DISCUSSION PAPER SERIES

DP11217

POLLUTION CLAIM SETTLEMENTS RECONSIDERED: HIDDEN INFORMATION AND BOUNDED PAYMENTS

Susanne Goldlücke and Patrick W. Schmitz

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

POLLUTION CLAIM SETTLEMENTS RECONSIDERED: HIDDEN INFORMATION AND BOUNDED PAYMENTS

Susanne Goldlücke and Patrick W. Schmitz

Discussion Paper 11217
Published 07 April 2016
Submitted 07 April 2016

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

POLLUTION CLAIM SETTLEMENTS RECONSIDERED: HIDDEN INFORMATION AND BOUNDED PAYMENTS

Abstract

A pollution-generating firm (the principal) can offer a contract to an agent (say, a nearby town) who has the right to be free of pollution. Subsequently, the agent privately learns the disutility caused by pollution. Then a production level and a payment from the principal to the agent are implemented as contractually specified. We explore the implications of a non-negativity constraint on the payment. For low cost types there is underproduction, while for high cost types there is overproduction. Hence, there may be too much pollution compared to the first-best solution (which is in contrast to standard adverse selection models).

JEL Classification: D82, D86, D23, D62, H23

Keywords: hidden information, externalities, Coasean bargaining, incentive contracting, limited liability

Patrick W. Schmitz - patrick.schmitz@uni-koeln.de
University of Cologne and CEPR

Susanne Goldlücke - susanne.goldluecke@uni-konstanz.de
University of Konstanz

Acknowledgements
Patrick Schmitz gratefully acknowledges financial support under the Institutional Strategy of the University of Cologne within the German Excellence Initiative (Hans-Kelsen-Prize 2015).

1 Introduction

We revisit the classical problem of internalizing a negative externality through Coasean bargaining. Consider a principal who can choose 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 relevant property rights, so in the absence of an agreement between the two parties the production level has to be zero.

Following Rob (1989), 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, 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 from pollution after the contract is written (see 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] However, we assume that payments from the principal to the agent must be non-negative. Such a constraint may be relevant if the agent has no resources or if it is politically

[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.

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 chooses a contractible action and there are more than two types.[2]

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 most efficient 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.[3]

[2] 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). See Laffont and Martimort (2002, section 3.5) for a discussion of two-type hidden information models with limited liability constraints.

[3] 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. In particular, see Lewis and Sappington (1989) for an adverse selection model with countervailing incentives due to type-dependence of the agent's reservation utility, and see Kessler et al. (2005) for an adverse selection model where a signal that is correlated with the type of a wealth-constrained agent can be verified ex post. In contrast to these papers, we follow Rob (1989) in assuming that the agent's reservation utility is zero and

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).[4] 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 hidden action literature.[5] 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

there are no verifiable ex post signals.

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

[5]See, for instance, Bolton and Dewatripont (2005, section 4.1.2). 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. Applications of principal-agent models where payments must be non-negative include e.g. the work by Hiriart and Martimort (2006) and Hiriart et al. (2010) on the regulation of risky activities and the recent work by Martimort and Straub (2016) on infrastructure contracts.

on the agent's ex post utility.[6] 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. However, he studies an adverse selection problem with multiple agents whose cost types (which are privately known 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.

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. The second-best output schedule is flat for low cost types and then decreasing, and always strictly below the first-best output schedule. In contrast, our model has only a one-sided bound on transfers, and the optimal output schedule is decreasing for low cost types and then flat, intersecting the first-best output schedule from below. While the principal in Levin's (2003) model cannot promise the large rewards that would be needed for high output levels, this problem is absent in our model, in which the principal loses only her ability to punish low output in an ex post efficient way.

[6] See e.g. Stole (1994) and Lawarrée and Van Audenrode (1996). Note that in standard hidden information problems, 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 chooses 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 chooses 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.

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. Production has a negative externality on the agent. If the principal was 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 that V is strictly increasing on $[0, x^{nl}]$, so that we can take $x^{\max} = x^{nl}$.

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 $c \in [c_L, c_H]$, with $c_H > c_L \geq 0$. Cost types are distributed according to a probability distribution function F, which is concave.[7] In the main part of the paper, we assume that the support of F is the whole interval $[c_L, c_H]$ and F is differentiable with density $f > 0$. We also consider the case that the support is finite and equal to $\{c_1, ..., c_n\}$ with $c_L = c_1 < ... < c_n = c_H$.

The principal proposes a contract, then the agent accepts or rejects. The agent's reservation utility is 0, which means that when no agreement between the parties is reached, the agent has the right to be free of pollution. After the agent has accepted the contract, he privately learns the realized cost parameter c. Finally, the contract is executed: The principal chooses output 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

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

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

[7] For all cost parameters c in the support of F with $c = \lambda c' + (1 - \lambda)c''$ for some $\lambda \in [0, 1]$ and $c', c'' \in [c_L, c_H]$, it holds that $F(c) \geq \lambda F(c') + (1 - \lambda)F(c'')$.

residual claimant for the profit and lets the principal receive the entire expected surplus.[8]

Remark 1 *If transfer payments are unbounded, the first-best solution will be attained with a payment* $t(x) = V(x) - S^{FB}$.

The optimal contract may require large transfer payments from the agent to the principal in the case of large external effects. Limits on the agent's wealth or political constraints may render this contract infeasible. In what follows, we thus assume that the agent can make no transfer payments to the principal; i.e., we impose the "limited liability" constraint $t \geq 0$ (adopting the wording of Pesendorfer, 1998). Note that this constraint would have no effect if the principal and the agent both knew the agent's cost type.[9] Yet, given that the constraint $t \geq 0$ 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, a general contract is of the form $(x(c), t(c))_{c \in [c_L, c_H]}$, 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 subsidies, and that induce truth-telling by the agent. Hence, the contract has to satisfy the participation constraint

$$E[t(c) - cx(c)] \geq 0, \tag{PC}$$

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

[9]In this case, the principal would offer the contract $t = cx \geq 0$, so the first-best solution would be achieved.

the limited liability constraint

$$t(c) \geq 0 \text{ for all } c, \tag{LL}$$

and the incentive compatibility constraint

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

We will look at each constraint in turn. First, if we solved this problem without the participation constraint, the solution would have $t = 0$, $x = x^{\max}$ and thus violate the participation constraint.

Lemma 1 *In the optimum, the participation constraint is binding.*

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.[10]

An output schedule x is *implementable* if there exist transfer payments such that (IC) is satisfied. The following standard result holds:

Lemma 2 *A function x is implementable if and only if*

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

The pointwise smallest payment function that implements x and satisfies (LL) is

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

This formula also holds for the case of a finite support if we extend functions $x : \{c_1, ..., c_n\} \to [0, x^{\max}]$ on the interval $[c_L, c_H]$ by defining $x(c) = x(c_{i+1})$ for $c \in (c_i, c_{i+1}]$.

Proof. First, assume that x is weakly decreasing. 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:

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

[10] See e.g. Laffont and Martimort (2002, section 4.3).

where we have used that x is weakly decreasing. Second, let (IC) be satisfied. Assuming $\hat{c} > c$, (IC) implies $t(c) - x(c)c \geq t(\hat{c}) - cx(\hat{c})$ and $t(\hat{c}) - x(\hat{c})\hat{c} \geq t(c) - \hat{c}x(c)$, which in turn implies $(\hat{c} - c)x(c) \geq (\hat{c} - c)x(\hat{c})$, hence x must be weakly decreasing. For the case of a finite support, it holds for any $c = c_l$ 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 that

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

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 lowest possible payment that implements x. For the case that the support is the interval $[c_L, c_H]$, condition (1) holds for any partition $c = c_l < c_{l+1} < ... < c_n = c_H$ of the interval $[c, c_H]$, so that in the limit as the partition becomes finer

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

Since transfers are weakly decreasing, they satisfy the limited liability constraint if and only if $t(c_H) \geq 0$, the smallest possible payments having $t(c_H) = 0$. ■

Since maximizing without the limited liability constraint would lead to the first-best solution, this constraint must be binding if the first-best surplus cannot be attained.

Lemma 3 *If for any function x^{FB} with*

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

it holds that

$$\int_{c_L}^{c_H} x^{FB}(c)F(c)dc \leq x^{FB}(c_H)c_H, \tag{2}$$

then the principal obtains the first-best surplus S^{FB}. Otherwise, the first-best surplus is not obtainable and the limited liability constraint is binding.

Proof. Plugging the pointwise lowest payment function that implements x^{FB}, as defined in Lemma 2, into the agent's expected payoff function yields

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

Changing the order of integration, we can write this as

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

If this expression is negative, the principal can gain the maximum amount S^{FB} by increasing the fixed payment to the agent until the participation constraint is satisfied. If this expression is positive, the participation constraint is not binding and hence x^{FB} cannot be optimal. ■

As in Lemma 2, this result also holds for the finite case.

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, such that the support of F is $\{c_L, c_H\}$. In many settings, the two-type case provides much of the intuition of the general case. Here, the principal can still achieve the first-best solution even when payments must be non-negative.

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

Yet, 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 shows that the two-type case is misleading in our setting.

Let the support of F be given by $\{c_L, c_M, c_H\}$, and let f_L, f_M, f_H denote the associated probabilities. We assume $x^{\max} = 1$ 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 $x^{FB}(c_H) = 0$.[11] Lemma 2 tells us that the cheapest way to make the cost type c_i choose quantity x_i with $x_L \geq x_M \geq x_H$ is to choose payments $t_H = 0$ for the

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

quantity x_H, $t_M = (x_M - x_H)c_M$ for the quantity x_M, and $t_L = t_M + (x_L - x_M)c_L$ for the quantity x_L. Lemma 3 clarifies that the first-best solution cannot be achieved here since $f_L(c_M - c_L) > 0$, so that indeed $t_H = 0$. 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 participation constraint

$$f_L((x_M - x_H)c_M - c_L x_M) - f_M c_M x_H - f_H c_H x_H = 0.$$

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

$$x_H = \frac{x_M f_L(c_M - c_L)}{(1 - f_H)c_M + f_H c_H}. \tag{3}$$

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

$$f_M(v - c_M) + \frac{f_L(c_M - c_L)(v - c_H)f_H}{(1 - f_H)c_M + f_H c_H} > 0, \tag{4}$$

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.[12] The intuition behind this result 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.[13]

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 (4) is relaxed. Similarly, as c_H increases, expected output goes down. More surprisingly,

[12] For example, if $f_L = f_M = f_H = \frac{1}{3}$, $c_L = 1, c_M = 2, c_H = 4$, and $v = 3$, we have $x_M = 1$ and $x_H = \frac{1}{8}$.

[13] For example, if $f_L = f_M = f_H = \frac{1}{3}$, $c_L = 1, c_M = 2, c_H = 4$, 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.

x_H increases in c_M for $x_M = 1$, so as long as condition (4) 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 (4) becomes true.

In the following, we assume that the cost parameter can take any value in the interval $[c_L, c_H]$ and is distributed according to a c.d.f. F with $F' = f > 0$.

Proposition 1 *If condition (2) does not hold, the optimal contract x^{LL}, t^{LL} must have the following properties:*

(i) There exists a $\lambda \in (0, 1)$, a cut-off $\bar{c}$, a value $\bar{x}$, and a non-increasing function

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

such that

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

(ii) It holds that $x^{LL}(c_L) = x^{FB}(c_L)$ and $x^{FB}(\bar{c}) > \bar{x} > x^{FB}(c_H)$.

Proof. Since (2) does not hold, the limited liability constraint must be binding. Since the participation constraint is binding, the objective function is equal to $E[V(x) - cx]$. Using (IC1), the participation constraint becomes

$$-x(c_H)c_H + \int_{c_L}^{c_H} x(c)F(c)dc = 0. \qquad \text{(ICPC)}$$

The maximization problem is

$$\begin{aligned} &\max_{x(.)} E[V(x) - cx] \qquad (5) \\ s.t. \quad &(ICPC), (IC2). \end{aligned}$$

We write $\bar{x}$ instead of $x(c_H)$ in (ICPC) and replace (IC2) by the weaker constraint $x(c) \geq \bar{x}$. We consider the following maximization problem over both x and $\bar{x}$:

$$\begin{aligned} &\max_{x(.),\bar{x}} \int_{c_L}^{c_H} (V(x(c)) - cx(c))f(c)dc \qquad (6)\\ s.t \quad &\int_{c_L}^{c_H} x(c)F(c)dc = \bar{x}c_H \\ &x(c) \geq \bar{x}. \end{aligned}$$

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 λ and a function $\mu \geq 0$ with $\mu(c)(x^*(c) - \bar{x}^*) = 0$ such that $x^*, \bar{x}^*$ solve

$$\max_{x(.),\bar{x}} \int_{c_L}^{c_H} (V(x(c)) - (c + \lambda \frac{F(c)}{f(c)})x(c) + \lambda c_H \bar{x} + \mu(c)(x(c) - \bar{x}))f(c)dc. \qquad (7)$$

It can be directly seen from (6) that $\bar{x}^*$ cannot be as low as 0 or as high as $x^{\max}$. To get an interior solution for $\bar{x}^*$, it must hold that $\lambda c_H = E[\mu]$ which implies that $\lambda \geq 0$. Since we assumed that the first-best surplus is not attainable for the principal, it must be that $\lambda > 0$.

For a given c, $V(x) - (c + \lambda \frac{F(c)}{f(c)})x$ has a unique maximizer if V is strictly concave. Taking into account that the set of maximizers can also be an interval, we denote the largest maximizer by

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

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 $\hat{x}(c') \leq \hat{x}(c)$ for $c' > c$. Hence, $x^*(c) = \max\{\hat{x}(c), \bar{x}^*\}$ is indeed weakly decreasing. Moreover, 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 $\hat{x}(c) = x^{\max} = x^{FB}(c)$.

If we set $\mu(c) = 0$ for $c \leq \bar{c}$ and $\mu(c) = -V'(\bar{x}^*) + (c + \lambda\frac{F(c)}{f(c)})$ else, then x^* indeed maximizes (7). From $\lambda c_H = \int_{\bar{c}}^{c_H} -V'(\bar{x}^*) + (c + \lambda\frac{F(c)}{f(c)})f(c)dc$ it follows that

$$\lambda = \frac{\int_{\bar{c}}^{c_H}(c - V'(\bar{x}^*))f(c)dc}{\bar{c}F(\bar{c}) + \int_{\bar{c}}^{c_H} cf(c)dc} < 1. \tag{8}$$

The fact that $\lambda > 0$ implies that $c_H > V'(\bar{x}^*)$, which means that $\bar{x}^* > x^{FB}(c_H)$ and $x^*(c_H) = \bar{x}^*$. Moreover, by definition $\bar{x}^* \leq \hat{x}(\bar{c})$ and therefore either $\bar{x}^* = \hat{x}(\bar{c}) < x^{FB}(\bar{c})$ or $\bar{x}^* < x^{FB}(\bar{c}) = x^{\max}$. ■

Compared to the first-best solution, high cost types need to accept higher levels of the externality, low cost types (weakly) lower levels. Due to the post-contractual asymmetric information, the principal has to make large payments to the agent for high levels of the externality. When there is limited liability, the principal cannot extract the low cost types' rent with a payment. Therefore, there are two forces in the model. On the one hand, the principal wants to reduce the low cost types' rents, which she does by reducing pollution if costs are low. On the other hand, the principal extracts these agent rents with generous pollution limits if costs are high.

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.[14] 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,

$$t(c) - cx(c) \geq 0 \text{ for all } c. \tag{IR}$$

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

Note that in the adverse selection model it is irrelevant whether or not we impose the limited liability constraint, because the non-negativity of the transfer payments is already implied by (IR).

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

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

and therefore corresponds to $\lambda = 1$ in the function $\hat{x}$ in Proposition 1. The proof of Proposition 1 shows that $\hat{x}$ is non-increasing in λ. ■

Figure 1 shows how the functions x^{FB}, x^{LL}, and x^{AS} compare in a numerical example.

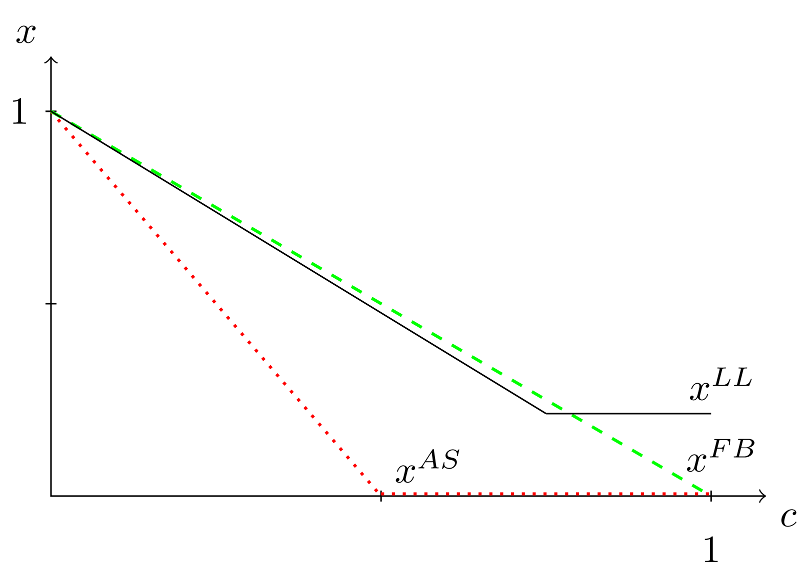

Figure 1. This figure shows x^{LL} and how it compares to x^{AS} (red) and x^{FB} (green) for the example $V(x) = x(1 - \frac{x}{2})$, $x^{\max} = 1$, and c uniformly distributed on $[0, 1]$.

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 as least as large as in the case of pre-contractual private information. 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 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 Figure 2.

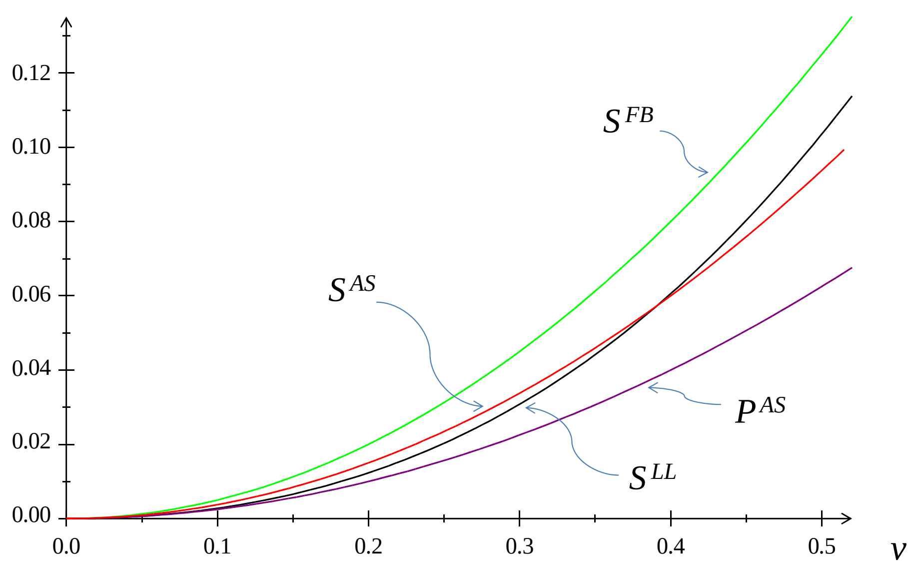

Figure 2. In the figure, $V(x) = vx$, $x^{\max} = 1$, and c is uniformly distributed on $[0, 1]$. The figure shows the expected total surplus in the case of post-contractual private information and limited liability (S^{LL}), compared to the expected total surplus (S^{AS}) and the principal's expected profit (P^{AS}) in the case of pre-contractual private information.

4.2 Extensions of the model in the linear case

In this section we assume that $V(x) = vx$ and $x^{\max} = 1$, so x can be interpreted as the probability of production. We assume that $v < c_H$, such that $x^{FB}(c_H) = 0$ and hence the optimal contract differs from the first-best solution. Proposition 1 tells us that there is a cut-off $\bar{c} < v$ such that the probability of production is equal to $x(c) = 1$ for cost types below $\bar{c}$, while for cost types above $\bar{c}$ it is equal to

$$\bar{x} = \frac{\int_{c_L}^{\bar{c}} F(c)dc}{c_H - \int_{\bar{c}}^{c_H} F(c)dc},$$

which is calculated from the binding (ICPC) constraint in the proof. Moreover, it must hold that $v = \bar{c} + \lambda \frac{F(\bar{c})}{\bar{c}}$, where λ is given by (8). The threshold $\bar{c}$ can then be found by solving

$$(v - \bar{c})f(\bar{c})(c_H - \int_{\bar{c}}^{c_H} F(c)dc) + F(\bar{c}) \int_{\bar{c}}^{c_H} (v - c)f(c)dc = 0. \tag{9}$$

The linear case can now be used to discuss some extensions of our model.

First, we have assumed that the reservation utility of the agent and the agent's wealth are equal (that both are set equal to zero is just a normalization). If we assume instead that the agent's wealth is given by $\bar{w} \geq 0$, the limited liability constraint becomes $t \geq -\bar{w}$. Our results can easily be adjusted to include this new parameter. In the linear case, the first-best solution can be achieved if

$$\int_{c_L}^{v} F(c)dc \leq \bar{w}.$$

If this is not the case, high cost types above $\bar{c}$ pay $\bar{w}$ and accept a production level of

$$\bar{x} = \frac{\int_{c_L}^{\bar{c}} F(c)dc - \bar{w}}{c_H - \int_{\bar{c}}^{c_H} F(c)dc},$$

while $\bar{c}$ is independent of $\bar{w}$ and still given by (9). If $\bar{w}$ decreases, production levels are higher for all cost types. The agent's expected utility stays the same, but very high cost types lose and low cost types gain (the cut-off type $\bar{c}$ gains).

Second, 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. As the number of agents becomes large, holding per-person profits constant by setting $V(x) = nvx$, the obstacles to bargaining due to the 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.[15] 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 $v > E[c_i]$.[16] 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.

[15] See Mailath and Postlewaite (1990) for a closely related limit result in the context of public goods. In contrast, Gresik and Satterthwaite (1989) 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.

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

Third, one could extend the model to include asymmetric information about the profitability of production. Suppose that v and c are independently distributed random variables and the principal privately knows the realization of v when she proposes the contract to the agent. The optimal contract would not be affected, since the agent's incentive and participation constraints only depend on the promised payment schedules.[17] However, private information about profitability might become meaningful if the principal could not commit to all contracts, but would go bankrupt if profit is negative. Again, the optimal contract is robust to such changes, but we can rationalize our assumption that there can be no subsidies with such a bankruptcy constraint and private information on the principal's side. Specifically, assume that the principal is either a profitable firm (making profit v) with probability p, or an unprofitable firm (making zero profit) with probability $1-p$. The principal then offers a menu of contracts and lets the agent choose as before, but an unprofitable firm only honors the contract if $t \leq 0$. Offering a contract to the agent becomes a signalling game, in which the profitable firm offering x^{LL}, t^{LL} and the unprofitable firm offering no contract is a separating equilibrium. Thus, by introducing the possibility that the principal has no interest in production and only wants to collect the subsidy in case of high costs, one can formalize another reason for the bound on payments in our model.

5 Concluding remarks

We have studied a contracting problem in which the principal can choose a verifiable production level, the agent learns his cost type after the contract is signed but before production takes place, and payments to the agent must be non-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 ad-

[17] Cf. Maskin and Tirole's (1990) discussion of informed principal problems for the case of quasi-linear preferences.

verse selection problem in which the agent learns his cost type before the contract is written.

Situations in which a firm's production causes negative externalities such as pollution abound in practice. Related examples include the siting of waste dumps, power plants, electricity pylons, wind turbines, homeless shelters, adult entertainment clubs, drug consumption rooms, or refugee hostels. 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. The principal would then choose $x^{\max}$, and it depends on the parameter constellation whether the expected total surplus would be smaller or larger than in the solution that we have characterized. Moreover, following Rob (1989) we have assumed that the principal can make a take-it-or-leave-it offer. To capture the specifics of some real-world applications, in future research it might be worthwhile to also study negotiation games in which the agent has some bargaining power.

In other contexts such as employer–employee relationships it is the agent who chooses the production level (say, by making a verifiable effort decision). In this case, our model can still be applied if the contractually specified effort level can be enforced even when the agent has no resources (which rules out monetary punishments for shirking). For instance, while the employee may not be able to make payments to the employer, the employee might still be willing to adhere to the contract, because the employer can threaten to pass the employee over for promotion in the future (or to fire him and write a bad reference letter), or the employee might fear of being judged a job hopper if he leaves early. Under such circumstances, our model can be applied and our upward distortion result provides a novel explanation for why employees sometimes work too much compared to what would be socially desirable.

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. 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 a interesting avenue for future research.[18]

[18] For an early paper that combines pre-contractual and post-contractual private information, see Riordan and Sappington (1987). Yet, they do not study the implications of bounded transfer payments.

References

Bolton, P., Dewatripont, M. (2005). *Contract Theory.* Cambridge, MA: MIT Press.

Coase, R.H. (1960). The problem of social cost. *Journal of Law and Economics* 3, 1–44.

d'Aspremont, C., Gérard-Varet, L.A. (1979). Incentives and incomplete information. *Journal of Public Economics* 11, 25–45.

Gresik, T.A., Satterthwaite, M.A. (1989). The rate at which a simple market converges to efficiency as the number of traders increases: An asymptotic result for optimal trading mechanisms. *Journal of Economic Theory* 48, 304–332.

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: Cambridge University Press, 71–155.

Hiriart, Y, Martimort, D. (2006). The benefits of extended liability. *RAND Journal of Economics* 37, 562–582.

Hiriart, Y, Martimort, D., Pouyet, J. (2010). The public management of risk: Separating ex ante and ex post monitors. *Journal of Public Economics* 94, 1008–1019.

Innes, R.D. (1990). Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory* 52, 45–67.

Kessler, A.S., Lülfesmann, C., Schmitz, P.W. (2005). Endogenous punishments in agency with verifiable ex post information. *International Economic Review* 46, 1207–1231.

Klibanoff, P., Morduch, J. (1995). Decentralization, externalities, and efficiency. *Review of Economic Studies* 62, 223–247.

Laffont, J.-J., Martimort, D. (2002). *The Theory of Incentives: The Principal-Agent Model.* Princeton, N.J.: Princeton University Press.

Lawarrée, J.P., Van Audenrode, M.A. (1996). Optimal contract, imperfect output observation, and limited liability. *Journal of Economic Theory* 71, 514–531.

Levin, J. (2003). Relational incentive contracts. *American Economic Review* 93, 835–857.

Lewis, T.R., Sappington, D.E.M. (1989). Countervailing incentives in agency problems. *Journal of Economic Theory* 49, 294–313.

Mailath, G.J., Postlewaite, A. (1990). Asymmetric information bargaining problems with many agents. *Review of Economic Studies* 57, 351–367.

Martimort, D, Straub, S. (2016). How to design infrastructure contracts in a warming world: A critical appraisal of public-private partnerships. *International Economic Review* 57, 61-88.

Maskin, E., Tirole, J. (1990). The principal-agent relationship with an informed principal: The case of private values. *Econometrica* 58, 379–409.

Pesendorfer, M. (1998). Pollution claim settlements under correlated information. *Journal of Economic Theory* 79, 72–105.

Riordan, M.H., Sappington, D.E.M. (1987). Awarding monopoly franchises. *American Economic Review* 77, 375–87.

Rob, R. (1989). Pollution claim settlements under private information. *Journal of Economic Theory* 47, 307–333.

Sappington, D. (1983). Limited liability contracts between principal and agent. *Journal of Economic Theory* 29, 1–21.

Stole, L.A. (1994). Information expropriation and moral hazard in optimal second-source auctions. *Journal of Public Economics* 54, 463–484.

Tirole, J. (1999). Incomplete contracts: Where do we stand? *Econometrica* 67, 741–781.