounting for the crowding out effect a la Frey (see, e.g., the empirical application to an environmental issue in Frey and Oberholzer-Gee (1997)) or the depreciation of the signal of a truly voluntary contribution as argued in Benabou and Tirole (2006).

Why is there this difference, given that: (i) both formulations yield the same contributions absent incentives, (ii) the enumeration of types is identical, $\theta = t$, and (iii) moreover, normalized, $E\theta = Et = 1$? The left hand sides of the relaxed program conditions are declining (due to concave objectives) and characterize the first best if equated to zero. Since $W(x, \theta) = tW(x, t)$, the left hand side of equation (25) exceeds its counterpart in equation (32) for $t = \theta < 1$ and the converse holds for $t = \theta > 1$. The right hand sides of the equations for the relaxed program, equations (32) and (25), determine the distortions due to the agency costs. These agency costs are declining with respect to types and they are lower for the t -types (from equation (32)),

$$\bar{t} - t < \frac{\beta(\bar{\theta} - \theta)}{\theta^2(1 - x)} \iff v > \delta.$$

The equivalence follows after using the identity $t = \theta$ and substituting the relaxed program $x^r(\theta)$. Therefore, combining a larger first best with lower agency costs implies:

Proposition 4. *More contributions (if positive) result for the θ -types for $\theta = t \leq 1$ and beyond. Since the first best relation is reversed for $t > 1$ and the agency costs decline, high t -types contribute more due to no distortion at the top, $x^1(t) > x^1(\theta)$ for $t = \theta \rightarrow \bar{t} = 1 + \delta$.*

3.2. Protecting a rainforest

3.2.1. Cost - t , that is, the evaluation of the rainforest - is private information

Subtracting from the objective V from equation (11) the subsidies using the definition in equation (15), and W from equation (12), the principal has to solve the following optimal

control problem:

$$\max_{\{x_0(t) \geq x(t) \geq 0\}} \int_t^{\bar{t}} [v(R - x(t)) + px(t) + t \ln(R - x(t)) - U(t)] dF, \tag{37}$$

$$\dot{U}(t) = W_t = \ln(R - x(t)), \tag{38}$$

$$U(t) \geq U_0(t) := pR + (\ln \frac{t}{p} - 1) \forall t \in [\underline{t}, \bar{t}]. \tag{39}$$

The control constraints stipulate that incentives are only offered for a reduction of logging ($x(t) \leq x_0(t)$) and that logging must be non-negative (no reforestation). The other constraints account, as in the above cases, for incentive compatibility equation (38) and individual rationality equation (39). The reservation price U_0 is again type dependent and characterizes the agent’s outside option from logging as he likes.

In this example, equation (22) implies the relaxed program,

$$x^r(t) = R + \frac{(\bar{t} - 2t)}{p - v} \geq x^l(t) = R - \frac{t}{p - v}, \tag{40}$$

which now exceeds the first best, because x is a negative action and the principal finds it too costly to reduce logging to the first best. Proper monotonicity, here $\dot{x} \leq 0$, requires $p > v$, and the participation constraint equation (39) binds again in many if not most cases at least for a subset of types.

Proposition 5. *Assuming that: (i) $p > v \iff x^r < 0$, i.e., the relaxed program (x^r) satisfies monotonicity, (ii) $R > (\bar{t}/p - v) = (1 + \delta/p - v) \iff x^r > 0 \forall t$ ensures positive logging, and (iii) the appreciation of the rainforest is sufficiently high,*

$$t > t^{\min} := \frac{p(1 + \delta)}{p + v}, \tag{41}$$

then the relaxed program is optimal, $x(t) = x^r(t)$. The relaxed program is continuously joined with the boundary solution x_0 from equation (14) at $t = t^{\min} \geq t$ and incentives are restricted to the types $t \geq t^{\min}$. They receive the compensation,

$$s(t) = \left\{ \begin{array}{l} 0 \\ \frac{2((p + v)t - p\bar{t}) + \bar{t}(p - v) \ln(\bar{t}/2t - \bar{t})(p - v/p + v)}{2(p - v)} \text{ for } t \begin{array}{l} < \\ \geq \end{array} \end{array} \right\} t^{\min}, \tag{42}$$

as a function of the type or conditional on logging (denoted by the function σ),

$$\sigma(x) = \frac{1}{2}((p + v)(R - x) - \bar{t}(\ln \frac{\bar{t}}{(p + v)(R - x)} - 1)) \text{ if } x < x_0, \tag{43}$$

in order to reduce their logging from x_0 to x^r equation (40).

The proof is straightforward. All properties along the relaxed program follow from the explicit solution. A necessary condition for the provision of incentives is that the principal values the protection less than the agent values logging, $v < p$, because

otherwise the value of logging is negative even leaving the agent’s own appreciation of the rainforest aside. The subsidies follow from the procedure outlined in section 3.1.

The restriction of incentives to a subset of types, $t > t^{\min}$, is a consequence of the participation constraint equation (39) and of the type dependency of the reservation price. More precisely and economically intuitive, handing out subsidies must be conditional on reduced logging, i.e., $x < x_0 = R - t/p$, arithmetically: starting at a type \tilde{t} with $x^r(\tilde{t}) > x_0(\tilde{t})$ violates the participation constraint because then $\dot{U}(\tilde{t}) = \ln(R - x^r(\tilde{t})) < \dot{U}_0 = \ln(R - x_0(\tilde{t}))$. The second constraint, $t > \bar{t}/2$, is necessary to ensure that some of the rainforest remains after logging, $x^r < R$, and this follows from the construction of $t^{\min} > \bar{t}/2 \Leftrightarrow 2p > v + p$ and the assumption $p > v$. A continuous joining of the boundary policy, $x = x_0$, with the relaxed program, x^r , at the point of intersection t^{\min} in equation (41) is optimal (i.e., it satisfies all necessary and sufficient optimality conditions).

However, the relaxed program can turn negative at high valuations or small forest areas (R). Therefore, the inequality (ii) assumed in proposition 5 ensures (positive) logging.

Corollary 2. *The share of agents receiving incentives,*

$$\frac{\bar{t} - t^{\min}}{\bar{t} - \underline{t}} = \frac{(1 + \delta)v}{2\delta},$$

increases with the principal’s appreciation of the rainforest (v) and is reduced by a larger variation of the agent types (δ).

The optimal incentive scheme is shown in figure 3, which compares the two different incentives resulting from the two observation equivalent characterizations of the agent. Logging is declining (linearly) in the agent’s type and this decline is compensated by increasing and convex subsidies to the agent with respect to types and logging and, of course, less logging earns higher subsidies. However, only a fraction ($t^{\min} = 8/9$ and thus only $2/3$) of the types receive incentives and they reduce logging only modestly although the principal’s valuation is significant (at one-half of the profit that the agent gets).

3.2.2. Benefit (i.e., importance of earnings) is private information

The alternative but observation equivalent assumption – the benefit from logging is the agent’s private information – leads again to a different incentive contract. The principal has to solve the following control problem (listing all arguments in order to highlight the difference from equations (37)–(39)),

$$\max_{\{x_0 \geq x \geq 0\}} \int_{\underline{\theta}}^{\bar{\theta}} \left[v(R - x(\theta)) + \frac{p}{\theta}x + \ln(R - x(\theta)) - U(\theta) \right] dF(\theta), \tag{44}$$

$$\dot{U}(\theta) = W_\theta = -\frac{px(\theta)}{\theta^2}, \tag{45}$$

$$U(\theta) \geq U_0(\theta) := \frac{pR}{\theta} + \ln \frac{\theta}{p} - 1 \quad \forall \theta \in [\underline{\theta}, \bar{\theta}], \tag{46}$$

and the constraints reflect incentive compatibility equation (45) and individual rationality equation (46) as in the above case.

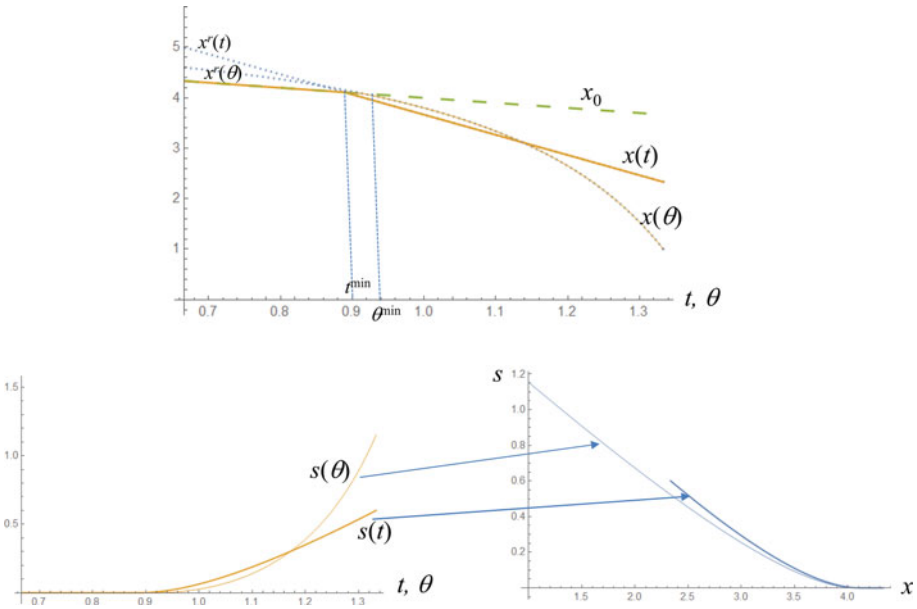


Figure 3. Logging the rain forest with either the conservation (t and bold) or revenue (θ) as the agent’s private information parameter, $p = 1$, $a = 1/2$, $R = 5$.

Proposition 6. Assuming $v < v^{\max} := \frac{p}{\theta} - \frac{1}{R}$ and $\theta \geq \theta^{\min} := \frac{\sqrt{p(p+4\bar{\theta}v)} - p}{2v} > \theta$, the relaxed program,

$$x^r(\theta) = R - \frac{\theta^2}{p\bar{\theta} - \theta^2v}, \tag{47}$$

determines the amount of logging. This program differs from its counterpart (proposition 5 and the agent’s valuation of the resource is private information) in terms of logging and participation: a contract in terms of θ is more inclusive, i.e., $t^{\min} \geq \theta^{\min}$, iff

$$(\bar{t} - 1) \geq \frac{v}{p}. \tag{48}$$

In terms of volume, $x^r(\bar{t}) > x^r(\bar{\theta}) \iff p > \bar{\theta}v$ and by continuity for types close to the top. Hence, a contract in terms of θ is (locally at least) more conservationist if the agent’s revenues exceed the donor’s valuation times the agent’s highest valuation of the rainforest (= least valuation of the revenues).

The proof is along the lines in 3.1.1 and thus brief. Starting now directly with equation (22),

$$-vx + \frac{p}{\theta} - \frac{1}{R - x} = \frac{p(\theta - \bar{\theta})}{\theta^2},$$

the solution $x^r(\theta)$ in equation (47) is verified. Comparing the marginal types and applying a few elementary but tedious operations yields equation (48).

However, condition equation (48) is not met in the example in figure 3 since $\delta = 1/3 < v/p = 1/2$. The panel at the top shows logging within, $x(t)$ and $x(\theta)$, outside of the contracts (x_0), and the relaxed programs outside of their range of applicability (x^r and dotted lines). Hence, the contract in terms of t includes slightly more types while the contract in terms of θ is significantly more conservationist for efficient types (as claimed) but not for intermediate types due to its concave shape. The bottom panels compare the subsidies depending on the type (left hand side) and on the amount of logging (right hand side). The subsidies cover a much larger range in terms of types and logging if the agent's private information is about his need for revenues (i.e., the benefit parameter θ).

3.3. Other examples

Although it does not fit our setup due to the additional complexity of multilateral externalities, the approaches to climate mitigation chosen in Martimort and Sand-Zantman (2013) and in Helm and Wirl (2014) highlight how different contracts can be. Helm and Wirl (2016) show that the two setups, private information attached to either benefits or costs, are observational equivalent but differ even in which of the types (low or high) is the efficient one.

Many other environmental examples abound. For example, the example of warm glow captures with few modifications aspects of energy conservation: b denoting the financial gain from energy efficiency investments at cost c and $t = \theta$ denoting the individually demanded payback time depending on the agents' subjective discount rate (see Wirl, 2019).

4. Concluding remarks

This paper addresses two different objectives. First, it considers the potential of environmental incentives in two different areas, protection of rainforests and contributions to public goods. Accounting for private information, the conclusion is that only a fraction of the usually reported potential will be realized if the recipients of incentives hold private information. The reason is that it makes no sense to incentivize inefficient types although they are the ones with the highest potential. Rather the opposite, subsidies should go to the already relatively efficient types who yield much less conservation (or contribution) given their own efforts in the absence of subsidies. In fact, given the information rent that they can earn, they are free riders at least to some extent. The reason why the inefficient types are pooled and excluded from subsidies is that handing out subsidies to inefficient types is costly and will achieve little for the money invested. And this dominates the theoretically high conservation potential.

This negative finding applies to many real world incentives. Actually the situation is much worse as current policy instruments and their implicit incentives do not even try to deter cheating, while the contracts derived in this paper are based on truth telling (the revelation principle). This neglect of cheating applies to Joint Implementation, the Clean Development Mechanism, to energy conservation across the board, in the past and currently under the label white certificates in the European Union (even the rebound is ignored (Bertoldi, 2011)). Therefore, most of the conservations that are attributed to these real world incentive schemes exist only on paper (and this under the idealistic assumption that transfers cause no deadweight loss).

A crucial feature of environmental problems is that the agents face a trade-off between benefit and cost. Adding private information, the question is whether to assign the

agent's private information to the benefit or the cost. This assignment is irrelevant in the absence of incentives, because the agent takes the same decision. Surprisingly, this seemingly arbitrary choice can matter for the design of environmental contracts. Therefore, one cannot attach arbitrarily (or by convenience) the private information within a trade-off but must argue where to attach it. This seems easy in some cases (e.g., in the case of rainforest protection) but not in all cases (e.g., contributions to a public good). Therefore, it should be worthwhile to revisit many of the proposed incentive mechanisms, whether the chosen private information parameter is well justified and, if not, how sensitive the contract is to this choice.

Acknowledgments. The author acknowledges helpful discussions with Carsten Helm and valuable comments from three anonymous referees.

References

- Andreoni J** (1990) Impure altruism and donations to public goods: a theory of warm-glow giving. *The Economic Journal* **100**, 464–477.
- Benabou R and Tirole J** (2006) Incentives and prosocial behavior. *American Economic Review* **96**, 1652–1678.
- Bertoldi P** (2011) Assessment and experience of white certificate schemes in the European Union. Available at <http://www.iea.org/media/workshops/2011/aupedee/PaoloBertoldi.pdf>.
- Frey BS and Oberholzer-Gee F** (1997) The cost of price incentives: an empirical analysis of motivation crowding-out. *The American Economic Review* **87**, 746–755.
- Fudenberg D and Tirole J** (1992) *Game Theory*, 2nd Printing. Cambridge, MA: MIT Press.
- Harstad B and Mideksa T** (2017) Conservation contracts and political regimes. *Review of Economic Studies* **84**, 1708–1734.
- Helm C and Wirl F** (2014) The principal-agent model with multilateral externalities: an application to climate agreements. *Journal of Environmental Economics and Management* **67**, 141–154.
- Helm C and Wirl F** (2016) Multilateral externalities: contracts with private information either about costs or benefits. *Economics Letters* **141**, 27–31.
- Kremen C, Niles JO, Dalton MG, Daily GC, Ehrlich PR, Fay JP, Grewal D and Guillery RP** (2000) Economic incentives for rain forest conservation across scales. *Science* **288**, 1828–1832.
- Ma C and Burton M** (2016) Warm glow from green power: evidence from Australian electricity consumers. *Journal of Environmental Economics and Management* **78**, 106–120.
- Martimort D and Sand-Zantman W** (2013) Solving the global warming problem: beyond markets, simple mechanisms may help!. *Canadian Journal of Economics* **46**, 361–378.
- Ruff LE** (1988) Least-cost-planning and demand-side-management: six common fallacies and one simple truth. *Public Utilities Fortnightly* **28**, 19–26.
- The Economist** (2013) Kenya's lions. Sad for Simba. The fight to save the big cat is on, *The Economist*, 26 January 2013.
- Wichman CJ** (2016) Incentives, green preferences, and private provision of impure public goods. *Journal of Environmental Economics and Management* **79**, 208–220.
- Wirl F** (2015) White certificates – energy efficiency programs under private information of consumers. *Energy Economics* **49**, 507–515.
- Wirl F** (2019) Incentivizing energy efficiency under private information: the social optimum. *The Energy Journal* **40**, 1–22.
- Wirl F, Huber C and Walker IO** (1998) Joint implementation: strategic reactions and possible remedies. *Environmental and Resource Economics* **12**, 203–224.

Cite this article: Wirl F (2022). Environmental incentives facing private information. *Environment and Development Economics* **27**, 167–183. <https://doi.org/10.1017/S1355770X21000048>