Erschienen in: European Economic Review ; 110 (2018). - S. 211-222
http://dx.doi.org/10.1016/j.euroecorev.2018.08.005

European Economic Review 110 (2018) 211–222

Contents lists available at ScienceDirect

European Economic Review

journal homepage: www.elsevier.com/locate/euroecorev

Pollution claim settlements reconsidered: Hidden information and bounded payments[☆]

Susanne Goldlücke[a,b], Patrick W. Schmitz[c,b,*]

[a] University of Konstanz, Germany
[b] CEPR, London, UK
[c] University of Cologne, Germany

ARTICLE INFO

Article history:
Received 19 March 2018
Accepted 16 August 2018
Available online 27 August 2018

JEL classification:
D86
D82
D23
D62
H23

Keywords:
Coasian contracting
Negative externalities
Hidden information
Limited liability
Overproduction

ABSTRACT

A principal's production decision imposes a negative externality on an agent. The principal may be a pollution-generating firm, the agent may be a nearby town. The principal offers a contract to the agent, who has the right to be free of pollution. Then the agent privately learns the disutility of pollution. Finally, a production level and a transfer payment are implemented. Suppose there is an upper bound (possibly zero) on payments that the agent can make to the principal. In the second-best solution, there is underproduction for low cost types, while there is overproduction for high cost types. In contrast to standard adverse selection models of pollution claim settlements, there may thus be *too much* pollution compared to the first-best solution.

1. Introduction

We revisit the classical problem of internalizing a negative externality through Coasian contracting. Consider a principal who can implement a verifiable production level and an agent who is negatively affected by the principal's production. In the prominent example discussed by Coase (1960), the principal is a cattle raiser and the agent is a farmer whose crops may be destroyed by straying cattle. Analogously, the principal may be a pollution-generating firm and the agent may be a nearby community. We suppose that the agent has the right to be free of pollution, so in the absence of an agreement between the two parties the production level has to be zero.

☆ Susanne Goldlücke: Department of Economics, University of Konstanz, Box 150, 78457 Konstanz, Germany. E-mail address: <susanne.goldluecke@uni-konstanz.de>. Patrick Schmitz: Department of Economics, University of Cologne, Albertus-Magnus-Platz, 50923 Köln, Germany. E-mail address: <patrick.schmitz@uni-koeln.de>. We would like to thank an associate editor and three referees for making valuable comments and suggestions. We also thank Oliver Gürtler, Eva Hoppe, Daniel Krähmer, and Daniel Müller for helpful discussions. Patrick Schmitz gratefully acknowledges financial support under the Institutional Strategy of the University of Cologne within the German Excellence Initiative (Hans-Kelsen-Prize 2015).

* Corresponding author at: University of Cologne, Department of Economics, Albertus-Magnus Platz, 50923 Cologne, Germany.
E-mail address: patrick.schmitz@uni-koeln.de (P.W. Schmitz).

https://doi.org/10.1016/j.euroecorev.2018.08.005

Following Rob's (1989) pioneering work on pollution claim settlements, we assume that the parties are risk-neutral and that the principal can make a take-it-or-leave-it offer to the agent. Suppose first that there are no relevant constraints on the transfer payments. Clearly, if there is complete information, according to the Coase Theorem the first-best solution that maximizes the parties' total surplus will be attained. Moreover, the first-best solution will also be attained if the agent becomes privately informed about his disutility of pollution after the contract is written (see Arrow, 1979, and d'Aspremont and Gérard-Varet, 1979). In both cases, the principal will extract the expected total surplus. Yet, as has been shown by Rob (1989), if the agent has private information already at the contracting stage, the principal faces a trade-off between rent extraction and achieving ex post efficiency, so there will typically be too little production and hence *too little pollution* compared to the first-best solution.

In the present paper, we analyze the case in which the agent becomes privately informed about his cost type only after the contract is written.[1] For example, in our framework the agent could be a municipality that contemplates granting a fracking permit. Fracking is a relatively new technology, so it is not yet fully known what the effects on the environment and human population will be. Similarly, carbon capture and storage is a promising new technology, which however exposes the host community to poorly understood leakage risks. The siting of such facilities may therefore require to negotiate a host community compensation agreement (Ter Mors et al., 2012). However, we assume that payments from the principal to the agent must be non-negative.[2] Such a constraint may be relevant if the agent has no resources or if it is politically infeasible to let a community that is harmed by pollution make positive payments to the polluter. Non-negativity constraints on payments are often imposed in hidden action models with limited liability (see e.g. Innes, 1990), but to the best of our knowledge they have not yet been studied in hidden information problems where the principal is in charge of a contractible action and there are more than two types.[3]

We show that while adding the non-negativity constraint to our problem has no effect in the two-type case, a novel kind of distortion away from the first-best solution can arise if there are more than two types. In particular, if the agent's cost type is continuously distributed, we find that for low levels of the agent's costs there is a downward distortion (except for the lowest cost type), while for high levels of the agent's costs there is an upward distortion. This finding contrasts with standard adverse selection models, where the solution usually involves a downward distortion only.

The intuition for the upward distortion in our hidden information setup is as follows. The fact that payments have to be non-negative means that the principal cannot extract the expected total surplus by letting the agent make a monetary transfer payment to the principal. Therefore, utility will instead be transferred from the agent to the principal through an inefficiently high production level. As a consequence, our model can provide a novel explanation for why in practice we may observe too much production and hence *too much pollution* compared to the first-best solution.[4]

Our result has interesting implications with regard to how the expected total surplus level depends on the sequence of events. Recall that in the absence of a constraint on feasible payments, the expected total surplus in standard adverse selection models with pre-contractual private information is smaller than the expected total surplus in corresponding hidden information models with post-contractual private information (where the first-best solution is achieved).[5] Yet, given that payments have to be non-negative, the expected total surplus can be *larger* in situations where the agent has learned his private information already at the contracting stage than in otherwise similar situations in which he will learn his private information after the contract has been written. Intuitively, when the agent learns his type after the contract has been signed, the principal will extract the expected total surplus by introducing ex post inefficient upward distortions, which can reduce the expected total surplus compared to a situation in which the principal must leave a rent to the agent since he has private information at the contracting stage already.

Related literature. Starting with Innes (1990), imposing a non-negativity constraint on payments from the principal to the agent has become a standard assumption in the moral hazard literature that studies optimal contracts when the agent is in charge of a hidden action.[6] In this literature, the principal must typically leave a limited liability rent to the agent in order

[1] According to the taxonomy proposed by Hart and Holmström (1987, p. 76), "adverse selection" models are characterized by pre-contractual private information, while "hidden information" models are characterized by post-contractual private information. In their wording, we thus study a hidden information variant of Rob's (1989) adverse selection problem.

[2] In the formal analysis, we will make the more general assumption that the agent cannot pay more than $\bar{t} \geq 0$ to the principal. For simplicity, in the introduction we focus our discussion on the case $\bar{t} = 0$.

[3] The introduction of a non-negativity constraint on payments would have no effect in standard adverse selection models with pre-contractual private information such as Rob (1989); cf. Section 4.1 below. See Laffont and Martimort (2002, section 3.5) for a discussion of two-type hidden information models with limited liability constraints.

[4] While Rob's (1989) classical model of pollution claim settlements can explain why pollution may be at inefficiently low levels from a social welfare perspective, many empirical studies find that in reality there may well be too much pollution, see e.g. OECD (2016).

[5] For a textbook exposition, see e.g. Laffont and Martimort (2002, sections 2.6 and 2.11).

[6] Some authors such as Tirole (1999, p. 745) and Laffont and Martimort (2002, p. 174) use the term "efficiency wage" model as a label for hidden action problems with resource-constrained agents. Note that limited liability also plays a central role in the sharecropping literature that studies the agrarian sector of less developed countries (see e.g. Shetty, 1988). As has been pointed out by Laffont and Martimort (2002, section 4.8.2), this literature typically assumes at the outset that sharing rules are linear, while the more recent contract-theoretic literature studies optimal contracts without imposing ad hoc constraints on the class of feasible contracts.

to induce him to choose high effort. To reduce this rent, the principal may prefer to implement a smaller effort level than she would do in a first-best world.[7]

Yet, in the literature on hidden information problems, the implications of bounded transfer payments have received less attention. In particular, following Sappington (1983), some authors have studied hidden information problems in which the agent chooses a verifiable action and is protected by limited liability in the sense of a lower bound on the agent's ex post utility.[8] Such constraints have effects similar to the individual rationality constraints in adverse selection models, which under standard assumptions lead to downward (but not upward) distortions.

Pesendorfer (1998) also imposes a non-negativity constraint on payments in a variant of Rob's (1989) model of pollution claim settlements. However, he studies an adverse selection problem with multiple agents whose cost types (which are privately known already at the contracting stage) are correlated across agents. Klibanoff and Morduch (1995) consider an adverse selection problem which shares with our model the feature that the production of one firm can have an external effect on another firm. In contrast to Rob (1989) and the present paper, they assume that firm 1 has the right to cause externalities by its production and it has private information about its profitability, while the impact of the externality on firm 2 is assumed to be public knowledge.

Alternative explanations of upward distortions of production levels can also be given in models where the agent has pre-contractual private information that differ from Rob's (1989) standard adverse selection setup. For example, an incentive to understate costs due to a reservation utility that is decreasing in marginal costs can lead to higher production levels for larger cost types. In particular, Lewis and Sappington (1989) study adverse selection with countervailing incentives due to type-dependence of the agent's reservation utility, Wirl and Huber (2005) analyze countervailing incentives created by the threat of a pollution tax, and Kessler et al. (2005) consider an adverse selection model where a signal that is correlated with the type of a wealth-constrained agent can be verified ex post. Of course, too much pollution can also be a consequence of insufficient regulation and impediments to Coasian bargaining like adverse selection on the firm's abatement costs or enforcement problems (see Baron, 1985, and Bontems and Bourgeon, 2005). In our setting, the agent has the right to be free of pollution, all pollution levels can be perfectly enforced, and inefficiently high pollution can arise because the principal has bargaining power but not enough instruments to efficiently extract rents.

Organization of the paper. The remainder of the paper is organized as follows. The basic model is presented in Section 2. We analyze the model in Section 3. In Section 4, we investigate implications of our results and we discuss some extensions of the model. Concluding remarks follow in Section 5. Some proofs have been relegated to the Appendix.

2. Model

There are two risk-neutral parties, a principal (a pollution-generating firm) and an agent (say, representing a nearby town), who enter a contractual relationship. The principal has a technology to produce any output $x \in [0, x^{\max}]$, yielding profit $V(x)$, for some differentiable and concave function V with $V(0) = 0$. Production has a negative externality on the agent. If the principal were not liable for the damage that production may cause, she would choose the quantity $x^{nl} = \arg\max_{x\in[0,x^{\max}]} V(x)$. We assume that x^{nl} is well-defined and take $x^{\max} = x^{nl}$, so that V is weakly increasing on $[0, x^{\max}]$.

An output level x leads to a (non-monetary) cost cx for the agent, where initially neither the principal nor the agent knows the realization of the cost parameter $c \in C$. Cost types are distributed according to a cumulative distribution function F, which is concave.[9] Unless stated otherwise, we assume that the support of F is an interval, $C = [c_L, c_H]$ with $c_H > c_L \geq 0$, and F is differentiable with density $f > 0$. We also consider the case that the support is finite, $C = \{c_1, \ldots, c_n\}$ with $c_L = c_1 < \ldots < c_n = c_H$. We study a Bayesian mechanism design problem; i.e., we impose no ad hoc restriction on the class of feasible contracts, the principal has full commitment power, and except for the realization of c all aspects of the model are common knowledge.[10]

The sequence of events is as follows. First, the principal proposes a contract, which consists of a menu of production levels and associated transfer payments. Then the agent accepts or rejects. We assume that when no agreement between the parties is reached, the agent has the right to be free of pollution. The agent's reservation utility is therefore either equal to zero or given by some alternative land use, and is denoted by $\bar{u} \geq 0$. The principal's reservation utility is normalized

[7] See Laffont and Martimort (2002, section 4.3) for a textbook exposition. For recent applications of hidden action models with bounded payments where limited liability constraints lead to deviations from first-best outcomes, see e.g. Ohlendorf and Schmitz (2012), Chen and Chiu (2013), Imhof and Kräkel (2014), Tamada and Tsai (2014), or Kräkel and Schöttner (2016). See also Lewis and Sappington (2000) for a model where a wealth-constrained agent exerts unobservable effort and is privately informed about his ability at the contracting stage.

[8] See e.g. Martimort (2006, p. 15) or the implementation stage in Khalil et al. (2006). Note that in standard hidden information problems without wealth constraints, it does not matter whether the verifiable action (say, the production level) is chosen by the agent or by the principal. We study the case in which the principal is in charge of implementing the action (so the contract can always be enforced, since the principal is solvent and thus not judgement-proof). If instead a wealth-constrained agent is in charge of the action, it may be impossible to enforce the contractually specified action. In this case, the wealth constraint may imply a lower bound on the agent's ex post utility.

[9] Concavity of F means that for all c, c', $c'' \in [c_L, c_H]$ with c in the support of F and $c = \lambda c' + (1-\lambda)c''$ for some $\lambda \in [0, 1]$, it holds that $F(c) \geq \lambda F(c') + (1-\lambda)F(c'')$. The support of F is the support of the corresponding probability distribution, i.e., the smallest closed set having probability 1.

[10] See Fudenberg and Tirole (1991, chapter 7) for an excellent introduction into the theory of Bayesian mechanism design.

to 0. After the agent has accepted the contract, he privately learns the realized cost parameter c.[11] Finally, the contract is executed: The agent chooses a production level and associated payment from the contractually agreed-upon menu. The principal then implements the production level x and makes a payment t, leading to payoffs $t - cx$ for the agent and $V(x) - t$ for the principal.

The maximum expected surplus (which would always be attained in a first-best world) is denoted by

$$S^{FB} = \int_{c_L}^{c_H} \max_{x \in [0, x^{\max}]} (V(x) - cx) dF(c),$$

where we assume that $S^{FB} \geq \bar{u}$.

It is well-known that the principal's problem has a simple solution if there are no constraints on what the agent can pay. The optimal contract makes the agent residual claimant for the profit and lets the principal receive the entire expected surplus.[12]

Remark 1. If transfer payments are unbounded, the first-best solution can be attained by a contract that lets the agent decide on x and that specifies the payment $t(x) = \bar{u} + V(x) - S^{FB}$ from the principal to the agent.

Observe that $t(x)$ may be negative; i.e., the optimal contract may require a positive transfer payment from the agent to the principal. Limits on the agent's wealth or political constraints may render this contract infeasible.[13] In what follows, we thus assume that the agent's ability to make transfer payments to the principal is restricted. Specifically, we impose (adopting the wording of Pesendorfer, 1998) a "limited liability" constraint $t \geq -\bar{t}$, where $\bar{t} \geq 0$ represents the agent's wealth or some other bound on transfers. Note that this constraint would have no effect if the principal and the agent both knew the agent's cost type.[14] Moreover, note that under our assumptions limited liability does not affect the implementability of the first-best output schedule. Yet, given that the constraint $t \geq -\bar{t}$ has to be satisfied, the presence of post-contractual private information on the agent's side may constitute a transaction cost, which may hinder the principal from appropriating the maximum expected surplus despite having all the bargaining power. As a consequence, an ex post inefficient outcome may result.

3. Analysis

According to the revelation principle (cf. Myerson, 1982), a general contract is of the form $(x(c), t(c))_{c \in C}$, specifying permitted output and associated payment for each reported cost type c. The principal proposes the contract that maximizes her expected profit $E[V(x) - t]$ among all contracts that the agent will accept, that require no payments from agent to principal above $\bar{t}$, and that induce truth-telling by the agent. Hence, the contract has to satisfy the participation constraint

$$E[t(c) - cx(c)] \geq \bar{u}, \tag{PC}$$

the limited liability constraints

$$t(c) \geq -\bar{t} \text{ for all } c, \tag{LL}$$

and the incentive compatibility constraints

$$t(c) - x(c)c \geq t(\hat{c}) - x(\hat{c})c \text{ for all } c, \hat{c}. \tag{IC}$$

We will look at each constraint in turn. First, we rewrite the incentive compatibility constraints in the standard way.

Lemma 1. *(IC) is equivalent to the following two conditions:*

$$t(c) = t(c_H) - x(c_H)c_H + x(c)c + \int_c^{c_H} x(\gamma) d\gamma, \tag{IC1}$$

and

$$x \text{ is weakly decreasing.} \tag{IC2}$$

For the case of a finite support, (IC1) defines the lowest possible payments to satisfy (IC) for given $t(c_H)$ if we extend functions $x : \{c_1, \ldots, c_n\} \to [0, x^{\max}]$ on the interval $[c_L, c_H]$ by defining $x(c) = x(c_{i+1})$ for $c \in (c_i, c_{i+1}]$.

Proof. See the Appendix. □

[11] Note that subsequent to signing the contract, the parties might have to make time-consuming relationship-specific investments before production can actually take place. For simplicity, in order to focus on ex post inefficiencies we do not model such investments explicitly (see e.g. Hart and Moore, 2008, for a similar approach).

[12] See e.g. Laffont and Martimort (2002, p. 57).

[13] The fact that local authorities suffer from binding budget constraints has often been emphasized by practitioners (see e.g. Committee on Climate Change, 2012).

[14] In this case, the principal would simply compensate the agent for his costs by offering the contract $t = \bar{u} + cx \geq 0$, so the first-best solution would be achieved.

Hence, the incentive compatibility constraints determine the transfer payments up to a constant. Next, we show how this constant is pinned down by the binding participation constraint.

Lemma 2. *In the optimum, the participation constraint is binding. A contract (x, t) with payment function t given by (IC1) satisfies (PC) with equality if*

$$t(c_H) = \bar{u} + x(c_H)c_H - \int_{c_L}^{c_H} x(c)F(c)dc. \tag{1}$$

As in Lemma 1, this result also holds for the finite case.

Proof. See the Appendix. □

Note that the binding participation constraint distinguishes our hidden information model from hidden action models with limited liability constraints, in which the agent typically receives a rent.[15]

Since transfers are weakly decreasing, they satisfy the limited liability constraints if and only if $t(c_H) \geq -\bar{t}$.

Lemma 3. *The principal's optimization problem is equivalent to maximizing*

$E[V(x) - cx]$ subject to the constraints (IC2) and

$$\int_{c_L}^{c_H} x(c)F(c)dc \leq x(c_H)c_H + \bar{t} + \bar{u}. \tag{LL-ICPC}$$

In particular, if any function x^{FB} with

$$x^{FB}(c) \in \arg\max_{x\in[0,x^{\max}]} (V(x) - cx)$$

satisfies (LL-ICPC), then the principal obtains the first-best payoff $S^{FB} - \bar{u}$. Otherwise, the first-best surplus is not obtainable and the new limited liability constraint (LL-ICPC) is binding. As in Lemma 1, this result also holds for the finite case.

Proof. See the Appendix. □

To analyze the effect of the limited liability constraint, it is natural to first consider the case in which the cost parameter can only take one of two values, $C = \{c_L, c_H\}$. In many settings, the two-type case provides much of the intuition of the general case. Yet, this is not true in our setup. Specifically, in the two-type case the principal can still achieve the first-best solution even when payments must be non-negative. This can be seen from Lemma 3, since condition (LL-ICPC) for the two-type case is $-x^{FB}(c_H)E[c] \leq \bar{t} + \bar{u}$.

Remark 2. The principal obtains the first-best surplus in the two-type case by proposing the contract $x(c) = x^{FB}(c)$, $t(c_H) = \bar{u} + x^{FB}(c_H)E[c]$ and $t(c_L) = t(c_H) + c_L(x^{FB}(c_L) - x^{FB}(c_H))$.

However, the fact that the limited liability constraint has no effect is an artefact of the binary case. Next, we provide an example with three types which illustrates why the two-type case is misleading in our setting.

Let $C = \{c_L, c_M, c_H\}$, and let f_L, f_M, f_H denote the associated probabilities. We assume $x^{\max} = 1$, $\bar{t} = 0$, $\bar{u} = 0$, and a linear profit function $V(x) = vx$ with $c_L < c_M < v < c_H$, such that $x^{FB}(c_L) = x^{FB}(c_M) = 1$ and $x^{FB}(c_H) = 0$.[16] Lemma 3 shows that the first-best solution cannot be achieved here (since $f_L(c_M - c_L) > 0$) and the principal solves

$$\max_{x_L, x_M, x_H \in [0,1]} f_L(v - c_L)x_L + f_M(v - c_M)x_M + f_H(v - c_H)x_H$$

subject to the binding (LL-ICPC) constraint

$$f_L x_M(c_M - c_L) + (f_L + f_M)x_H(c_H - c_M) = x_H c_H.$$

We see that there is no distortion at the top, $x_L = 1$, and the constraint implies

$$x_H = \frac{x_M f_L(c_M - c_L)}{(1 - f_H)c_M + f_H c_H}. \tag{2}$$

Plugging this into the objective function, we can conclude that if

$$f_M(v - c_M) + \frac{f_L(c_M - c_L)(v - c_H)f_H}{(1 - f_H)c_M + f_H c_H} \geq 0, \tag{3}$$

the solution has $x_M = 1$, and otherwise it has $x_M = 0$.

This simple example shows that with more than two types, the first-best solution is not necessarily attained anymore. Moreover, if f_M is large enough, the production level is *larger* than in the first-best solution.[17] The intuition behind this result

[15] In these models, the hidden action is typically an effort level chosen by the agent. The principal can reduce the agent's limited liability rent by implementing an effort level that is smaller than the first-best benchmark. See e.g. Laffont and Martimort (2002, Section 4.3).

[16] The other conceivable cases would result in the first-best allocation.

[17] For example, if $f_L = f_M = f_H = 1/3$, $c_L = 1$, $c_M = 2$, $c_H = 4$, and $v = 3$, we have $x_M = 1$ and $x_H = 1/8$.

is that due to the limited liability constraint, the principal cannot extract the agent's rent with a subsidy for low production levels, so that instead the agent makes a non-monetary "transfer" by allowing an inefficiently high level of pollution. Note that $x_M = x_H = 0$ is also possible; i.e., there can also be a downward distortion of the production level.[18]

In terms of comparative statics, it is straightforward to see that as v increases, expected output weakly increases: While x_H does not change, condition (3) is relaxed. Similarly, as c_H increases, expected output goes down. More surprisingly, x_H increases in c_M for $x_M = 1$, so as long as condition (3) is satisfied, expected output *increases* in c_M. The reason is that a larger medium cost type finds it more attractive to overstate his cost, so that the principal has to make this option less attractive. The opposite holds true for an increase in c_L, which decreases x_H, but can *increase* expected output if condition (3) becomes true.

Next we consider the case that the cost parameter can take any value in the interval $[c_L, c_H]$.

Proposition 1. *The optimal output schedule x^{LL} satisfies:*

(i) *If $\bar{t} + \bar{u}$ is sufficiently large such that (LL-ICPC) in Lemma 3 is satisfied for $x = x^{FB}$, then $x^{LL} = x^{FB}$ and the principal obtains the first-best payoff $S^{FB} - \bar{u}$.*

(ii) *Else there exists a weakly decreasing function $\hat{x}$, a cut-off $\bar{c} \in [c_L, c_H]$, and a value $\bar{x}$ such that*

$$x^{LL}(c) = \begin{cases} \hat{x}(c) & \text{for all } c \leq \bar{c}, \\ \bar{x} & \text{for all } c > \bar{c}. \end{cases}$$

The function $\hat{x}$ satisfies $\hat{x}(c) \leq x^{FB}(c)$, with strict inequality if $c \neq c_L$ and $0 < x^{FB}(c) < x^{\max}$. The threshold value $\bar{x}$ satisfies $x^{FB}(\bar{c}) \geq \bar{x} \geq x^{FB}(c_H)$, with strict inequality if $x^{FB}(c_H) > 0$.

Proof. Since *(i)* is already shown in Lemma 3, we have to solve the optimization problem given there for the case that no x^{FB} satisfies (LL-ICPC), which must therefore be binding. In order to do this, we write $\bar{x}$ instead of $x(c_H)$ in (LL-ICPC) and replace (IC2) by the weaker constraint $x(c) \geq \bar{x}$. We consider the following maximization problem over both x and $\bar{x}$:

$$\begin{aligned} &\max_{x(.),\bar{x}} \int_{c_L}^{c_H} (V(x(c)) - cx(c)) f(c) dc \\ &s.t. \quad \int_{c_L}^{c_H} x(c) F(c) dc \leq \bar{x} c_H + \bar{t} + \bar{u} \\ &\qquad x(c) \geq \bar{x} \end{aligned} \tag{4}$$

This is a relatively simple problem with a concave objective function and linear constraints. Let $x^*, \bar{x}^*$ denote the solution of this problem. If x^* turns out to be a weakly decreasing function with $x^*(c_H) = \bar{x}^*$, then x^* must also be the solution of the original maximization problem.

There must exist a Lagrange multiplier $\lambda \geq 0$ and a function $\mu \geq 0$ with $\mu(c)(x^*(c) - \bar{x}^*) = 0$ such that $x^*, \bar{x}^*$ solve

$$\max_{x(.),\bar{x}} \int_{c_L}^{c_H} (V(x(c)) - \left(c + \lambda \frac{F(c)}{f(c)}\right) x(c) + \lambda(\bar{x} c_H + \bar{t} + \bar{u}) + \mu(c)(x(c) - \bar{x})) f(c) dc. \tag{5}$$

If $\mu(c) > 0$ for some c, then $x^*(c) = \bar{x}^*$, and if $\mu(c) = 0$, then $x^*(c) \geq \bar{x}^*$ and $x^*(c)$ is a maximizer of $V(x) - (c + \lambda \frac{F(c)}{f(c)})x$. If this maximizer is unique, we define

$$\hat{x}(c) = \arg \max_{x \in [0, x^{\max}]} \left(V(x) - \left(c + \lambda \frac{F(c)}{f(c)}\right) x \right),$$

and if the set of maximizers is an interval, then $\hat{x}(c)$ is taken to be the upper interval limit. Our assumption that F is concave implies that as a function of (x, c), $V(x) - (c + \lambda \frac{F(c)}{f(c)})x$ has strictly decreasing differences, which implies that $\hat{x}(c)$ is weakly decreasing. Hence, there must be some cut-off $\bar{c} \in [c_L, c_H]$ such that $x^*(c) = \hat{x}(c)$ for $c \leq \bar{c}$ and $x^*(c) = \bar{x}^*$ for larger levels of the agent's costs.

Similarly, as a function of (x, λ), $V(x) - (c + \lambda \frac{F(c)}{f(c)})x$ has strictly decreasing differences, which implies that $\hat{x}(c) \leq x^{FB}(c)$. Considering the derivative $V'(x) - (c + \lambda \frac{F(c)}{f(c)})$, we see that the stronger condition $\hat{x}(c) < x^{FB}(c)$ holds unless $c = c_L$ or $\hat{x}(c)$ is a corner solution with either $\hat{x}(c) = 0 = x^{FB}(c)$ or $\hat{x}(c) = x^{\max} = x^{FB}(c)$.

Since x^* can nowhere be smaller than $x^{FB}(c_H)$ nor everywhere equal to $x^{\max}$, it has to hold that $x^{\max} > \bar{x}^* \geq x^{FB}(c_H)$. This can either mean that $\bar{x}^* = 0 = x^{FB}(c_H)$ or an interior solution for $\bar{x}^*$. In the first case, $x^*(c) = \hat{x}(c)$, and the conditions for the threshold $\bar{x}^*$ in the proposition hold with equality. In the latter case, it must hold that $\lambda c_H = E[\mu] > 0$. The function μ then has to be $\mu(c) = 0$ for $c \leq \bar{c}$ and $\mu(c) = -V'(\bar{x}^*) + (c + \lambda \frac{F(c)}{f(c)})$ else, and from $\lambda c_H = \int_{\bar{c}}^{c_H} (c - V'(\bar{x}^*) + \lambda \frac{F(c)}{f(c)}) f(c) dc$ it follows that

$$\lambda = \frac{\int_{\bar{c}}^{c_H} (c - V'(\bar{x}^*)) f(c) dc}{\bar{c} F(\bar{c}) + \int_{\bar{c}}^{c_H} c f(c) dc} < 1. \tag{6}$$

[18] For example, if $f_L = f_M = f_H = 1/3$, $c_L = 1$, $c_M = 2$, $c_H = 4$, and $v = 2.1$, we have $x_M = x_H = 0$. Note that although the output schedule is very different from the first-best solution, the expected surplus is very close to the first-best surplus.

Since $\lambda \geq 0$, this expression for λ implies that $c_H > V'(\bar{x}^*)$, which means that $\bar{x}^* > x^{FB}(c_H)$. Moreover, either $\bar{x}^* = \hat{x}(\bar{c}) < x^{FB}(\bar{c})$ or $\bar{x}^* < x^{FB}(\bar{c}) = x^{\max}$. □

Consider part (ii) of Proposition 1. The optimal contract provides the agent with a menu of output levels and corresponding payments that the principal must make to the agent. The induced output schedule has no distortion at the lowest cost type, then it decreases until it becomes flat, intersecting the first-best output schedule from below.[19] In particular, the smallest level of output in the menu is chosen by all agents with costs larger than a threshold $\bar{c}$. For very high cost types, production is thus larger than in the first-best solution.

Recall that in our hidden information setup, the agent learns his cost type only after the contract has been signed; i.e., in contrast to adverse selection models with pre-contractual private information no rent will be left to the agent.[20] In the absence of a limited liability constraint, the principal would push the agent's expected utility down to his reservation utility with a suitable payment schedule, which would be an efficient way of transferring utility between the parties. Yet, given limited liability, the principal pushes the agent down to his reservation utility by transferring utility in an inefficient way from the agent to the principal. Instead of using a monetary payment, the agent sometimes has to suffer from an inefficiently high level of production, which benefits the principal. The optimal contract thus no longer maximizes the expected total surplus, instead there is a trade-off between maximizing the total surplus and minimizing the loss caused by the inefficient way of transferring utility between the parties, which yields the novel pattern of distortions derived in Proposition 1.

The optimal contract allows the following interpretation. The principal and the agent agree on a compensation scheme that is tied to the principal's production level, $t(x)$, where $x \in [\bar{x}, x^{FB}(c_L)]$ and $t(x) = t^{LL}(c)$ with $x^{LL}(c) = x$. The agent retains the right to set a limit on production, which however cannot be lower than the threshold $\bar{x}$. For a given limit $\hat{x}$ that the agent chooses, the principal can choose any $x \leq \hat{x}$ and has to pay $t(x)$. The principal will then choose the maximally allowed level $\hat{x}$, and the agent will, after learning his cost type c, set the limit equal to $x^{LL}(c)$.

It has become commonplace that local governments and developers negotiate formal contracts regarding projects with potential negative externalities, in which they specify monetary or in-kind compensation for the host community (Selmi, 2010). From a legal perspective, such contracts can be problematic because the local government is required to keep the power to react to bad news and always be able to safeguard the well-being of the community. As Selmi (2010, p. 619) argues, "local governments in the twenty-first century have very limited financial resources" and will therefore not be able to simply breach a contract for land use if circumstances change. In the context of our model, this corresponds to the fact that limited liability renders a contract as in Remark 1 infeasible. The optimal contract, which is a combination of a guaranteed production level, a pre-determined compensation schedule for production above this level, and a flexible cap on production, balances the need for certainty for the developer (e.g. to protect its relationship-specific investment) and the need for the community to be able to react to changed circumstances in the future.[21]

There is a concern and some empirical evidence in the literature that poorer communities and those with less political participation have to bear more pollution (see e.g. Hamilton, 1993, and Pargal and Wheeler, 1996). Richer communities in our model correspond to an agent with a larger $\bar{t}$, and more activism against potential pollution would correspond to a larger $\bar{u}$, as the local government gains political support if it does not allow the land use by the polluting facility. In our model, a decrease in $\bar{t}$ and $\bar{u}$ have the same effect on the expected level of pollution, which can be positive or negative. While a decrease in $\bar{u}$ would make the community worse off, a decrease in $\bar{t}$ would leave its expected utility unchanged, since the larger expected compensation from the community to the producer would offset any increase in pollution levels.

4. Discussion

4.1. Post-contractual vs. pre-contractual information

If the agent knows the realization of his cost type already at the contracting stage, our model corresponds to a standard adverse selection problem.[22] In this case, the participation constraint (PC) has to be replaced by an individual rationality constraint that ensures participation for every possible realization of the agent's cost,

$$t(c) - cx(c) \geq \bar{u} \text{ for all } c. \tag{IR}$$

[19] Technically, our optimization problem is related to Levin (2003). To find the optimal relational contract in a repeated principal-agent setting, he has to solve a static problem of hidden information in which transfers are limited by the continuation value of the relationship. In contrast to our model with only a one-sided bound on transfers, in Levin (2003) the second-best output schedule is flat for low cost types and then decreasing, and always strictly below the first-best output schedule.

[20] Observe that this feature distinguishes our model from countervailing incentives models such as Lewis and Sappington (1989) and Maggi and Rodriguez-Clare (1995), in which the agent has private information already at the contracting stage and the agent's reservation utility is type-dependent.

[21] See also Kenney et al. (2004), who document cases of community benefits agreements, some of which specify extensive monitoring of output and environmental damage, combined with linear fees and annual production caps. They also illustrate the need to react to later circumstances. For instance, an agreement between a local activist group and a coal mine specified an annual production cap of 5 million tons and a penalty for higher output levels. However, when it turned out later that longer trains could transport more without any additional safety hazards, the cap was amended to allow 5.25 million tons with no additional remuneration for the community.

[22] See e.g. Laffont and Martimort (2002, p. 134).

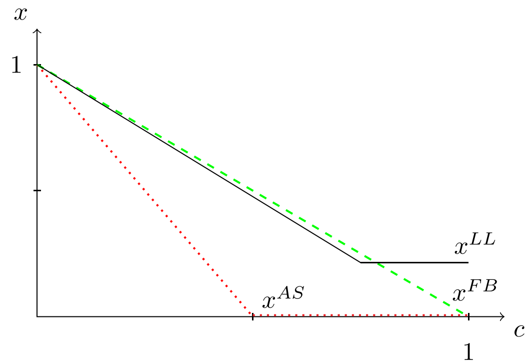

Fig. 1. This figure shows x^{LL} and how it compares to x^{AS} (dotted, red) and x^{FB} (dashed, green) for the example $V(x) = x(1 - \frac{x}{2})$, $x^{\max} = 1$, $\bar{t} = \bar{u} = 0$, and c uniformly distributed on [0,1].

Note that in the adverse selection model it is irrelevant whether or not we impose the limited liability constraint, because (IR) already implies (LL) since we assumed that $\bar{u} \geq 0 \geq -\bar{t}$.

Proposition 2. *Let $x^{AS}(c)$ denote the allocation from the optimal contract with pre-contractual private information. It holds that $x^{LL}(c) \geq x^{AS}(c)$ for all $c \in [c_L, c_H]$.*

Proof. In the case of pre-contractual private information, the optimal output plan is given by

$$x^{AS}(c) \in \arg\max_x V(x) - \left(c + \frac{F(c)}{f(c)}\right)x,$$

and therefore corresponds to $\lambda = 1$ in the function $\hat{x}$ in the proof of Proposition 1, which also shows that $\hat{x}$ is decreasing in λ. □

Fig. 1 shows how the functions x^{FB}, x^{LL}, and x^{AS} compare in a numerical example. Payments t^{LL} in this example are a convex and decreasing function, equal to zero where x^{LL} is flat. While in this example $t^{LL}(c) \geq t^{AS}(c)$ for all cost types, this does not hold in general.

Since the agent's participation constraint (PC) is binding with post-contractual private information, while the agent gets an information rent in the case of pre-contractual private information, the agent is better off in the latter case. The principal is worse off in the case of pre-contractual private information, because she could have chosen to implement x^{AS} in the case of post-contractual private information, but preferred not to do so. Recall that in the absence of a limited liability constraint, the first-best solution is attained in the case of post-contractual private information, so the expected total surplus in this case is at least as large as in the case of pre-contractual private information (S^{AS}). In contrast, when payments must be non-negative, then the surplus comparison becomes ambiguous. In particular, the expected total surplus in the case of post-contractual private information (S^{LL}) can now be strictly *smaller* than in the case of pre-contractual private information, where ex post inefficient upward distortions do not occur. This result is illustrated in Fig. 2.

In our hidden information model, we have assumed that the agent's costs only become known with experience, after the principal has built the hazardous facility and maybe even started production. Another possibility is that the agent can become informed about his true costs over time also in the absence of an agreement. We can then also relax the assumption of full commitment by the principal and assume that negotiations can take place both before (ex ante) and after (ex post) the agent has learned his cost type. Suppose the agent's reservation utility in the ex post negotiations is zero. In order to accept a contract at the ex ante stage already, the agent must now be offered at least $\bar{u} = E[u^{AS}(c)]$, where $u^{AS}(c) = \int_c^{c_H} x^{AS}(\gamma)d\gamma$ denotes the agent's information rent that he would get in the ex post negotiations. Since x^{LL} maximizes the expected surplus given the constraints in Lemma 3, and x^{AS} also satisfies these constraints for $\bar{u} = E[u^{AS}(c)]$, it holds in this case that $S^{LL} \geq S^{AS}$. Hence, the parties will indeed sign a contract before the agent learns his cost. The agent's option value of waiting until he learns his disutility may or may not be large enough to overcome the distortion in production levels, depending on whether $\int_{c_L}^{c_H}(x^{FB}(c) - x^{AS}(c))F(c)dc \leq x^{FB}(c_H)c_H + \bar{t}$ or not. For instance, in the numerical example illustrated in Fig. 1 (where $x^{FB}(c_H) = \bar{t} = 0$), a distortion due to hidden information arises even if the agent could get a new offer after he has learned his private information.

4.2. Bargaining power

Following Rob (1989) we have assumed that the pollution-generating firm can make a take-it-or-leave-it offer. This is the distribution of bargaining power for which the transaction costs that we study (i.e., hidden information when transfer payments are bounded) matter most. To capture the specifics of some real-world applications, it may be worthwhile to also study the case in which the agent has some bargaining power. We therefore now assume that the bargaining stage results in

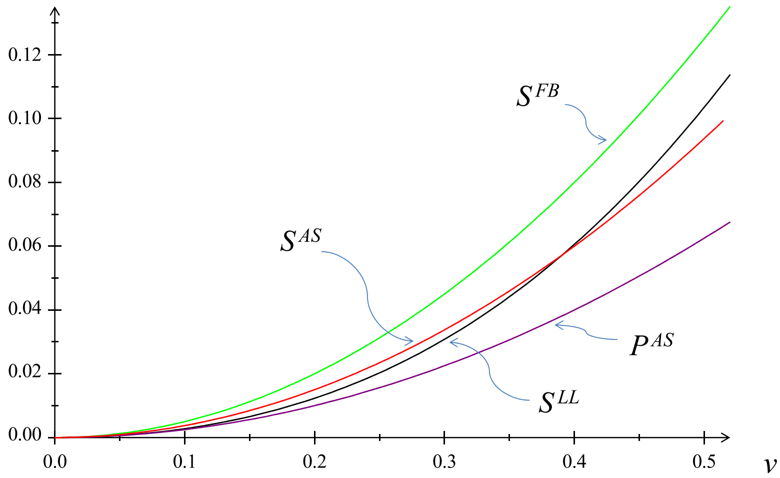

Fig. 2. In the figure, $V(x) = vx$, $x^{max} = 1$, $\bar{t} = \bar{u} = 0$, and c is uniformly distributed on [0,1]. The figure shows the expected total surplus in the case of post-contractual private information and limited liability (S^{LL}), compared to the expected total surplus (S^{AS}) and the principal's expected profit (P^{AS}) in the case of pre-contractual private information.

the generalized Nash bargaining solution, where $\alpha \in [0, 1]$ denotes the agent's bargaining power. The agent's disagreement payoff d_A is given by his reservation utility $\bar{u}$, while the principal's disagreement payoff d_P is given by her reservation utility zero. Agent and principal bargain over the set of payoffs that are attainable with a contract that satisfies (PC), (LL) and (IC), which has to be a convex set.

For a fixed agent's payoff $u_A \in [0, S^{FB}]$, the principal's maximum payoff is $S^{LL}(u_A) - u_A$, which means that the Pareto frontier of feasible expected payoffs consists of all payoff pairs (u_A, u_P) with $u_A \in [0, S^{FB}]$ and $u_P = S^{LL}(u_A) - u_A$. Note that $S^{LL}(u_A) - u_A$ is decreasing in u_A with derivative $\lambda(u_A) - 1$, where $0 \le \lambda(u_A) < 1$ is the Lagrange multiplier from the optimization problem in the proof of Proposition 1 for $\bar{u} = u_A$, and is equal to zero if $S^{LL}(u_A) = S^{FB}$. As u_A increases, the expected total surplus $S^{LL}(u_A)$ increases, and once u_A is large enough such that the condition in Lemma 3 holds, the Pareto frontier becomes linear and the first-best solution is attained. Moreover, the function $S^{LL}(u_A) - u_A$ is indeed concave, since for all contracts (x, t) and (x', t') that yield expected payoff pairs (u_A, u_P) and (u'_A, u'_P), and for all $\mu \in [0, 1]$, the contract $(\mu x + (1-\mu)x',\ \mu t + (1-\mu)t')$ also satisfies (LL) and (IC) and leaves the agent with a payoff equal to $\mu u_A + (1-\mu)u'_A$ and the principal with a payoff of at least $\mu u_P + (1-\mu)u'_P$.

The generalized Nash bargaining solution is the maximizer of $(u_A - d_A)^{\alpha}(u_P - d_P)^{1-\alpha}$ over all feasible expected payoff pairs (u_A, u_P). Since it will select a Pareto efficient point, the agent's bargaining payoff can be found by maximizing $(u_A - \bar{u})^{\alpha}(S^{LL}(u_A) - u_A)^{1-\alpha}$ over u_A. Hence, the Nash bargaining solution yields a negotiation payoff u_A for the agent that is implicitly given by

$$\bar{u} + \frac{\alpha}{1 - \lambda(u_A)(1-\alpha)}\left(S^{LL}(u_A) - \bar{u}\right) = u_A.$$

In particular, the agent's negotiation payoff is equal to $u_A = \bar{u} + \alpha(S^{FB} - \bar{u})$ if

$$\bar{u} + \alpha(S^{FB} - \bar{u}) \ge \int_{c_L}^{c_H} x^{FB}(c)F(c)dc - x^{FB}(c_H)c_H - \bar{t}. \tag{7}$$

Hence, if the agent's bargaining power or his disagreement payoff is sufficiently large, the first-best solution is attained.

4.3. Many agents

One may ask what happens if there are many agents, especially in light of the negative limit result in Rob (1989). In that paper, the principal bargains with n agents who have pre-contractual private information about their costs. Suppose that each agent's cost type c_i is independently drawn from the same distribution and the reservation utilities are zero. Moreover, let us hold per-person profits constant by setting $V(x) = nvx$, where $v < c_H$ and $x^{max} = 1$, so that x can be interpreted as the probability of production. As the number of agents becomes large, the obstacles to bargaining due to pre-contractual asymmetric information may become insurmountable. In particular, Rob (1989) shows that there are circumstances such that when n goes to infinity, then the ratio of realized to potential welfare (i.e., the expected total surplus given the optimal contract under adverse selection divided by the expected total surplus in the first-best solution) converges to zero.[23]

[23] See Mailath and Postlewaite (1990) for a closely related limit result in the context of public goods. In contrast, Rustichini et al. (1994) show that increasing the number of agents can mitigate the problems caused by pre-contractual private information in private-good settings. See also Pesendorfer (1998), who shows that Rob's (1989) negative result does not hold if the agents' types are correlated.

In contrast, a positive limit result can be established in our setting with post-contractual private information and limited liability. Specifically, when the number of agents goes to infinity, then the ratio of realized to potential welfare converges to one. To see this, suppose that $v > E[c_i]$.[24] Consider the following simple mechanism. The principal proposes to choose $x = 1$ and to pay $E[c_i]$ to each agent. Note that the payments are non-negative and each agent's participation constraint is binding. The expected total surplus attained by this simple mechanism, $n(v - E[c_i])$, is a lower bound on the expected total surplus that will be achieved by the principal's optimal mechanism. By the law of large numbers, for the ratio of realized to potential welfare it holds that

$$\lim_{n\to\infty} \frac{n(v - E[c_i])}{E[\max\{nv - \Sigma_i c_i, 0\}]} = \lim_{n\to\infty} \frac{v - E[c_i]}{E[\max\{v - \frac{1}{n}\Sigma_i c_i, 0\}]} = 1.$$

Intuitively, when the number of agents goes to infinity, then the ex post efficient decision is already known with probability one ex ante, such that the principal can simply compensate each agent for his expected costs.

5. Concluding remarks

We have provided a new perspective on the classical topic of Coasian contracting to internalize a negative externality. In our model, the principal can implement a verifiable production level, the agent learns the realization of the costs caused by pollution after the contract is signed, and payments to the agent are not allowed to become (too) negative. We have shown that for high cost types there may be an *upward* distortion of the production level. Moreover, the expected total surplus in our hidden information model can be *smaller* than the expected total surplus in an otherwise similar adverse selection problem in which the agent learns his cost type before the contract is written.

Situations in which a firm's or a government agency's decisions may cause negative externalities abound in practice. Related examples include the siting of waste dumps, power plants, electricity pylons, or wind turbines. Similarly, communities might suffer from having adult entertainment clubs, drug consumption rooms, homeless shelters, or refugee hostels in their backyards. Also in these applications it is well possible that the agent's disutility (e.g., due to nuisance caused by noise) is learned only after the facility is built, and the extent of the nuisance is increasing in the verifiable occupancy rate.

Yet, while we have assumed that the agent has the right to be unaffected by externalities, in some applications this might not be the case.[25] The principal would then choose $x^{\max}$, which in our model would lead to an expected total surplus smaller than in the solution we have characterized.[26] Moreover, in some applications the federal government may be the principal and a local authority may be the agent. In this case, one might also want to consider the possibility that the principal's objective function puts some weight $\beta \in (0, 1)$ on the agent's utility (cf. Baron and Myerson, 1982).

Recall that our model is tailored to the case in which the principal implements the production level, while the agent is negatively affected by production, which is a standard setup in environmental problems. In other contexts such as employer–employee relationships it is the agent who implements the production level by making a costly effort decision. However, even when effort is verifiable, our model might not be directly applicable in such a setup. The reason is that the contractually specified effort level might not be enforceable, since the agent cannot be fined for breaking the contract when he has no resources (which rules out monetary punishments).

Yet, there may also be circumstances under which our model could be applied in an employment setting, since the employee might be willing to adhere to the contract for reputational concerns (say, because the employer could threaten to pass the employee over for promotion in the future, or the employee might fear being judged a job hopper if he leaves early). Moreover, both the limited liability assumption and the assumption that the cost of reaching a given output level is private information and only learned on the job are very natural in the employment setting. Thus, our upward distortion result could provide a novel explanation for why employees sometimes work too much compared to what would be socially desirable.[27]

Finally, while there is a large contract-theoretic literature on adverse selection models in which the agent has pre-contractual private information, hidden information models in which the agent becomes privately informed after the contract has been signed have received somewhat less attention.[28] In environmental problems in practice, it is likely the case that the parties have already some information when the contract is written, while additional information is learned later on. Studying hybrid models with both pre-contractual and post-contractual private information and bounded payments might be an interesting avenue for future research.[29]

---

[24] If $v \le E[c_i]$, then the expected total surplus in the first-best solution converges to zero when n goes to infinity.

[25] In settings with pre-contractual private information, Samuelson (1985) and McKelvey and Page (2002) study whether the polluter or the pollutee should have the relevant property rights. See also Matouschek (2004) and the Segal and Whinston (2013) for related models.

[26] The reason is that x^{LL} maximizes expected surplus subject to some constraints as stated in Lemma 3, and $x^{\max}$ also satisfies these constraints.

[27] The fact that overwork may be a severe problem has been suggested by several empirical studies, see e.g. Galinsky et al. (2005).

[28] See e.g. the recent work by Iossa and Martimort (2015) and the literature discussed there.

[29] For an early paper that combines pre-contractual and post-contractual private information, see Riordan and Sappington (1987). See also Crémer and Khalil (1994) and Crémer et al. (1998), where the agent can gather pre-contractual information about his type (which otherwise he will learn after the contract is signed). Yet, these papers do not study the implications of bounded transfer payments.

Appendix

Proof of Lemma 1.

First, assume that (IC1) and (IC2) hold. With the transfer function defined by (IC1), the agent's utility $u(c) = t(c) - x(c)c$ is equal to $u(c) = u(c_H) + \int_c^{c_H} x(\gamma)d\gamma$, and (IC) is satisfied:

$$u(c) - u(\hat{c}) = \int_c^{\hat{c}} x(\gamma)d\gamma \geq (\hat{c} - c)x(\hat{c}),$$

where we have used that x is weakly decreasing. Second, let (IC) be satisfied. Assuming $\hat{c} > c$, (IC) implies $t(c) - x(c)c \geq t(\hat{c}) - x(\hat{c})c$ and $t(\hat{c}) - x(\hat{c})\hat{c} \geq t(c) - x(c)\hat{c}$, which in turn implies $(\hat{c} - c)x(c) \geq (\hat{c} - c)x(\hat{c})$, hence x must be weakly decreasing. To show that (IC1) is also satisfied, consider first the case of a finite support. It holds that $u(c_i) - u(c_{i+1}) \geq (c_{i+1} - c_i)x(c_{i+1})$ as well as $(c_{i+1} - c_i)x(c_i) \geq u(c_i) - u(c_{i+1})$, which implies for any $c = c_l$

$$\sum_{i=l}^{n-1} x(c_{i+1})(c_{i+1} - c_i) \leq u(c) - u(c_H) \leq \sum_{i=l}^{n-1} x(c_i)(c_{i+1} - c_i). \tag{8}$$

It follows that for fixed $t(c_H)$,

$$t(c) = u(c_H) + x(c)c + \sum_{i=l}^{n-1} x(c_{i+1})(c_{i+1} - c_i)$$

is the pointwise smallest possible payment function such that (x, t) satisfies (IC). For the case that the support is the interval $[c_L, c_H]$, condition (8) holds for any partition $c = c_l < c_{l+1} < \ldots < c_n = c_H$ of the interval $[c, c_H]$, so that in the limit as the partition becomes finer, we get

$$\int_c^{c_H} x(\gamma)d\gamma = u(c) - u(c_H).$$

Proof of Lemma 2.

Assume that a contract (x, t) with $E[t(c) - cx(c)] > \bar{u}$ solved the principal's problem. Note that it cannot be that $x(c) = x^{\max}$ and $t(c) = -\bar{t}$, because this contract would violate the participation constraint ($-x^{\max}E[c] - \bar{t} < \bar{u}$). There would exist a sufficiently small $\varepsilon > 0$ such that the contract $(\tilde{x}, \tilde{t})$ with $\tilde{x}(c) = (1 - \varepsilon)x(c) + \varepsilon x^{\max}$ and $\tilde{t}(c) = (1 - \varepsilon)t(c) - \varepsilon\bar{t}$ satisfies (PC). This new contract would also satisfy (LL) and (IC), and it would yield a strictly higher utility for the principal. Hence, the participation constraint must be binding.

Plugging the payment function given by (IC1) into the agent's expected payoff function yields

$$E[t(c) - cx(c)] = t(c_H) - x(c_H)c_H + \int\int_{[\gamma \geq c]} x(\gamma)d\gamma dF(c).$$

Changing the order of integration (applying Fubini's theorem), we can write the right-hand side as

$$t(c_H) - x(c_H)c_H + \int_{c_L}^{c_H} F(\gamma)x(\gamma)d\gamma.$$

Hence, setting $E[t(c) - cx(c)] = \bar{u}$ yields the expression for $t(c_H)$ as claimed in the Lemma.

Proof of Lemma 3.

With $t(c)$ defined by (IC1) and $t(c_H)$ by (PC) as in Lemma 2, the objective function is equal to $E[V(x) - cx]$ and the limited liability constraint $t(c_H) \geq -\bar{t}$ takes the form of (LL-ICPC). If (LL-ICPC) holds strictly and the principal receives less than $S^{FB} - \bar{u}$, then due to concavity of the objective function, the principal's profit can be increased if x is replaced by $\varepsilon x^{FB} + (1 - \varepsilon)x$ for some small $\varepsilon > 0$ such that (LL-ICPC) is still satisfied.

References

Arrow, K., 1979. The property rights doctrine and demand revelation under incomplete information. In: Boskin, M.J. (Ed.), Economics and Human Welfare: Essays in Honor of Tibor Scitovsky. Academic Press, pp. 23–29.

Baron, D.P., 1985. Regulation of prices and pollution under incomplete information. J. Public Econ. 28, 211–231.

Baron, D.P., Myerson, R., 1982. Regulating a monopolist with unknown costs. Econometrica 50, 911–930.

Bontems, P., Bourgeon, J.M., 2005. Optimal environmental taxation and enforcement policy. Eur. Econ. Rev. 49, 409–435.

Committee on Climate Change. How Local Authorities Can Reduce Emissions and Manage Climate Risk, 2012; London, UK, https://www.theccc.org.uk/wp-content/uploads/2012/05/LA-Report_final.pdf.

Chen, B.R., Chiu, Y.S., 2013. Interim performance evaluation in contract design. Econ. J. 123, 665–698.

Coase, R.H., 1960. The problem of social cost. J. Law Econ. 3, 1–44.

Crémer, J., Khalil, F., 1994. Gathering information before the contract is offered: The case with two states of nature. Eur. Econ. Rev. 38, 675–682.

Crémer, J., Khalil, F., Rochet, J.C., 1998. Strategic information gathering before a contract is offered. J. Econ. Theory 81, 163–200.

d'Aspremont, C., Gérard-Varet, L.A., 1979. Incentives and incomplete information. J. Public Econ. 11, 25–45.

Fudenberg, D., Tirole, J., 1991. Game Theory. MIT Press, Cambridge, MA.

Galinsky, E., Bond, J.T., Kim, S.S., Backon, L., Brownfield, E., Sakai, K., 2005. Overwork in America: When the Way We Work Becomes Too Much. Families and Work Institute, New York.

Hamilton, J.T., 1993. Politics and social costs: estimating the impact of collective action on hazardous waste facilities. RAND J. Econ. 24, 101–125.
Hart, O., Holmström, B., 1987. The theory of contracts. In: Bewley, T. (Ed.), Advances in Economics and Econometrics, Econometric Society Monographs, Fifth World Congress. Cambridge University Press, Cambridge, pp. 71–155.
Hart, O., Moore, J., 2008. Contracts as reference points. Q. J. Econ. 123, 1–48.
Imhof, L., Kräkel, M., 2014. Bonus pools and the informativeness principle. Eur. Econ. Rev. 66, 180–191.
Innes, R.D., 1990. Limited liability and incentive contracting with *ex-ante* action choices. J. Econ. Theory 52, 45–67.
Iossa, E., Martimort, D., 2015. Pessimistic information gathering. Games Econ. Behav. 91, 75–96.
Kenney, D.S., Stohs, M., Chavez, J., Fitzgerald, A., Erickson, T., 2004. Evaluating the Use of Good Neighbor Agreements for Environmental and Community Protection. Natural Resources Law Center, University of Colorado School of Law.
Kessler, A.S., Lülfesmann, C., Schmitz, P.W., 2005. Endogenous punishments in agency with verifiable ex post information. Int. Econ. Rev. 46, 1207–1231.
Khalil, F., Kim, D., Shin, D., 2006. Optimal task design: to integrate or separate planning and implementation. J. Econ. Manag. Strat. 15, 457–478.
Klibanoff, P., Morduch, J., 1995. Decentralization, externalities, and efficiency. Rev. Econ. Stud. 62, 223–247.
Kräkel, M., Schöttner, A., 2016. Optimal sales force compensation. J. Econ. Behav. Organ. 126, 179–195.
Laffont, J.J., Martimort, D., 2002. The Theory of Incentives: The Principal-Agent Model. Princeton University Press, Princeton, NJ.
Levin, J., 2003. Relational incentive contracts. Am. Econ. Rev. 93, 835–857.
Lewis, T.R., Sappington, D.E.M., 1989. Countervailing incentives in agency problems. J. Econ. Theory 49, 294–313.
Lewis, T.R., Sappington, D.E.M., 2000. Motivating wealth-constrained actors. Am. Econ. Rev. 90, 944–960.
Maggi, G., Rodriguez-Clare, A., 1995. On countervailing incentives. J. Econ. Theory 66, 238–263.
Mailath, G.J., Postlewaite, A., 1990. Asymmetric information bargaining problems with many agents. Rev. Econ. Stud. 57, 351–367.
Martimort, D., 2006. An agency perspective on the costs and benefits of privatization. J. Regul. Econ. 30, 5–44.
Matouschek, N., 2004. Ex post inefficiencies in a property rights theory of the firm. J. Law Econ. Organ. 20, 125–147.
McKelvey, R.D., Page, T., 2002. Status quo bias in bargaining: an extension of the Myerson–Satterthwaite theorem with an application to the coase theorem. J. Econ. Theory 107, 336–355.
Myerson, R.B., 1982. Optimal coordination mechanisms in generalized principal-agent problems. J. Math. Econ. 10, 67–81.
OECD, 2016. The Economic Consequences of Outdoor Air Pollution. OECD Publishing, Paris.
Ohlendorf, S., Schmitz, P.W., 2012. Repeated moral hazard and contracts with memory: the case of risk-neutrality. Int. Econ. Rev. 53, 433–452.
Pargal, S., Wheeler, D., 1996. Informal regulation of industrial pollution in developing countries: evidence from indonesia. J. Polit. Econ. 104, 1314–1327.
Pesendorfer, M., 1998. Pollution claim settlements under correlated information. J. Econ. Theory 79, 72–105.
Riordan, M.H., Sappington, D.E.M., 1987. Awarding monopoly franchises. Am. Econ. Rev. 77, 375–387.
Rob, R., 1989. Pollution claim settlements under private information. J. Econ. Theory 47, 307–333.
Rustichini, A., Satterthwaite, M.A., Williams, S.R., 1994. Convergence to efficiency in a simple market with incomplete information. Econometrica 62, 1041–1063.
Samuelson, W., 1985. A comment on the Coase theorem. In: Roth, A.E. (Ed.), Game-Theoretic Models of Bargaining. Cambridge University Press, pp. 321–339.
Sappington, D., 1983. Limited liability contracts between principal and agent. J. Econ. Theory 29, 1–21.
Segal, I., Whinston, M.D., 2013. Property rights. In: Gibbons, R., Roberts, J. (Eds.), Handbook of Organizational Economics. Princeton University Press, pp. 100–158.
Selmi, D.P., 2010. The contract transformation in land use regulation. Stanf. Law Rev. 63, 591–646.
Shetty, S., 1988. Limited liability, wealth differences and tenancy contracts in agrarian economies. J. Dev. Econ. 9, 1–22.
Tamada, Y., Tsai, T.S., 2014. Delegating the decision-making authority to terminate a sequential project. J. Econ. Behav. Organ. 99, 178–194.
Ter Mors, E., Terwel, B.W., Daamen, D.D., 2012. The potential of host community compensation in facility siting. Int. J. Greenh. Gas Control 11, 130–138.
Tirole, J., 1999. Incomplete contracts: where do we stand? Econometrica 67, 741–781.
Wirl, F., Huber, C., 2005. Voluntary internalisations facing the threat of a pollution tax. Rev. Econ. Dyn. 9, 337–362.