Nº 311
Julio 2006

Documento de Trabajo

ISSN (edición impresa) **0716-7334**
ISSN (edición electrónica) **0717-7593**

A Simple Auction Mechanism for the Optimal Allocation of the Commons

Juan Pablo Montero

www.economia.puc.cl

Versión impresa ISSN: 0716-7334
Versión electrónica ISSN: 0717-7593

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA

Oficina de Publicaciones
Casilla 76, Correo 17, Santiago
www.economia.puc.cl

A SIMPLE AUCTION MECHANISM FOR THE OPTIMAL ALLOCATION OF THE COMMONS

Juan-Pablo Montero*

Documento de Trabajo Nº 311

Santiago, Julio 2006

* jmontero@faceapuc.cl

INDEX

ABSTRACT		1
1.	INTRODUCTION	1
2.	THE MODEL	6
	2.1. Notation and first-best allocation	6
	2.2. Kwerel’s scheme	6
3.	THE AUCTION MECHANISM	8
	3.1. Single firm	9
	3.2. Multiple firms: Non-cooperative equilibrium	12
	3.3. Properties of the mechanism	15
4.	COLLUSION	18
	4.1. The optimal collusive agreement	18
	4.2. Implementing the collusive agreement	19
5.	EXTENSIONS	23
	5.1. Imperfect substitutability of licenses	23
	5.2. Private externalities	25
6.	FINAL REMARKS	27
APPENDIX A: PROOF OF PROPOSITION 1		28
APPENDIX B: PROOF OF PROPOSITION 5		30
REFERENCES		32

A simple auction mechanism for the optimal allocation of the commons

Juan-Pablo Montero*

May 1, 2006

Abstract

Efficient regulation of the commons requires information about the regulated firms that is rarely available to regulators (e.g., cost of pollution abatement). Different mechanisms have been proposed for inducing firms to reveal their private information but for reasons I discuss in the paper, I find these mechanisms of limited use. I propose a much simpler mechanism that implements the first-best for any number of firms: a uniform price sealed-bid auction of an endogenous number of (transferable) licenses with a fraction of the auction revenues given back to firms. Paybacks, which decrease with the number of firms, are such that truth-telling is a dominant strategy regardless of whether firms behave non-cooperatively or collusively. (*JEL* D44, D62, D82)

1 Introduction

Regulatory authorities generally find that part of the information they need for implementing an efficient regulation is in the hands of those who are to be regulated. Regulating externalities such as access to common resources (e.g., clean air, water streams, fisheries, etc.) is not the exception. Environmental regulators, for example, know little about firms' pollution abatement costs, so without communicating with firms they would be unable to establish the efficient level of pollution. Different mechanisms have been proposed for inducing firms to reveal their private

*Associate Professor of Economics at the Catholic Univertity of Chile (jmontero@faceapuc.cl) and Research Associate at the MIT Center for Energy and Environmental Policy Research. This article was written while I was visiting the Harvard's Kennedy School of Government under a Repsol YPF-Harvard Kennedy School Research Fellowship. I would like to thank Tony Creane, Evan Kwerel, Matti Liski, Marty Weitzman, seminar participants at Harvard University, Texas A&M University, University of Maryland and University of Massachusetts-Amherst, and specially Bill Hogan for many reactions and comments. Of course, all errors and omissions are mine.

information but for reasons I discuss below I find these mechanisms of limited use. In this paper, I propose a simpler and more effective mechanism: a uniform price sealed-bid auction of an endogenous number of (transferable) licenses with a fraction of the auction revenues given back to firms.[1] The mechanism is developed under the additional assumption that firms know nothing about the other firms' characteristics (they may be even unaware of the number of firms being regulated).

Following Weitzman (1974), several authors have looked for ways in which to improve upon his fixed tax or license scheme. Roberts and Spence's (1976) hybrid tax/license scheme can in principle implement the first-best when there is an infinitely large number of firms and the regulator is free to impose a tax schedule (as opposed to a fixed tax) and issue a continuum of license types, with each type clearing at a different price. Building also upon the assumption of perfect competition in the license market, Kwerel (1977) develops a much simpler subsidy/license scheme that supposedly implements the first-best; although not in dominant strategies. Relaxing the perfect competition assumption, Dasgupta, Hammond and Maskin (1980) propose a Vickrey-Clarke-Goves (VCG) mechanism that implements the first-best in dominant strategies.[2] Maintaining the few-firms assumption and adding complete information by firms, in a more recent paper Duggan and Roberts (2002) advance a quantity-based scheme in which each firm chooses the number of licenses for itself and for its "neighbor."[3]

With the exception of Kwerel (1977), the fact that these first-best mechanisms are highly non-linear, firm-specific and/or depending on a complete information assumption complicates their practical implementation.[4] In fact, we do not see anything like it being applied in practice or, at least, under consideration. In Kwerel's scheme, on the other hand, the regulator allocates a fixed number of transferable licenses to firms (possibly proportional to historic use as typically done in practice) and establishes a subsidy per license to be paid to any firm holding licenses in excess of its needs. Both the total number of licenses and the subsidy level are calculated on the basis of the information provided by firms. Although this scheme has been criticized because it requires perfect competition in the license market,[5] it has the advantage of being constructed

[1]Licenses are generally refered to as permits or allowances in water and air pollution control, as rights in water supply management and as quotas in fisheries management. In this paper, I will use the term license throughout.

[2]See also Spulber (1988).

[3]In the context of pure private externalities, Varian (1994) also offers a simple multistage mechanism in which firms announce Pigouvian taxes. Firms are, as in Duggan and Roberts (2002), assumed to have complete information about other firms' characteristics.

[4]Evidence of significant information asymmetries across firms facing a commons problem is provided, for example, by Wiggins and Libecap (1985).

[5]This assumption seems adequate for many of the existing license markets (Tietenberg, 2003). Nevertheless,

upon instruments currently used in practice.[6] The problem with Kwerel's scheme, however, is that it does not work. I will show below that the combination of a free allocation of licenses with a subsidy creates perverse incentives for firms not to reveal their true types, as originally claimed, but to over-report their demand for licenses to the maximum extent possible.

One might argue that a quick way to fix Kwerel's scheme, while maintaining its relatively simple structure, is by allocating the licenses via a uniform-price auction in which each firm bids a demand schedule indicating the number of licenses willing to purchase at any given price. Unfortunately, these multi-unit auctions have their own problems as well. As first recognized by Wilson (1979) in his pioneer "auctions of shares" article (see also Milgrom [2004]), even when there is a large number of bidders, uniform price auctions can exhibit Nash equilibria with prices far below the competitive price (the price that would prevail if all bidders submit their true demand curves). The reason for this is that uniform pricing creates strong incentives for bidders to (non-cooperatively) reduce their demand schedules in order to depress the price they pay for their inframarginal units. Therefore, anticipating a low-price equilibrium at the auction (most likely zero, or the reserve price if there is any), firms would find it again profitable to over-report their demand functions by as much as possible in order to acquire a large volume of licences that can then be sold back to the government at a price much higher than the auction clearing price.[7]

Different solutions have been advanced in dealing with this low-price equilibria phenomenon. One radical solution is to give up the uniform-price format altogether and opt for a discriminatory-price format (e.g., Ausubel, 2004; Vickrey, 1961). But within the uniform-price format, different authors have also been looking for ways in which changing auction rules could eliminate underpricing. Kremer and Nyborg (2004), for example, propose changing the allocation rule (i.e., the way the asset is divided when there is excess demand at the clearing price) from the usual marginal pro rata tie-breaking rule to a total pro rata rule.[8] More recently,

Kwerel's scheme does become unworkable if firms' pollutant are not perfect substitutes in the social damage function (Dasgupta et el., 1980).

[6] Note that in Kwerel's truth-telling equilibrium subsidy payments are never triggered.

[7] Some readers may argue that removing the subsidy and having it replaced by a price floor (of equal magnitude) would finally solve matters. As we shall see, it does not do so for the exact opposite reasons. Anticipating the reserve price as the auction clearing price, firms would have incentives to deviate from truth-telling by under-reporting their demand curves in order to decrease the reserve price. The expectation of a (non-cooperative) low-price equilibrium acts as coordinating monopsonist device.

[8] For the change in the allocation rule to have an effect on clearing prices, bidders must be allowed to submit discountinuous demand schedules, which, by construction, is not possible in Wilson (1979). But unlike in Kremer and Nyborg (2004) where bidders have a constant valuation for the asset, in our context this allocation rule change is of little help because bidders do have fairly continuos downward sloping demand curves.

McAdams (2005) eliminates underpricing by letting the auctioneer not to commit to a fixed quantity and reserve price ex-ante. Bidders only learn about the total quantity sold by the auctioneer once the auction is concluded.

In this paper I propose a mechanism that builds upon a conventional uniform price sealed-bid auction but introduces two key ingredients. First, I let the total number of licenses be endogenous to the demand schedules submitted by firms. This is a most natural thing to do in our context because the regulator is clueless about the efficient number of licenses to be allocated before communicating with firms. But unlike in McAdams (2005), this "flexible supply" feature by itself does not fully solve the underpricing problem.[9] Hence, I introduce a second ingredient: paybacks or rebates. Part of the auction revenues are returned to firms not as lump sum transfers but in a way that firms would have incentives to bid truthfully. While paybacks may seem odd in other contexts,[10] they are not new in existing auctions for "protecting the commons".[11] Furthermore, an auction with paybacks seem to be a natural point of departure for any license-type regulatory proposal given the mixed experience with allocating licenses (grandfather allocation vs auction allocation) that is observed in existing programs across a variety of areas including air-pollution control, water supply management and fisheries management (Tietenberg, 2003).[12]

The incorporation of these two key ingredients —endogenous supply and paybacks— into a uniform price auction is not only simple but remarkably effective in implementing the first-best. Paybacks are structured in such a way that truth-telling is a dominant strategy for firms regardless of whether they behave non-cooperatively or collusively. All firms pay the exact same price at the margin for licenses (in or off equilibrium) but their paybacks may differ unless they have identical demand functions. Paybacks, however, will rapidly fall with the number of firms and as the number of firms grow large the auction scheme converges to the Pigouvian solution (Pigou, 1920).

There is a close connection between the auction proposed in this paper and the discriminatory auctions of Vickrey (1961) and Ausubel (2004). Despite the quite different structures,

[9] It only works in the limit, when there is an infinitely large number of firms so that paybacks are virtually zero. Part of the reason why "flexible supply" is not sufficient is because I work with a very different set of assumptions than McAdam (2005). I let firms to be asymmetric, to have downward sloping demand curves and to know nothing about other firms' characteristics. In addition, my auctioneer's objective function is not to maximize revenues but social welfare.

[10] Not surprisingly they are absent in recent books by Milgrom (2004) and Kemplerer (2004).

[11] See, for example, the US EPA auction for sulfur dioxide allowances (Joskow et al, 1998).

[12] See also Cramton and Kerr (1999) for related arguments.

the total payment in the auction mechanism is not different than the payment rules of Vickrey (1961) and Ausubel (2004). The reason is that all these three auctions are constructed upon the same principle. In Vickrey and Ausubel auctions each bidder is asked to pay an amount equal to the "pecuniary externality" she exerts on the other competing bidders while here each bidder is asked to pay an amount equal to the (additional) externality she imposes on society.

The auction mechanism can be certainly viewed as a VCG mechanism in that it makes each firm to pay only for the externality she exerts on society. It departs, however, from the conventional VCG mechanism, as implemented by Dasgupta-Hammond-Maskin (DHM),[13] in at least three fundamental ways. First, payments in the auction mechanism do not follow the two-part structure of DHM (there is, for instance, no such constant term constructed upon other firms' reports). Second, unlike the auction mechanism, the DHM mechanism fails to allocate licenses efficiently across firms when the aggregate supply of licenses is fixed. This is because each individual firm is no longer pivotal under DHM, i.e., its report does not affect the aggregate supply (see, e.g., Milgrom, 2004).[14] This is quite an important distinction because in many commons problems the aggregate supply is likely to be fixed, either because of the presence of some genuine threshold or, more likely, because the auctioneer/regulator has no control over the aggregate supply. Third, the DHM mechanism fails to deliver the first-best when firms are acting collusively unless the constant term is set to zero. When we do that, collusive and non-cooperative behavior are indeed no different but payments become so large (i.e., the full social cost) that the (now Groves) mechanism becomes of little practical value.

The rest of the article is organized as follows. In the next section I present the modelling assumptions and a brief discussion of Kwerel's scheme. In Section 3 I present the auction mechanism; first for a single firm and then for multiple firms. In Section 4 I study how the mechanism performs under collusive behavior. In Section 5 I extend the mechanism to accommodate for firm-specific (i.e., no transferable) licenses and for the existence of private externalities. Concluding remarks are in Section 6.

[13] Recall that in DHM each firm, say i, faces a tax schedule equal to the total pollution damages plus the clean-up cost of all the remaining firms minus a constant term that is independent of firm i's report. Although DHM is not specific about the latter, we know that in a VCG mechanism the constant term is made equal to the social cost (total damages and clean-up costs) associated to the efficient allocation in the absence of firm i.

[14] To be more specific let x be the aggregate supply and n the number of (asymmetric) firms. Under the DHM scheme, one of the multiple (inefficient) Nash equilibria is for each firm to submit a null report and then produce x/n (or alternatively, report that there are no clean up costs up to x/n).

2 The model

To facilitate the exposition I will develop the model for the case of a classical pollution externality (which would correspond to an auction of shares with variable supply). In subsequent sections I extend the model to other commons problems including those in which licenses are firm-specific and where firms impose (private) externalities on each other.

2.1 Notation and first-best allocation

Consider $n \geq 1$ firms ($i = 1, ..., n$). All firms are assumed to have inverse demand functions for pollution of the form $P_i(x_i)$ with $P_i'(x_i) < 0$, where x_i is firm i's pollution level that is accurately monitored by the regulator (In some cases I will work with the demand function, which is denoted by $X_i(p)$ with $X_i'(p) < 0$, where p is the price of pollution). Function $P_i(\cdot)$ is only known by firm i, neither by the regulator nor by other firms. The aggregate demand curve for pollution is denoted by $P(x)$, where $x = \sum_{i=1}^n x_i$ is total pollution. The social damage caused by pollution x is $D(x)$ with $D(0) = 0$, $D'(x) > 0$ and $D''(x) \geq 0$. $D'(x)$ can be interpreted more generally as the regulator's supply function for licenses. We may want to assume that $D(x)$ is publicly known but it is actually not necessary.

In the absence of regulation firm i would emit x_i^0 , where $P_i(x_i^0) = 0$. Hence, firm i's cost of reducing emissions from x_i^0 to some level $x_i < x_i^0$ is $C_i(x_i) = \int_{x_i}^{x_i^0} P_i(z)dz$ (note that $-C_i'(x_i) \equiv P_i(x_i)$), and the minimum total cost of achieving pollution level $x < x^0$ is $C(x) = \int_x^{x^0} P(z)dz$.

The regulator's objective is to minimize the sum of clean-up costs and damages from pollution, i.e., $C(x) + D(x)$. Therefore, the socially optimal or first-best pollution level $x^* < x^0$ satisfies

$$P(x^*) = D'(x^*) = P_i(x_i^*) \quad \text{for all } i = 1, ..., n \tag{1}$$

But the regulator cannot directly implement the first-best allocation because he does not know the demand functions $P_i(\cdot)$. He must then look for mechanisms in which it is in the firms' best interest to communicate their private information to him. Kwerel's scheme is supposedly one of such mechanisms for the case in which there is a large number of firms.

2.2 Kwerel's scheme

To appreciate the workings of my auction scheme it is useful to start by understanding firms' incentives under Kwerel's scheme. This latter proves to be interesting in itself because, as we shall see below, the scheme does not work as intended.

Kwerel's mechanism is based on the combination of two instruments: a free allocation of a total of l transferable licenses and a subsidy of s per license to be paid to any firm holding licenses in excess of its emissions.[15] The regulator asks firms to report their demand curves after they are informed that the parameters l are s are to be set according to

$$s = \hat{P}(l) = D'(l) \tag{2}$$

where $\hat{P}(\cdot)$ is the aggregate demand curve built upon individual reports $\hat{P}_i(x_i)$.

Assuming that the market for licenses is competitive, it must hold in equilibrium that $P_i(x_i) \equiv -C_i'(x_i) = p$ and $x_i = l_i$ for all $i = 1, ..., n$, where p denote the market price of licenses. Firms equate marginal abatement costs to the market price and keep a number of licenses just to cover their emissions. Kwerel argues that this simple scheme induces each firm i to report its true demand curve $P_i(\cdot)$ as long as it believes all other firms are telling the truth. In other words, truth-telling is a (Bayesian) Nash equilibrium.

Kwerel's logic can be easily explained with the aid of Figure 1. Figure 1a depicts the situation in which a firm or a group of firms over-report their demand curves such that the reported aggregate demand curve is $\hat{P}(x)$ instead of the true curve $P(x)$.[16] The license and subsidy parameters take the values of $\hat{l}$ and $\hat{s}$, respectively, which are above their first-best levels l^* and s^*. Since the government is buying back licenses at price $\hat{s}$, the market equilibrium price of licenses is not p' (as if no license were sell back to the government) but $p = \hat{s} > p^*$. On aggregate, firms sell back $\hat{l} - \hat{x}$ licenses, so total pollution falls below its first-best level to $\hat{x} < l^*$. Figure 1b, on the other hand, depicts the under-reporting situation. Given the reported aggregate demand curve $\tilde{P}(x)$, the license and subsidy parameters take now the values of $\tilde{l}$ and $\tilde{s}$, respectively. The market equilibrium price is $p = \tilde{p} > p^*$ and total pollution is $x = \tilde{l} < l^*$.

From inspection of these two cases it should become evident that no matter what firms report

[15] To be fair, Kwerel (1977) is never explicit about allocating the licenses for free (neither about auctioning them off); he instead talks about licenses that are issued and the existence of a market for licenses. Besides, there is no much we can infer from the firm's cost minimizing problems laid out in pags. 596 and 597 because of the price-taking behavior (i.e., any grandfathered allocation can be omitted from the minimization problem since it is a lump-sum transfer with no effect on the firm's abatement decision). But had the author been aware that the scheme would work differently depending on whether the licenses are grandfathered or auctioned off, he would have explicitly said so. I thank the author for clarifying to me that the paper was in fact developed having in mind a free allocation of licenses, which is also the interpretation that one grasps from reading other papers of the time (e.g., Roberts and Spence, 1976; Dasgupta et al., 1980).

[16] There is an implicit assumption in Kwerel's work that each firms is able to move the curve.

to the regulator the market equilibrium price of licenses is given by $p = \max\{P(x), D'(x)\}$. Hence, the minimum possible equilibrium price for licenses is $p^{\min} = P(x^*) = D'(x^*)$, which is obtained when all firms report their true demand curves. Based on this observation, Kwerel closes his proof by arguing that since each firm's compliance cost is an increasing function of p, no firm has incentives to move the aggregate demand curve from its actual value, whatever it is, when it believes that all the other firms are telling the truth.[17]

Unfortunately, Kwerel's logic has a fundamental flaw in that it neglects the revenues accruing to firms from selling licenses back to the government. After taking these extra revenues into account, it is not difficult to show that firms have no incentives to tell the truth but to over-report their demand curves. To see this, simply go back to Figure 1a and compare the total compliance costs when reporting the true aggregate demand curve $P(x)$ with those when reporting $\hat{P}(x)$. Total costs under truth-telling correspond exclusively to abatement costs (area $x^0 l^* E$) while under over-reporting to abatement costs (area $x^0 \hat{x} A$) minus license sales to the government (area $\hat{l}\hat{x}AB$). Clearly, the regulation has turned out to be quite a profitable business for firms, and more so the higher the degree of over-reporting. More generally, it can be established

Proposition 1 *The unique (Nash-equilibrium) outcome in Kwerel's scheme is for firms to over-report their demand curves as to ensure the maximum possible number of licenses and subsidy level.*

Proof. See Appendix A. ■

Despite its fundamental flaw, Kwerel's scheme has an element that I also use in constructing the auction mechanism that I present next. Under this new scheme firms are also communicated in advance that the information they report to the regulator will be used in a form similar to expression (2); although with some key differences.

3 The auction mechanism

The best way to appreciate the workings of the auction mechanism is in the context of a single firm. I will then show how the scheme extends to the case of multiple (non-cooperative) firms.

[17] Kwerel also mentions the existence of multiple "offsetting-lies" Nash equilibria in which two or more firms send false reports that, on aggregate, add to the true demand curve $P(x)$. Without knowing the actual $P(x)$, however, it is hard to see how firms could coordinate in one of these "offsetting-lies" Nash equilibria.

3.1 Single firm

Consider a single firm with demand curve $P(x) \equiv -C'(x)$. The auction scheme operates as follows:

(i) firm is informed in advance about the auction rules (including the way the paybacks are computed)

(ii) firm is asked to bid a non-increasing inverse demand schedule $\hat{P}(x)$ (or equivalently, a non-increasing demand schedule $\hat{X}(p)$)

(iii) the auctioneer/regulator clears the auction (i.e., determines p and l) according to

$$p = \hat{P}(l) = D'(l) \tag{3}$$

(iv) the firm receives l licenses and pay p for each license

(v) the firm gets a fraction $\alpha(l)$ of the auction revenues back (i.e., payback is $\alpha(l)pl$)

With the aid of Figures 1a and 1b we can see that is not socially optimal for the regulator to set the fraction $\alpha(l)$ equal to either 1 or 0. If the regulator keeps no revenue for himself (i.e., $\alpha(l) = 1$), the firm has incentives to over-report by as much as to postpone any abatement effort. As shown in Figure 1a, by submitting $\hat{P}(x)$ instead of $P(x)$, the firm is able to reduce its compliance cost from area $x^0 l^* E$ to area $x^0 \hat{x} A$ minus $\hat{l}\hat{x}AB$ and abate far less than the first-best level. Conversely, if the regulator keeps all the auction revenues for himself (i.e., $\alpha(l) = 0$), the firm has incentives to under-report to some optimal extent. As shown in Figure 1b, by submitting $\tilde{P}(x)$ instead of $P(x)$, the firm is able to reduce its compliance cost from area $x^0 0 p^* E$ to area $x^0 0 \tilde{s} FBE$. The firm's optimal under-reporting in this case balances at the margin the gains from getting a lower price for licenses with the losses from higher abatement. In terms of Figure 1b, the optimal under-reporting is found by maximizing the difference between area $F\tilde{s}p^*A$ and area ABE.

To find the function $\alpha(l)$ that just induces the firm to submit its true demand curve and hence allows the regulator to implement the first-best, we proceed by backward induction. Given some function $\alpha(l)$, the firm's problem is to find the demand schedule $\hat{P}(x)$ that solves

$$\min C(l) + pl - \alpha(l)pl \tag{4}$$

subject to (3).

Using the auction clearing equation (3) we can replace p by $D'(l)$ in (4), and since there is a one-to-one correspondence between a demand schedule $\hat{P}(x)$ and the number of licenses l, at least in the range of prices the firm expects the auction to clear, the firm's first order condition is given by

$$C'(l) + D'(l) + D''(l)l - \alpha'(l)D'(l)l - \alpha(l)(D''(l)l + D'(l)) = 0 \tag{5}$$

Note that if we plug $\alpha(l) = 1$ into (5), the latter reduces to $P(l = x^0) = 0$, which is the no-regulation solution. If, on the other hand, we plug $\alpha(l) = 0$ into (5), we obtain the monopsonist solution $l^{\#}$, where $0 < l^{\#} < l^*$.

Anticipating (5), the regulator's problem is to find the function $\alpha(l)$ that induces the firm to deliver the first-best allocation, i.e., $C'(l^*) + D'(l^*) = 0$ (or $P(l^*) = D'(l^*)$). Such function solves the differential equation

$$\alpha'(l) + \alpha(l)\left(\frac{D''(l)l + D'(l)}{D'(l)l}\right) = \frac{D''(l)}{D'(l)} \tag{6}$$

The function $\alpha(l)$ that results from solving (6) is the function the regulator informs the firm along with the other auction rules in step (i) above.

Interestingly enough, the function $\alpha(l)$ only depends on the shape of the marginal damage function and not on the demand schedule submitted by the firm. If, for example, $D'(x) = hx$, where h is some constant parameter, the differential equation (6) reduces to $\alpha'(l) + 2\alpha(l)/l = 1/l$, which has a unique solution: $\alpha(l) = 1/2$.[18] The reason why $\alpha(l)$ does not depend on the firm's bid —a property that is not unique to the single-firm case but extends also to the multi-firm case— is simply because the firm's cost minimization problem pays explicit attention to the auction clearing equation (3).

Furthermore, if the function $\alpha(l)$ is such that the firm is delivering the first-best l^* it must be the case that the firm is solving the regulator's problem up front. In other words, the last two terms of eq.(4) must add to $C(l)$, which leads to (recall that $p = D'(l)$)

Proposition 2 *The payback function is given by*

$$\alpha(l) = 1 - \frac{D(l)}{D'(l)l}$$

[18] Strictly speaking the general solution is $\alpha(l) = 1/2 + K/l^2$ where K is an integration constant. But since $\alpha(l)$ cannot go unbounded as l approaches zero, K must be set to zero.

Proof. *Let*

$$g(l) = \exp \int \frac{D''(l)l + D'(l)}{D'(l)l} dl = \exp \int d\ln(D'(l)l) = D'(l)l$$

Then the solution to the differential equation (6) for $0 \leq l < \infty$ *is given by*

$$\alpha(l) = \frac{1}{g(l)} \left(K + \int g(l) \frac{D''(l)}{D'(l)} dl \right) = \frac{1}{D'(l)l} \left(K + \int D''(l) l dl \right)$$

where K *is an integration constant. Integrating by parts, we obtain*

$$\alpha(l) = \frac{1}{D'(l)l} (K + D'(l)l - D(l))$$

and setting the constant term K *to zero*[19] *finishes the proof.* ■

Since $D'(l)$ is a non-decreasing function of l it is clear that $0 \leq \alpha(l) \leq 1$, so the final price paid by the firm for each license (i.e., $(1 - \alpha)p$), is at most equal to marginal damage $D'(l)$. Plugging back the function $\alpha(l)$ of Proposition 2 into the firm's objective function (4) it is immediately seen that the new auction scheme has indeed converted the firm's problem into the regulator's by making the firm bear the full cost of the pollution damages.

The idea of requiring the firm to pay $D(l)$ is certainly not new. For example, the DHM mechanism for the case of a single firm reduces precisely to the regulator informing the firm that it faces a tax function $T(x) = D(x)$, where x is the firm's observed pollution level (note also that because there is only one firm the regulator does not need the firm to report any cost/demand information to him). The auction mechanism implements the same result but in a different way. Here the regulator asks the firm to submit a demand schedule that it is then used to compute the optimal number of licenses and the price to be charged for each license. In other words, the pollution cap is set ex-ante, i.e, before pollution occurs. It that sense the auction scheme fully decouples the regulatory design from the enforcement/monitoring activity like in any other quantity-based regulation —whether it is based on standards or transferable licenses— which today is by far the more prevalent type of regulation for protecting the commons.[20] Other differences will become evident as we consider more than one firm.

[19] *Note that* $\lim_{l \to 0} D(l)/D'(l)l = 1$*, so for* $\alpha(l)$ *not to explode as* l *approaches zero,* K *must be set to zero (see also footnote 19).*

[20] It is easy, however, to transform the DHM into a quantity-based mechanism by simply allocating licenses at the reported first-best levels and asking firms to pay for them according to their individual tax schedules.

But before moving to the multiple-firm case it is worth mentioning that if our single firm knows the function $D(x)$, it does not need to truthfully bid its entire demand schedule but only the portion relevant to the auction clearing. It could for instance submit the perfectly inelastic demand schedule $\hat{X}(p) = l^*$.[21] More importantly, although I have developed the auction mechanism as if the firm knew the damage function $D(x)$, it should be clear by now that the firm does not actually need to know $D(x)$, and hence $\alpha(l)$, for the auction mechanism to work in providing incentives for truthful revelation. We only require the firm to believe that it is facing a regulator committed to implement the first-best for whatever function $D(x)$ he has in mind. And if the firm does indeed know little about $D(x)$, it will truthfully bid its (almost) entire demand schedule to make sure that for any possible function $D(x)$ chosen by the regulator it will get the first-best level of licenses.

3.2 Multiple firms: Non-cooperative equilibrium

Let us now consider the incomplete information game of $n \geq 2$ firms with $P_i(x_i)$ being firm i's private information. The auction mechanism extends as follows

(i) firms are informed in advance about the auction rules (including the way payback functions are computed)

(ii) firm i $(= 1, ..., n)$ is asked to bid a non-increasing inverse demand schedule $\hat{P}_i(x_i)$ (or, equivalently, a non-increasing demand schedule $\hat{X}_i(p)$)

(iii) the regulator computes the residual marginal damage function (i.e., residual supply function) for each firm i using other firms' demand schedules, that is

$$D_i'(x_i) \equiv D'(x) - \hat{P}_{-i}(x_{-i}) \tag{7}$$

where $\hat{P}_{-i}(x_{-i})$ is the demand schedule from aggregating all other bids besides firm i's and $x_i \equiv x - x_{-i} \geq 0$ (note that since $x_i \geq 0$, $D_i'(x_i)$ is only defined at and above the point at which $D'(x) = \hat{P}_{-i}(x_{-i})$; see Figure 2)[22]

(iv) entering $D_i'(l_i)$ and $D_i(l_i) \equiv \int_0^{l_i} D_i'(l)dl$ into Proposition 2 the regulator computes a function $\alpha_i(l_i)$ for each firm

[21] Here the relevant range has reduced to a single point, that is $p^* = D'(l^*)$.

[22] In terms of demand and supply functions, firm i's residual supply function would be $X_i^{rs}(p) = X^s(p) - \hat{X}_{-i}^d(p)$, where $D_i'(X_i^{rs}(p)) = p$, $D'(X^s(p)) = p$ and $\hat{P}_{-i}(\hat{X}_{-i}^d(p)) = p$.

(v) the regulator clears the auction by determining a price p_i and number licenses l_i for each bidder i according to

$$p_i = \hat{P}_i(l_i) = D_i'(l_i)$$

(vi) firm i purchases l_i licenses at a price p_i each, and soon after gets a payback of $\alpha_i(l_i)p_i l_i$

Proposition 3 *In the auction mechanism described above it is optimal for each firm i to bid its true demand curve $P_i(x_i)$ regardless of what other firms bid.*

Proof. It follows immediately from the construction of $\alpha_i(l_i)$ in Proposition 2. ■

Truth-telling is a (weakly) dominant strategy for firms, so there is no need for them to form beliefs about other firms' types and/or actions. The immediate implication of Proposition 2 is that this auction scheme implements the first-best with each firm facing the same price at the margin (i.e., $p_i = p^*$ for all i) and getting exactly the first-best allocation of licenses (i.e., $l_i = x_i^*$), which eliminates (efficiency) reasons for trading licenses after the auction.[23] These efficiency properties can be readily seen in Figure 2: if $\hat{P}_i(x_i) = P_i(x_i)$ and $\hat{P}_{-i}(x_{-i}) = P_{-i}(x_{-i})$, then $l_i = x_i^*$, $l = x^*$ and $\hat{p} = p^*$.

Although in principle the auctioneer (i.e., regulator) goes bidder after bidder determining individual prices p_i, these prices are all the same regardless of how truthful firms are (in terms of Figure 2: $p_1 = ... = p_n = \hat{p}$). But because of the existence of paybacks, one may still question the uniform-price format of the auction. Unless firms have identical demand curves, final prices, $(1 - \alpha_i)p$, will differ across firms (in and off equilibrium). But in equilibrium they will differ in a most equitable way. From Proposition 2 we know that $\alpha_i(l_i)$ is such that

Proposition 4 *In the auction mechanism described above firm i's total payment in equilibrium is exactly equal to the residual (or additional) damage caused by its pollution, i.e., $(1 - \alpha_i(l_i^*))p^* l_i^* = D_i(l_i^*)$, where $D_i(l_i) \equiv \int_0^{l_i} D_i'(z)dz$.*

This is a very important result for the purposes of practical implementation. It is hard to imagine of a more equitable and efficient way to allocate the total cost of the pollution externality. Proposition 4 also allows us to see the close connection that exists between the uniform-price auction mechanism and the (also private-value) discriminatory auctions of Vickrey (1961) and Ausubel (2004). In both of these discriminatory auctions, where the total number of licenses

[23] Note also that there is no need for formally modelling the after-auction market since no firm can profit from it by misreporting at the auction.

is known ex-ante, bidder i pays a different price for each item she gets and possibly different prices than j's, but the total payment of i is equal to the area under the her residual supply function.[24] Despite the quite different structures, it should not be surprising that the total payment in the uniform-price auction mechanism follows the exact payment rule of Vickrey (1961) and Ausubel (2004). The reason is that all these three auctions are constructed upon the same principle. In Vickrey and Ausubel auctions each bidder is asked to pay an amount equal to the "pecuniary externality" she exerts on the other competing bidders while here each bidder is asked to pay an amount equal to the (residual) externality she imposes on society.

Based on the payment rule of Proposition 4, the auction mechanism can be certainly viewed as a VCG mechanism in that it makes each firm to pay only for the externality she exerts on society. It is, however, structurally different than the more conventional VCG mechanism as implemented by DHM. Letting θ_i identify firm i's true type, in DHM firm i faces a tax schedule $T_i(x_i)$ equal to

$$T_i(x_i) = D(x_i + \sum_{j \neq i} x_j^*(\hat{\theta}_i, \hat{\boldsymbol{\theta}}_{-i})) + \sum_{j \neq i} C_j(x_j^*(\hat{\theta}_i, \hat{\boldsymbol{\theta}}_{-i}), \hat{\theta}_j) - A_i(\hat{\boldsymbol{\theta}}_{-i}) \tag{8}$$

where $\hat{\theta}_i$ is firm i's report to the regulator, $\hat{\boldsymbol{\theta}}_{-i}$ is the vector of firms $j \neq i$'s reports, $x_j^*(\hat{\theta}_i, \hat{\boldsymbol{\theta}}_{-i})$ is firm j's first-best pollution level as dictated by the reports of all firms, and A_i is a constant term independent of firm i's report. Although DHM is not specific about the A_i's,[25] we know that for DHM to be a VCG mechanism the constant term A_i is made equal to the efficient social cost had firm i not existed, that is

$$A_i(\hat{\boldsymbol{\theta}}_{-i}) = D(x_{-i}^{**}) + \sum_{j \neq i} C_j(x_j^{**}(\hat{\boldsymbol{\theta}}_{-i}), \hat{\theta}_j)$$

where $x_{-i}^{**} \equiv \sum_{j \neq i} x_j^{**}(\hat{\boldsymbol{\theta}}_{-i})$ and $x_j^{**}(\hat{\boldsymbol{\theta}}_{-i})$ is firm j's first-best pollution level in the absence of firm i.

Firm i's total payments under the auction mechanism and under the DHM mechanism will be exactly the same in equilibrium (but not off-equilibrium),[26] although computed in structurally different ways. This can be easily shown with the aid of Figure 2. Firm i's

[24] If l is the volume of licenses to be auctioned off, firm i's residual supply function is $l - \hat{X}_{-i}(p)$, where $\hat{X}_{-i}(p)$ is the sum of the bids submitted by all other bidders.

[25] One does not need to be specific in order to implement the first-best if there are no restrictions on firms' payments.

[26] But different off equilibrium they are not the same.

total payment under the auction mechanism is the shaded area to the left while the payment under the DHM mechanism is the shaded area to the right (recall that $l_i = \hat{x} - \hat{x}_{-i}$). Despite the similarities, the structural differences will prove critical in explaining why in some other, yet important, circumstances the auction mechanism continues to perform equally well (i.e., delivering the efficient allocation) while the DHM mechanism does not.

3.3 Properties of the mechanism

The main property of the auction mechanism is that it delivers the first-best for any number of firms and irrespective of how well-informed firms are. There are four additional properties (related to the evolution of paybacks, perfectly elastic/inelastic supply curve, off-equilibrium behavior, and budget balancing) that may be important for purposes of practical implementation. First, while the payment rule of Proposition 4 allows for final prices to differ across firms, it does it in an equitable way, and more so if as we increase the number of firms, the α_i's rapidly fall towards zero. Although the α_i's are never exactly equal to zero, when there is a large number of firms, firm i has virtually no effect on the equilibrium price, so $D_i'(x_i^*) \approx D_i'(0)$. And by replacing $D_i''(x_i) = 0$ into (6), the differential equation's unique solution happens to be $\alpha_i(l_i) = 0$; hence, the auction scheme has converged to the Pigouvian principle for taxing externalities (Pigou, 1920).

To illustrate how rapidly the auction's payment rule approaches Pigou's, let us consider a numerical example. Suppose there are n symmetric firms with linear demand curves. The aggregate demand curve is $P(x) = \bar{p}(1 - x/x^0)$, where $\bar{p}$ is the choke price and x^0 is the unregulated level of pollution. The damage function is $D'(x) = hx$. Solving as a function of the number of firms, we obtain

$$\alpha(n) = \frac{1}{2}\frac{\bar{p}}{(n-1)hx^0 + n\bar{p}} \tag{9}$$

If we further let the slopes of the aggregate demand and marginal damage curves be the same (i.e., $h = \bar{p}/x^0$), then eq. (9) reduces to $\alpha(n) = 1/(4n-2)$. The rebate for three firms is 10%, for ten firms is 2.6%, and for 100 firms is less than 0.3%.

Second, it is well known that if $D'(x)$ is constant a first-best policy is to charge a Pigouvian tax equal to D'. The auction mechanism is equivalent to this tax policy in that paybacks are exactly equal to zero but it still has the (practical) advantage that the socially efficient amount of pollution is revealed ex-ante, i.e., before pollution occurs.

More interestingly, if the supply curve is totally inelastic, say, at $x = \bar{x}$, whether because there is a genuine threshold at $\bar{x}$ or, more likely, because the regulator has no control over x, the auction mechanism still retains its truth-telling properties (note that each firm faces a elastic residual supply function).[27] This latter has two important implications. On the one hand, it makes the auction mechanism readily comparable to the (fixed-supply) private-value auctions of Vickrey and Ausubel. In fact, the three auctions yield the same outcome in terms of revenues and allocations; although they are implemented in quite different ways. On the other hand, it introduces a sharp distinction between the auction mechanism and the DHM mechanism. DHM fails to yield the efficient outcome because each individual firm is no longer pivotal, i.e., its report does not affect the aggregate supply. DHM can lead, for example, to the inefficient Nash equilibrium in which each of the (no necessarily symmetric) n firms submit a null report (or that there are no clean up costs up to $\bar{x}/n$) and then emit $\bar{x}/n$. This equilibrium strategy reduces payments to zero and clearly no firm wants to deviate from it (note that above $\bar{x}$ firms face an infinitely large pollution fee).

Third, unlike in the single-firm case, in the context of multiple firms is in each firm's best interest to bid truthfully not only that portion of the demand curve around its first-best allocation x_i^* but rather a large portion of its demand curve. Even if each firm knows $D(x)$, it can no longer anticipate $l_i^* = x_i^*$ with precision because it does not know other firms' demand curves $P_{-i}(x_{-i})$ (it may be even unaware of the number of firms being regulated). To be more precise, a firm will only find it strictly optimal to bid truthfully the portion of its demand curve that is relevant for the auction clearing. Thus, if firms assign zero probability to the event that the clearing price will fall below some value, say $\underline{p}$, firms can just bid an almost perfectly elastic (or inelastic for that matter) demand curve for $p \leq \underline{p}$.[28] While this off-equilibrium behavior has no consequences on the clearing price, and hence, on implementing the first-best allocation, it does have an effect on firms' total payments. But because demand schedules are non-increasing in p, a firm's total payment will never be greater (and generally smaller) than the Pigouvian payment.

Finally, the auction mechanism is, like any other VCG mechanism, a non-budget-balanced mechanism both in and off equilibrium (unless $\hat{X}_i(p) = 0$ for all i). Although there is no

[27] There is, however, one exception: the "single" firm will no longer submit its true demand curve but $\hat{P}(x) = 0$. But since x is fixed and there is only one firm, this has no allocative implications.

[28] Going back to the fixed-supply case discussed in the preceding paragraph, the existence of $\underline{p} > 0$ could eventually break the equivalence with the Vickrey and Ausubel auctions in terms of revenues (but not in terms of allocations).

efficiency reasons for balancing the budget there may be political economy reasons for doing so (Tietenberg, 2003).[29] As first pointed out by Groves and Ledyard (1977), if there are at least three agents it should be possible to balance the budget for a variety of mechanisms. The basic idea is to distribute the surplus or deficit generated by each agent ($D_i(l_i)$ in our case) among the other agents in some lump-sum manner as to avoid any incentive effects. Behind this idea lies an implicit "separability" condition that in our case would allow us to either make the payment $D_i(l_i)$ independent of some firm j's report (i.e., $\hat{P}_j(x_j)$), as in Duggan and Roberts (2002), or to perfectly disentangle the contribution of each firm $j \neq i$'s report to firm i's payment, as in Varian (1994). By construction, the auction mechanism lacks of such separability; hence, there is no way in which the mechanism can be modified to achieve perfect budget-balancing while retaining its first-best properties.[30]

There exists, however, an approximate solution. Building upon the idea of Groves and Ledyard (1977), let denote by $D_j^{-i}(l_j^{-i})$ the total payment that firm j would have hypothetically faced under the same auction mechanism but in the absence of firm i's demand schedule, where l_j^{-i} is the corresponding number of licenses allocated to j. The regulator can thus fashion a lump-sum compensation rebate R_i for firm i using these influence-free hypothetical payments. For example,

$$R_i = \frac{1}{n-1} \sum_{j \neq i} D_j^{-i}(l_j^{-i})$$

where $n \geq 2$.

This solution assures a perfectly balanced budget (i.e., $\sum_{i=1}^n R_i = \sum_{i=1}^n D_i(l_i)$) only in the limiting case of a large number of firms; otherwise, $\sum R_i$ could be smaller, greater or equal than $\sum D_i(l_i)$. The ratio $\rho \equiv \sum R_i / \sum D_i(l_i)$ will ultimately depend on the number of firms and shape of the demands and marginal damage curves. For linear curves, for example, it can be shown that for three (symmetric) firms ρ can be anywhere between 0.60 and 1.50, for ten firms anywhere between 0.90 and 1.11 and for 100 firms anywhere between 0.99 and 1.01. Thus, a regulator that cannot run a deficit, i.e., constrained to return at most $\sum D_i(l_i)$ to firms, can inform in advance that it will return only some fixed fraction of the total $\sum D_i(l_i)$ (in the case

[29] We may also want to take into consideration the general equilibrium reasons of Bovenberg and Goulder (1996) for not balancing the budget.

[30] The reason why the complete-information mechanisms of Duggan and Roberts (2002) and Varian (1994) can balance the budget is because they are based on discrete announcements by firms. In the former firms announce quantities while in the latter they announce prices. In the auction mechanism firms announce a continuum of quantity-price pairs.

of 10 firms this fraction could be 90%).

4 Collusion

The workings of bidding rings or auction cartels have received a fair amount of theoretical and empirical attention in the auction literature (e.g., McAfee and McMillan, 1992; Kemplerer, 2004). In this section I discuss how the auction scheme proposed in this paper performs under collusive behavior, if sustainable, and whether it requires of any adjustment in order to preserve its first-best properties.

The way to implement a collusive agreement in our multi-unit auction is not very different from the description of McAfee and McMillan (1992) for a single-unit auction except for some elements that I will explain below. Cartel firms need to both coordinate on their bidding schedules and agree on the procedure for sharing the cartel profits. But because cartel members do not know each other's demand curves, in implementing the collusive agreement the cartel organization must itself overcome an adverse-selection problem: it must induce its members to truthfully reveal their private information. In other words, the cartel itself faces an internal mechanism design problem. I will first present the optimal (i.e., maximal profits) collusive agreement and then an internal mechanism the cartel can use to implement such outcome.

4.1 The optimal collusive agreement

Imagine for a moment that there is a relatively large number of independent production plants. In the non-cooperative equilibrium each plant i operates at its first-best level x_i^* and receives virtually no payback. Imagine now that all those plants belong to a single holding company subject to the same auction scheme. The holding company is clearly better off because is not only operating at the same level (x_i^* at plant $i = 1, ..n$) but also receiving a strictly positive payback (1/2 of the auction revenues if $D'(x) = hx$). A good collusive agreement would then be for plants to coordinate as if they were acting as a single entity. It can be established more generally, however, that

Proposition 5 *The optimal collusive agreement for a cartel of $m \leq n$ firms is to submit only one serious bid with the true aggregate demand curve of the cartel, say $P_c(x_c)$. One cartel member submits the serious bid while all the other members submit empty demand schedules. The optimal collusive agreement delivers the first-best allocation.*

Proof. See Appendix B. ■

It is remarkable that the auction mechanism, without any sort of adjustment, can deliver the first-best even when firms are colluding. There are three interrelated reasons for that. First, paybacks for any given level of licenses are largest when the cartel faces the total supply function instead of a series of residual supply functions. Second, clean-up costs for any given level of licenses are lowest when they can be split cost-effectively across all firms in the cartel. Third, the single-firm analysis has already shown that the level of licenses that minimizes overall costs (clean-up costs and payments) is the first-best level.

These same three reasons also help explaining why cartel profits are increasing with the number of cartel members. Unlike in the single-unit auctions of McAfee and McMillan (1992), where the addition of a "low-valuation" member only contributes to dissipate cartel rents, in our multi-unit auction the most profitable cartel is an all-inclusive cartel (i.e., $m = n$). Existing members may eventually restrict additional participation insofar as it helps to prevent detection by antitrust authorities.

There are two additional observations. First, from looking at expression (8) it is not difficult to see that a collusive agreement under the DHM mechanism would depart from the first-best allocation. Because of the positive constant term $A_i(\hat{\boldsymbol{\theta}}_{-i})$ firms would coordinate on some over-reporting of their types (i.e., demand curves).[31] If we set the constant term to zero, however, collusive and non-cooperative behavior would be no different but payments would be so large that the (now Groves) mechanism would be of little practical value.

Second, there are many other, yet less profitable, collusive agreements under the auction mechanism. For example, it is possible to show that firms would benefit if they can coordinate on some demand reduction. A sub-optimal agreement like this would certainly move us away from the first best. But there is no good reason for firms to ever coordinate on a sub-optimal agreement if they can enforce the optimal agreement, as I argue in the next section.

4.2 Implementing the collusive agreement

The arguments made thus far have assumed that the cartel submits only one serious bid and this is the aggregate demand of cartel members. To do this, however, the cartel has to induce

[31] Note that the optimal (all-inclusive) collusive agreement under DHM would be defined by the n-report vector $\hat{\boldsymbol{\theta}}$ that solves

$$\min n \cdot \left(\sum_{i=1}^{n} C_i(x_i(\hat{\boldsymbol{\theta}}, \theta_i)) + D(x) \right) - \sum_{i=1}^{n} A_i(\hat{\boldsymbol{\theta}}_{-i})$$

where $x = \sum x_i$.

its members to truthfully reveal their individual demand curves. In addition, collusive profits have to be shared among the members in a way that the members would wish to participate in the cartel and not to deviate at the auction.

McAfee and McMillan (1992) provides a complete description of the two commonly observed forms of cartel organization: weak cartels, cartels whose members are unable to make transfer payments among themselves, and strong cartels, cartels whose members can both make transfer payments and exclude new entrants. While either type of organization may eventually arise in the single-unit auctions of McAfee and McMillan (1992), for the multi-unit auctions studied in this paper we must restrict attention to strong cartels.

It is hard to imagine how a weak cartel could achieve any degree of cooperation in the multi-unit incomplete-information environment of the auction mechanism. In the absence of after-auction transfers, a week cartel must conform itself with a sub-optimal collusive agreement in which each member (tacitly or not) agree on shading their bids to some extent. Since a firm's dominant strategy at the auction is to bid truthfully, a necessary condition for the sustainability of such an agreement is that cartel members can detect deviations at the auction. But unlike in a single-object auction, where "weak-cartel" members coordinate on bidding the seller's reserve price, detecting deviations in the multi-unit auction mechanism requires cartel members to have information on each and every member's demand curve. In the absence of transfers, it is impossible for the cartel to devise an internal (incentive-compatible) mechanism that can provide cartel members with such information prior to the auction.

Paying exclusive attention to strong cartels may seem restrictive but I would argue that in our context it is the most natural type of organization of the two because licenses are readily transferable either through bilateral contracts or in the (after-auction) spot market.[32] Thus, the transferable nature of licenses provides the cartel with the needed flexibility to structure transfer payments without necessarily facilitating its detection.[33]

Let us then consider a potential strong cartel of $m \leq n$ members indexed as $j = 1, ..., m$. In implementing the optimal collusive agreement of Proposition 5, the cartel organization must solve two intertwined problems. First, it must put in place an internal scheme that induces firms to truthfully reveal their individual demand curves to the cartel organization, which, as in

[32] Although in equilibrium we should not observe any trade of licenses after the auction, in reality there will be always some trading activity as firms accomodate to factor price variation between auctions.

[33] Although collusion is less of a concern here (Proposition 5), antitrust authorities may still dislike it in that it can facilitate the spread-out of collusion through multi-market contacts.

McAfee and McMillan (1992), we will call the *cartel mechanism*. Second, the cartel must ensure obedience to the cartel mechanism, that is, it must be equipped with the ability to detect and credibly punish deviators.

Suppose for now that cartel has solved the second problem (I will come back to this shortly) and focus on the first problem. In so doing, consider the following cartel mechanism. Prior to the official auction, cartel members first agree on how to divide cartel profits by determining shares $\lambda_j > 0$, where $\sum_{j=1}^{m} \lambda_j = 1$. One plausible criteria can be historic use of the resource, that is $\lambda_j \approx x_j^0 / \sum_{k=1}^{m} x_k^0$ (in McAffee and McMillan (1992) $\lambda_j = 1/m$). Cartel profits are defined as the difference between the payment associated to the optimal agreement and the sum of the non-cooperative payments that cartel members would have faced at the auction for the same demand schedules reported to the cartel mechanism. After λ_j's are set, cartel members report their demand schedules $\tilde{P}_j(x_j)$ to the cartel mechanism. Let $\tilde{P}_c(x_c)$ denote the aggregate demand curve reported by cartel members. The cartel mechanism selects an arbitrary member to be the *serious bidder*, say bidder 1, which bids $\hat{P}_1(x_1) = \tilde{P}_c(x_1)$. Remaining cartel members bid $\hat{X}_j(p) = 0$ for all $j = 2, ..., m$.

The cartel mechanism also establishes the way licenses and payments are transferred across cartel members posterior to the auction. Let denote by l_c the number of licenses received by the serious bidder at the auction and by $D_c(l_c)$ the corresponding payment, where $D_c(x_c)$ is the residual damage function faced by the cartel. The cartel mechanism establishes that each cartel member j will receive an amount of licenses exactly equal to what he would have individually obtained at the auction for the demand curve that he reported to the mechanism. Member j's total payment for the l_j licenses will be equal to what he would have individually paid at the auction, $D_j(l_j)$, minus a fraction λ_j of the cartel profits. Payment $D_j(l_j)$ is computed as in the auction mechanism, that is, using the aggregate demand curve reported by the remaining cartel members, $\tilde{P}_{-j}(x_{-j})$. More precisely, $D_j(x_j) \equiv \int_0^{x_j} D_j'(z)dz$, where $D_j'(x_j) \equiv D_c'(x_c) - \tilde{P}_{-j}(x_{-j})$ for all j.[34]

Proposition 6 *Assuming that the cartel members can agree on the λ_j's, it is a dominant*

[34] In an all-inclusive cartel, i.e., $m = n$, $D_c(x_c)$ is known prior to the auction ($D_c(x_c) = D(x)$). In a partial cartel, i.e., $m < n$, $D_c(x_c)$ is only learned posterior to the auction. Although the immediate auction results (i.e., l_c, $D_c(l_c)$ and $D_c'(l_c) = \hat{p}$) will provide the cartel with insufficient information to fully reconstruct the curve $D_c(x_c)$, I see no reason why the serious bidder, or any bidder for that matter, cannot request information on $D_c(x_c)$ from the auctioneer. Furthermore, and as double-check practice, it may be a good idea to provide bidders with their own residual supply functions (in case of a mistake in the clearing of the auction, there will be always some bidder that will find it profitable to report so). If bidders are not entitled to request such information, the cartel can alternatively use a first-order (linear) approximation for $D_c'(x_c)$.

strategy Nash equilibrium for them to report truthfully to the cartel mechanism, i.e., $\tilde{P}_j(x_j) = P_j(x_j)$ *for all* $j = 1, ..., m \leq n$.

Proof. Given members' obedience to the cartel mechanism for whatever demand schedules they choose to report, cartel member $j = 1, ..., m \leq n$ will report the demand schedule $\tilde{P}_j(\cdot)$ that solves (recall the one-to-one correspondence between reporting $\tilde{P}_j(\cdot)$ and requesting l_j licenses)

$$\min_{l_j} C_j(l_j) + D_j(l_j) - \lambda_j \left(\sum_{k=1}^{m} D_k(l_k(l_j)) - D_c(l_c) \right)$$

where $l_k(l_j)$ is member k's license allocation as a function of j's allocation and $l_c = \sum_{k=1}^{m} l_k$. The first-order condition is

$$C_j'(l_j) + D_j'(l_j) - \lambda_j \left(\sum_{k=1}^{m} (D_k'(l_k) - D_c'(l_c)) \frac{dl_k}{dl_j} \right) = 0$$

But from the auction clearing condition we have $\hat{p} = D_c'(l_c) = D_k'(l_k)$ for all $k = 1, ..., m$, which finishes the proof. ■

Before moving onto the cartel's second implementation problem, that of obedience with the cartel mechanism, let briefly touch on three issues. First, the cartel mechanism proposed above is not the only (incentive-compatible) mechanism available to the cartel. It has the advantage, however, that by sharing the format of the auction mechanism it makes it easier for members to understand its workings.

Second, I have little to add on how firms will come to an agreement on the λ_j's other than pointing out that I see no reason for negotiations to fall apart because it is all about splitting a pie of an ex-ante unknown size. In other words, the bargaining process for setting the λ_j's does not involved the type of information asymmetries that are usually associated to negotiation failure (e.g., Wiggins and Libecap, 1985).

Third, the cartel mechanism (whether the one proposed here or any other) must be executed in its entirety prior to the official auction, except for the actual transfer of licenses (and payments) across cartel members. Unlike in a single-object auction where the cartel mechanism (e.g., "knockout" auction) that decides which of the cartel members will keep the object can be conducted either before or after the official auction (McAfee and McMillan, 1992), in our multi-unit environment this is simply not possible for both information and incentive reasons. On the one hand, the serious bidder must be informed of $\tilde{P}_c(x_c)$ before coming to the official auction. On the other hand, the use of any ex-post bidding procedure for determining how to allocate l_c and $D_c(l_c)$ across cartel members (e.g., a knockout auction with a structure similar to the auction mechanism but for an inelastic supply) will necessary distort member's (ex-ante) incentives in communicating with the cartel mechanism.

Let us now look at the cartel's second problem that of enforcement with the cartel mechanism. In the absence of an outside enforcer, the enforcement needed to ensure that all cartel members comply with the cartel mechanism may come from a grim trigger strategy in the infinitely repeated auction or in some infinitely repeated multimarket contact. A deviating bidder would be threatened with noncooperative profits in all future auctions (or "contacted" markets) should she disobey the cartel mechanism. Unlike in the single-object auction of McAfee and McMillan (1992), the cartel in our multi-unit context requires of no patience from its members to maintain cooperation through the official auction. The optimal deviation of bidder j, whether it is the serious bidder ($j = 1$) or any of the non-serious bidders ($j = 2, ..., m$), is to bid an empty demand schedule to the cartel mechanism (so as to reduce its residual supply curve at the official auction to the maximum extent possible) and then bid its true demand curve $P_j(x_j)$ at the official auction. But this deviation leaves the deviating bidder strictly worse off in an amount exactly equal to its share λ_j of the otherwise cartel profits.[35]

5 Extensions

I will show next how the auction mechanism easily accommodates to other commons problems such as those involving non-uniformly mixed pollutants (i.e., firms' pollutants are not perfect substitutes in the damage function) and private externalities.

5.1 Imperfect substitutability of licenses

Consider the case in which social damage is no longer a function of total pollution but, as in Dasgupta et al. (1980) and Duggan and Roberts (2002), of the firms' pollution vector. There are $n \geq 2$ firms with (privately known) demand and cost functions $P_i(x_i)$ and $C_i(x_i)$,

[35] Like in McAfee and McMillan (1992), there is still the possibility that posterior to the official auction the serious bidder renege the cartel mechanism and run away with the object (l_c licenses). Since no other license holder is willing to sell for less than the auction clearing price $\hat{p}$ at the open market that may develop after the auction, the serious bidder could appropriate all cartel profits by simply posting a price $\hat{p}$ for the l_c licenses. One may argue, however, that since firms are free to sign bilateral contracts for transfering licenses, the run-away problem could be managed by the writing of bilateral contracts contigent on auctions results (i.e., $\hat{p}$, l_c, average paybacks, etc.) and to be called upon should the serious bidder renege the cartel mechanism.

respectively, where $i = 1, ..., n$. Pollution damages are denoted by the differentiable and convex function $D(\mathbf{x})$, where $\mathbf{x} = (x_1, ..., x_n)$ is the pollution vector. Without perfect substitutability of pollutants, and hence of licenses, we do not want to insist on a uniform-price auction because it may be socially optimal that each firm faces a different price for licenses at the margin. For the same reason the regulator wants to make licenses to be firm-specific as to prevent any trading of licenses after the auction.

Let $\mathbf{x}^* = (x_1^*, ..., x_n^*)$ be the first-best allocation vector (which is interior and unique); then $\mathbf{x}^*$ satisfies the first-order conditions

$$-C_i'(x_i^*) \equiv P_i(x_i^*) = \frac{\partial D(\mathbf{x}^*)}{\partial x_i} \tag{10}$$

for all $i = 1, ..., n$.

For the auction mechanism to deliver the first-best allocation, the payment rule of Proposition 4 implies that firm i's residual damage curve as a function of x_i must be

$$D_i(x_i) \equiv \int_0^{x_i} \frac{\partial D(x_1^*(y), ..., x_{i-1}^*(y), y, x_{i+1}^*(y), ..., x_n^*(y))}{\partial y} dy \tag{11}$$

where $x_j^*(y)$ is the first-best allocation to firm $j \neq i$ when y licenses are allocated to firm i. It is easy to see that if firm i's total payment is given by (11), the solution to firm i's problem, i.e., find the number of licenses l_i that minimizes $C_i(l_i) + D_i(l_i)$, satisfies the first-order condition (10).

To compute firm i's residual damage curve the auctioneer/regulator will use the bids from the remaining $n-1$ firms to solve a system of $n-1$ first-order conditions

$$\hat{P}_j(x_j) = \frac{\partial D(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n)}{\partial x_j} \tag{12}$$

for $j = 1, .., n$ and $j \neq i$. Solving the system of equations (12) leads to $n-1$ functions of the form $x_j^*(x_i)$ for all $j \neq i$. These functions are then entered into $D(x_i, \mathbf{x}_{-i}^*(x_i))$ to finally obtain firm i's residual damage function (11).

Given definition (11), the auction works exactly as before. The regulator clears the auction by determining a price p_i and number licenses l_i for each bidder i according to

$$p_i = \hat{P}_i(l_i) = D_i'(l_i) \equiv \frac{\partial D(l_i, \mathbf{x}_{-i}(l_i))}{\partial l_i}$$

and soon after gives i a rebate of $\alpha_i(l_i)p_il_i$, where (Proposition 2) $\alpha_i(l_i) = 1 - D_i(l_i)/l_iD_i'(l_i)$ with $0 \leq \alpha_i(l_i) \leq 1$.

Before closing this section, it is worth mentioning that collusion, whether implemented by a weak cartel or a strong cartel, is not possible in this imperfect-substitution world. Weak cartels cannot detect deviations for the same reasons laid out before and strong cartels cannot make transfers because licenses are by definition not transferable.

5.2 Private externalities

Consider now the case in which firms not only impose costs on society but also impose costs (or benefits) on other firms. Fishing in open sea and grazing goats in public land are two "commons" examples but the analysis here applies more generally to any private externality problem. There are $n \geq 2$ firms. Firm i's production is denoted by x_i and its (differentiable) profit function by $\Pi_i(x_1, ..., x_i, ..., x_n)$ where $\partial\Pi_i(\cdot)/\partial x_i > 0$ and $\partial^2\Pi_i(\cdot)/\partial x_i^2 < 0$. For concreteness, let us focus on the case of pure private negative externalities, i.e., $\partial\Pi_i(\cdot)/\partial x_j < 0$ for all $j \neq i$ (it is relatively straightforward to generalize the scheme to the presence of both social and private externalities).

Let $\mathbf{x}^* = (x_1^*, ..., x_n^*)$ be the first-best or joint-profit-maximizing allocation vector (which is interior and unique); then $\mathbf{x}^*$ satisfies the first-order conditions

$$\frac{\partial\Pi_i(\mathbf{x}^*)}{\partial x_i} + \sum_{j\neq i}\frac{\partial\Pi_j(\mathbf{x}^*)}{\partial x_i} = 0 \text{ for all } i = 1, ..., n \tag{13}$$

Had the regulator known the size of the externality exerted by each firm at the first-best level, i.e., $\sum_{j\neq i}\partial\Pi_j(\mathbf{x}^*)/\partial x_i$, he would have just charged a Pigouvian tax equal to $\tau_i^* \equiv \sum_{j\neq i}\tau_{ij}^* \equiv \sum_{j\neq i}\partial\Pi_j(\mathbf{x}^*)/\partial x_i$ to firm i's output, where τ_{ij}^* measures the (first-best) marginal damage that i imposes on j. But since regulators generally do not have such information, Varian (1994) has provided them with the following simple multistage mechanism. First, all firms simultaneously announce the magnitude of the vector of Pigouvian taxes to be faced by each firm (including itself). Then the regulator uses firms' announcements to compute transfers from/to firms as a function of the production vector $\mathbf{x}$. Finally, output $\mathbf{x}$ is decided. Varian shows that transfers can be structured in a way that the (unique) subgame-perfect equilibrium of this game is that each firm reports the first-best Pigouvian tax vector and that $\mathbf{x} = \mathbf{x}^*$. As expressed by Varian (1994) in the concluding paragraph of the paper, however, the main problem with this multistage mechanism is that it requires complete information by the firms.

The auction mechanism proposed in this paper does not require firms to possess any such information. It assumes that $\Pi_i(\mathbf{x})$ is firm i's private information. In the specific context of private externalities, the auction mechanism operates as follows. Firms are asked to submit (non-increasing) demand schedules $\hat{P}_i(x_1, ..., x_i, ..., x_n)$ for $i = 1, ..., n$.[36] The regulator/auctioneer uses that information to recover "reported" profit functions

$$\hat{\Pi}_i(x_i, \mathbf{x}_{-i}) = \int_0^{x_i} \hat{P}_i(x_1, ..., x_{i-1}, y, x_{i+1}, ..., x_n)dy$$

which then he uses to compute the residual damage functions as dictated by Proposition 4

$$D_i(x_i) \equiv \sum_{j \neq i} \hat{\Pi}_j(x_1^{**}, ..., x_{i-1}^{**}, 0, x_{i+1}^{**}, ..., x_n^{**}) - \sum_{j \neq i} \hat{\Pi}_j(x_1^*(x_i), ..., x_i, ..., x_n^*(x_i)) \tag{14}$$

for all $j = 1, ..., n$ and $j \neq i$.

The first sum in (14) is the "reported" first-best profits of all firms but i in the absence of firm i and the second sum is the first-best profits of all firms but i when firm i is allowed to produced $x_i > 0$. As in the basic model, expression (14) tracks down the (first-best) profit losses that the presence of firm i, as measured by x_i, causes on all other agents. Again, it is not difficult to see that if firm i's total payment is given by (14), the solution to firm i's problem, i.e., find the number of licenses $l_i = x_i$ that maximizes $\Pi_i(l_i, \mathbf{x}_{-i}) - D_i(l_i)$, satisfies the first-order condition (13). The computation of functions $x_j^*(x_i)$ for all $j \neq i$ is as in the previous section: the auctioneer will use the bids from the $j \neq i$ firms and solve the $n - 1$ first-order conditions as a function of x_i.

A simple example may help here (to make it more interesting I will allow for corner solutions). Consider two firms 1 and 2 (or i and j) with profit functions $\Pi_i(x_i, x_j) = (\theta_i - x_i - x_j)x_i \geq 0$, where the value of θ_i is firm i's private information. For $\theta_i > \theta_j$ the socially optimal solution is

$$x_i^* = \frac{\theta_i}{2} \text{ and } x_j^* = 0$$

(and for $\theta_i = \theta_j = \theta$ the efficient solution is $x_i^* + x_j^* = \theta/2$).

In the absence of regulation firms will produce beyond this joint-profit maximizing level

[36] Note that in many commons problems these demand schedules will reduce to $\hat{P}_i(x_i, x_{-i})$, where $x_{-i} \equiv \sum_{j \neq i} x_j$.

(we may have a total collapse of the resource in that $\theta_i < x_i + x_j$ for $i = 1, 2$).[37] The auction mechanism corrects the externalities as follows. Firms are asked to report their demand curves or types to the auctioneer, say $\hat{\theta}_1$ and $\hat{\theta}_2$, knowing beforehand that the regulator/auctioneer will use this information to determine allocations

$$l_i = \begin{cases} \hat{\theta}_i/2 & \text{if } \hat{\theta}_i > \hat{\theta}_j \\ 0 & \text{if } \hat{\theta}_i < \hat{\theta}_j \end{cases}$$

and total payments

$$D_i = \begin{cases} \hat{\theta}_j^2/4 & \text{if } \hat{\theta}_i > \hat{\theta}_j \\ 0 & \text{if } \hat{\theta}_i < \hat{\theta}_j \end{cases}$$

for $i = 1, 2$. If $\hat{\theta}_i = \hat{\theta}_j = \hat{\theta}$ the regulator flips a coin for deciding who gets the $\hat{\theta}/2$ licenses for a total payment of $\hat{\theta}^2/4$ (we assume that the winning firm opts to produce despite making zero profits).

By letting firm i face a payment equal to firm j's (first-best) profits had firm i not existed (i.e., $\Pi_j = \theta_j^2/4$), it is in firm i's best interest to submit a truthful bid (i.e., $\hat{\theta}_i = \theta_i$) regardless of what firm j bids. This is not surprising since the auction mechanism has collapsed to a single-object second-price auction.[38]

6 Final remarks

I have developed a simple auction mechanism for the optimal regulation of a commons resource when the regulator lacks information about the characteristics of the firms that are being regulated. The mechanism is developed under the additional assumption that firms know nothing about other firms' characteristics. The mechanism is not only simple in that it is based on commonly used instruments (i.e., transferable licenses) but also remarkably effective in that it delivers the first-best allocation regardless of the number of firms and on whether they are acting noncooperatively or collusively. The mechanism yields the efficient allocation even when

[37] Suppose that is common information that θ_i's are i.i.d. over the support $[\underline{\theta}_i, \bar{\theta}_i]$, the Bayesian Nash equilibrium is

$$x_i = \frac{\theta_i}{2} - \frac{2E[\theta_i] - E[\theta_j]}{6}$$

for $i = 1, 2$ and where $E[\cdot]$ is the expected value operator.

[38] Since there are multiple socially optimal solutions for the case in which $\theta_i = \theta_j$, one may be inclined to replace the coin-flipping allocation by a more equitable allocation such as the following: if $\hat{\theta}_i = \hat{\theta}_j = \hat{\theta}$, then $l_i = l_j = \hat{\theta}/4$ and $D_i = D_j = \hat{\theta}^2/8$. This allocation rule still yields a truth-telling Nash equilibrium but no longer in dominant strategies. If, for example, firm i believes $\hat{\theta}_j > \theta_i > \theta_j$, it may be optimal for i to bid $\hat{\theta}_i = \hat{\theta}_j > \theta_i$.

the aggregate supply of licenses is fixed – either because there is a genuine threshold or, more likely, because the regulator has no control over the aggregate supply.

There are two aspects not treated in the paper that may prove relevant in some commons problems. First, there may be cases in which emissions (or resource use, more generally) cannot be perfectly monitored. Then, in addition to the adverse selection problem of not observing a firm's type (e.g., abatement costs) the regulator must now overcome the moral hazard problem of not perfectly observing the firm's action. In a recent paper, Montero (2005) compares the performance of two instruments —grandfathered transferable licenses and performance standards— in such information environment. He finds that in some cases a standard-alone policy can welfare dominate a licenses-alone policy. In many cases, though, the optimal policy is to combine licenses and standards. It would be interesting to study how the auction mechanism extends to the case of imperfect monitoring and to ask whether and to what extent it remains (second-best) optimal to auctioning off the licenses subject to a minimum performance standard.

Second, in existing license programs the allocation horizon of licenses is not necessarily tight to the nature of the investments. If investments depreciate rather quickly, say, in a year, and the magnitude of the social externality is likely to fluctuate from year to year, it appears optimal to run a new auction every year. If, on the other hand, investments are long-lived and mostly irreversible, but the authority allocates licenses via annual auctions, the time-consistent solution would be for firms to under-invest relative to the first-best level (recall that in the presence of irreversibility ex-post marginal costs can be significantly lower than ex-ante marginal costs). Since the nature of an investment is to a large extent firm's private information, it remains to be seen whether and to what extent the auction mechanism can handle this additional adverse selection problem. The regulator/auctioneer must use firms' bids not only to allocate licenses efficiently across firms at any point in time but also across time, i.e., it must also resolve how often firms, either individually or a as group, must come to the auction. Perhaps it is optimal to let firms store current licenses for future use, as permitted in most existing license markets, and make future auction allocations contingent upon the amount of licenses stored.

Appendix A: Proof of Proposition 1

Let $1 > \lambda_i > 0$ be the fraction of licenses allocated to firm i $(= 1, ..., n)$, so firm i receives an initial allocation of $\lambda_i l$, where l is the total number of licenses and $\sum_{i=1}^{n} \lambda_i = 1$ (as commonly observed in practice, λ_i could be proportional to historic emissions, that is $\lambda_i \approx x_i^0/x^0$). The license market is assumed perfectly competitive (i.e., n large). Let $\bar{s}$ be the maximum value the subsidy can take, which by construction fixes the maximum number of licenses to $\bar{l}$, where $D'(\bar{l}) = \bar{s}$. The regulator sets $\bar{s}$ sufficiently high that is always above the first-best level p^* for any possible realization of $P(x)$; otherwise, there is no point in using the scheme (in Kwerel (1977) s is unbounded but in reality we cannot let it go to infinity).

We will demonstrate that the pair $(\bar{s}, \bar{l})$ is the unique Nash-equilibrium outcome of Kwerel's scheme. From the arguments in the text we do not need to consider the case of under-reporting. Thus, for a reported aggregate demand curve $\hat{P}(x) \geq P(x)$, the subsidy level is s and the market price of licenses is $p = s$; hence, firm i's total compliance costs as a function of s becomes

$$TC_i(s) = C_i(X_i(s)) + s \cdot (X_i(s) - \lambda_i l(s)) \tag{A1}$$

where $l(s) = D'^{-1}(s)$. The first term of (A1) is abatement cost and the second term is the net cost of purchasing licenses (which is negative when the firm is a net seller of licenses). Consider first the case in which the regulator sets $\bar{s}$ "close" to infinity. It is not difficult to see that no firm has incentives to move the outcome away from the pair $(\bar{s}, \bar{l})$. Since $X_i(s = \bar{s}) = 0$ for all $i = 1, ..n$ (firms either shut down operations or install backstop zero-emission technologies), all firms become net sellers to the government and their total costs, $TC_i(s = \bar{s}) = C_i(0) - \lambda_i \bar{s}\bar{l} < 0$, reach the minimum (recall that $l'(s) > 0$). Consequently, all firms will submit infinitely large demand curves $\hat{P}_i(x_i)$ so as to ensure that $s = \bar{s}$.

Consider now the case in which $\bar{s}$ is "only large" in the sense that for some or all firms $X_i(\bar{s}) > 0$. We have that

$$\frac{dTC_i(s)}{ds} = C_i' \cdot \frac{dX_i(s)}{ds} + X_i(s) - \lambda_i l(s) + s \cdot \left(\frac{dX_i(s)}{ds} - \lambda_i \frac{1}{D''(l(s))} \right) \tag{A2}$$

But $C_i' = -s$, so evaluating (A2) at $s = \bar{s}$ reduces to

$$\left. \frac{dTC_i(s)}{ds} \right|_{s=\bar{s}} = X_i(\bar{s}) - \lambda_i l(\bar{s}) - \frac{\lambda_i \bar{s}}{D''(\bar{l})}$$

Since with over-reporting $\sum_{i=1}^{n} X_i(\bar{s}) < l(\bar{s})$, there will be many firms for which $dTC_i(\bar{s})/ds < 0$ (note that if firms are symmetric it is immediate that $dTC_i(\bar{s})/ds < 0$ for all firms; if firms are heterogeneous it may be still be the case that $dTC_i(\bar{s})/ds < 0$ for all firms). These firms have no incentives to move the outcome away from $s = \bar{s}$, so they will over-report by as much as possible as to ensure that $s = \bar{s}$. And since firms for which $dTC_j(\bar{s})/ds > 0$ cannot fully counterbalance these over-reporting incentives because at best they can report $\hat{X}_j(p) = 0$, the equilibrium outcome will necessarily be the pair $(\bar{s}, \bar{l})$.

Appendix B: Proof of Proposition 5

Without loss of generality let us parametrize firm i's inverse demand function as $P_i(x_i) \equiv P(x_i, \theta_i)$ where θ_i is an index of type and $\partial P/\partial\theta > 0$ (similarly, the parametrization for the demand function is $X_i(p) \equiv X(p, \theta_i)$ where $\partial X/\partial\theta > 0$). There is a one-to-one correspondence between a reported demand schedule $\hat{P}_i$ and a reported type $\hat{\theta}_i$. Consider for the moment only two firms, i and j. The firms' reports $\hat{\theta}_i$ and $\hat{\theta}_j$ conducive to the most profitable collusive agreement are found by solving

$$\min_{\hat{\theta}_i,\, \hat{\theta}_j} C(x_i,\theta_i)+C(x_j,\theta_j)+[1-\alpha_i(l_i(\hat{\theta}_i,\hat{\theta}_j))]\hat{p}(\hat{\theta}_i,\hat{\theta}_j)l_i(\hat{\theta}_i,\hat{\theta}_j)+[1-\alpha_j(l_j(\hat{\theta}_i,\hat{\theta}_j))]\hat{p}(\hat{\theta}_i,\hat{\theta}_j)l_j(\hat{\theta}_i,\hat{\theta}_j)) \tag{B1}$$

subject to

$$x_i + x_j = l_i(\hat{\theta}_i, \hat{\theta}_j) + l_j(\hat{\theta}_i, \hat{\theta}_j) = l(\hat{\theta}_i, \hat{\theta}_j) \tag{B2}$$

where $\hat{p}(\hat{\theta}_i, \hat{\theta}_j) \equiv \hat{p}$ is the auction clearing price as a function of firms' bids and $l_i(\hat{\theta}_i, \hat{\theta}_j) \equiv l_i$ is the number of licenses allocated to firm i. In what follows I will omit $\hat{\theta}_i$ and $\hat{\theta}_j$ unless it would otherwise cause confusion. From Proposition 2 we know that

$$[1 - \alpha_i(l_i)]\hat{p}l_i = D_i(l_i) = \hat{p}l_i - \int_{\hat{p}_j(\hat{\theta}_j)}^{\hat{p}} [X^s(p) - X(p, \hat{\theta}_j)]dp \tag{B3}$$

where $X^s(p)$ is the social supply function, i.e., $D'(x)$, so $X^s(p) - X(p, \hat{\theta}_j)$ is the residual supply faced by firm i, i.e., $D_i'(x_i, \hat{\theta}_j)$; and $\hat{p}_j(\hat{\theta}_j) \leq \hat{p}$ is the (hypothetical) clearing price in the absence of firm i's bid (in terms of Figure 2, $\hat{p}_j(\hat{\theta}_j)$ corresponds to $\hat{p}_{-i}$). The first-order condition for (B1) is (allowing for corner solutions)

$$\frac{\partial C(x_i,\theta_i)}{\partial x_i}\frac{dx_i}{dl}\frac{\partial l}{\partial \hat{\theta}_i} + \frac{\partial C(x_j,\theta_j)}{\partial x_i}\frac{dx_j}{dl}\frac{\partial l}{\partial \hat{\theta}_i} + \frac{\partial D_i(l_i)}{\partial \hat{\theta}_i} + \frac{\partial D_j(l_j)}{\partial \hat{\theta}_i} \geq 0 \tag{B4}$$

Recall that $-\partial C(x_i, \theta_i)/\partial x_i = P(x_i, \theta_i)$. To obtain an expression for dx_i/dl use (B2) and note that collusion optima requires

$$P(x_i, \theta_i) = P(x_j = l - x_i, \theta_j) = P(l, \theta_{i+j}) \tag{B5}$$

where $P(l, \theta_{i+j})$ is the true aggregate demand function. Totally differentiating (B5) with respect to l and rearranging leads to

$$\frac{dx_i}{dl} = \frac{P'_j}{P'_i + P'_j} \tag{B6}$$

where $P'_i \equiv \partial P(x_i, \theta_i)/\partial x_i$. To obtain expressions for $\partial D_i(l_i)/\partial\hat{\theta}_i$ and $\partial D_j(l_j)/\partial\hat{\theta}_i$ (or $\partial D_i(l_i)/\partial\hat{\theta}_j$), on the other hand, note that from (B3) we have

$$\frac{\partial D_i(l_i)}{\partial\hat{\theta}_i} = \frac{\partial\hat{p}}{\partial\hat{\theta}_i} l_i + \hat{p}\frac{\partial l_i}{\hat{\theta}_i} - \frac{\partial\hat{p}}{\partial\hat{\theta}_i}[X^s(\hat{p}) - X(\hat{p}, \hat{\theta}_j)] \tag{B7}$$

But $X^s(\hat{p}) - X(\hat{p}, \hat{\theta}_j) = l_i$, so (B7) reduces to

$$\frac{\partial D_i(l_i)}{\partial\hat{\theta}_i} = \hat{p}\frac{\partial l_i}{\hat{\theta}_i} \tag{B8}$$

Similarly,

$$\frac{\partial D_i(l_i)}{\partial\hat{\theta}_j} = \frac{\partial\hat{p}}{\partial\hat{\theta}_j} l_i + \hat{p}\frac{\partial l_i}{\hat{\theta}_j} - \frac{\partial\hat{p}}{\partial\hat{\theta}_j} l_i + \frac{\partial\hat{p}_j(\hat{\theta}_j)}{\partial\hat{\theta}_j} \cdot 0 + \int_{\hat{p}_j}^{\hat{p}} \frac{\partial X(p, \hat{\theta}_j)}{\partial\hat{\theta}_j} dp$$

Rearranging and inverting i by j leads to

$$\frac{\partial D_j(l_j)}{\partial\hat{\theta}_i} = \hat{p}\frac{\partial l_j}{\partial\hat{\theta}_i} + \int_{\hat{p}_i}^{\hat{p}} \frac{\partial X(p, \hat{\theta}_i)}{\partial\hat{\theta}_i} dp \tag{B9}$$

Note that since $\hat{p}_i \leq \hat{p}$ the last term of (B9) is non-negative. Plugging (B6), (B8) and (B9) into (B4), using (B5) and rearranging, the two FOCs become

$$\frac{\partial l(\hat{\theta}_i, \hat{\theta}_j)}{\partial\hat{\theta}_i}[-P(l(\hat{\theta}_i, \hat{\theta}_j), \theta_{i+j}) + \hat{p}(\hat{\theta}_i, \hat{\theta}_j)] + \int_{\hat{p}_i(\hat{\theta}_i)}^{\hat{p}(\hat{\theta}_i, \hat{\theta}_j)} \frac{\partial X(p, \hat{\theta}_i)}{\partial\hat{\theta}_i} dp \geq 0 \quad \text{for } i \text{ and } j \tag{15}$$

By inspection of (15) one arrives at two possible solutions. One solution is for i to report $\hat{\theta}_i = \theta_{i+j}$ (i.e., the true aggregate demand curve) and for j to report the corner $\hat{\theta}_j = \emptyset$ (i.e., $\hat{X}_j = 0$ for all p). If so, $P(l, \theta_{i+j}) = \hat{p} = \hat{p}_i > \hat{p}_j = D'(0)$, and hence, the FOC for i equals to zero and the FOC for j is strictly positive. The second solution is just the inverse. Both solutions are equally optimal (for the firms) and, more importantly, they implement the first-best in that firms find it in their best collusive interest to submit the aggregate true curve. Extending the proof to the case of more than two firms and to the possibility of partial collusion (i.e., collusion among a subset of firms) is straightforward.

References

[1] Ausubel, L. (2004), An efficient ascending-bid auction for multiple objects, *American Economic Review* 94, 1452-1475.

[2] Bovenberg, L. and L. Goulder (1996), Optimal environmental taxation in the presence of other taxes: general equilibrium analysis, *American Economic Review* 86, 985-1000.

[3] Cramton, P. and S. Kerr (1999), Tradeable carbon permit auctions: How and why to auction not grandfather, Working paper, Maryland University.

[4] Dasgupta, P., P. Hammond and E. Maskin (1980), On imperfect information and optimal pollution control, *Review of Economic Studies* 47, 857-860.

[5] Duggan, J. and J. Roberts (2002), Implementing the efficient allocation of pollution, *American Economic Review* 92, 1070-1078.

[6] Groves, T. and J. Ledyard (1977), Optimal allocations of public goods: A solution to the free rider problem, *Econometrica* 45, 783-809.

[7] Joskow, P., R. Schmalensee and E.M Bailey (1998), The market for sulfur dioxide emissions, *American Economic Review* 88, 669-685.

[8] Kemplerer, P. (2004), *Auctions: Theory and Practice*, Princeton University Press, Princeton, NJ.

[9] Kremer, I. and N. Nyborg (2004), Divisible-good auctions: The role of allocation rules, *RAND Journal of Economics* 35, 147-159.

[10] Kwerel, E. (1977), To tell the truth: Imperfect information and optimal pollution control, *Review of Economic Studies* 44, 595-601.

[11] McAdams, D. (2005), Adjustable supply in uniform price auctions: The value of non-commitment, Working paper, MIT.

[12] McAfee, P. and J. McMillan (1992), Bidding rings, *American Economic Review* 82, 579-599.

[13] Milgrom, P. (2004), *Putting Auction Theory to Work*, Cambridge University Press, New York, NY.

[14] Montero, J.-P. (2005), Pollution markets with imperfectly observed emissions, *RAND Journal of Economics* 36, 645-660.

[15] Pigou, A.C. (1920). *The Economics of Welfare*, Macmillan, London.

[16] Roberts, M. and M. Spence (1976), Effluent charges and licenses under uncertainty, *Journal of Public Economics* 5, 193-208.

[17] Spulber, D. (1988), Optimal environmental regulation under asymmetric information, *Journal of Public Economics* 35, 163-181.

[18] Tietenberg, T. (2003), The tradable-permits approach to protecting the commons: Lessons for climate change, *Oxford Review of Economics Policy* 19, 400-419.

[19] Varian, H. (1994), A solution to the problem of externalities when agents are well-informed, *American Economic Review* 84, 1278-1293.

[20] Vickrey (1961), Counterspeculation, auctions and competitive sealed tenders, *Journal of Finance* 16, 8-37.

[21] Weitzman, M. (1974), Prices vs quantities, *Review of Economic Studies* 41, 477-491.

[22] Wiggins, S. and G. Libecap (1985), Oil field unitization: Contractual failure in the presence of imperfect information, *American Economic Review* 75, 368-385.

[23] Wilson, R. (1979), Auctions of shares, *Quarterly Journal of Economics* 93, 675-689.

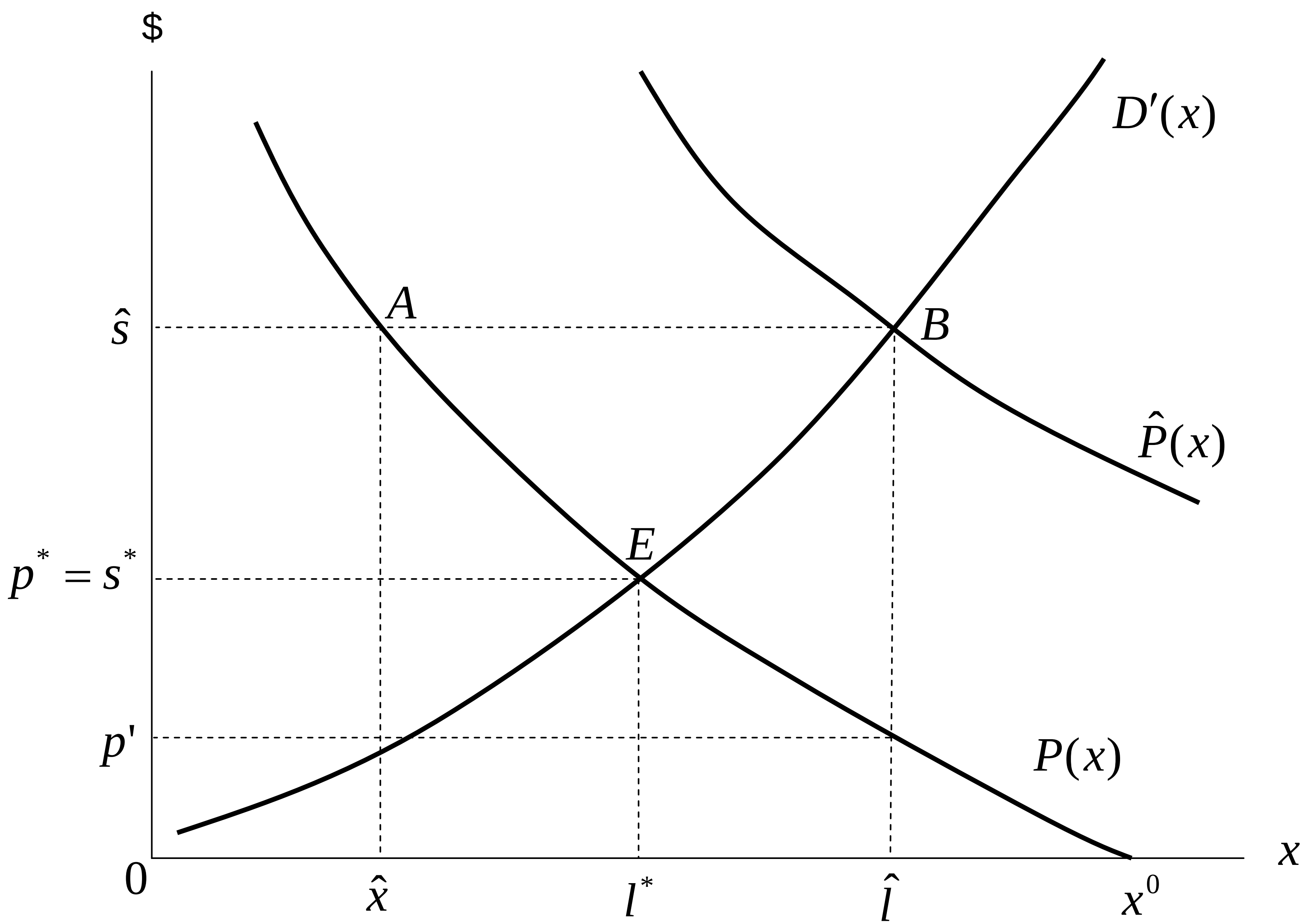

Figure 1a: Incentives to over-report

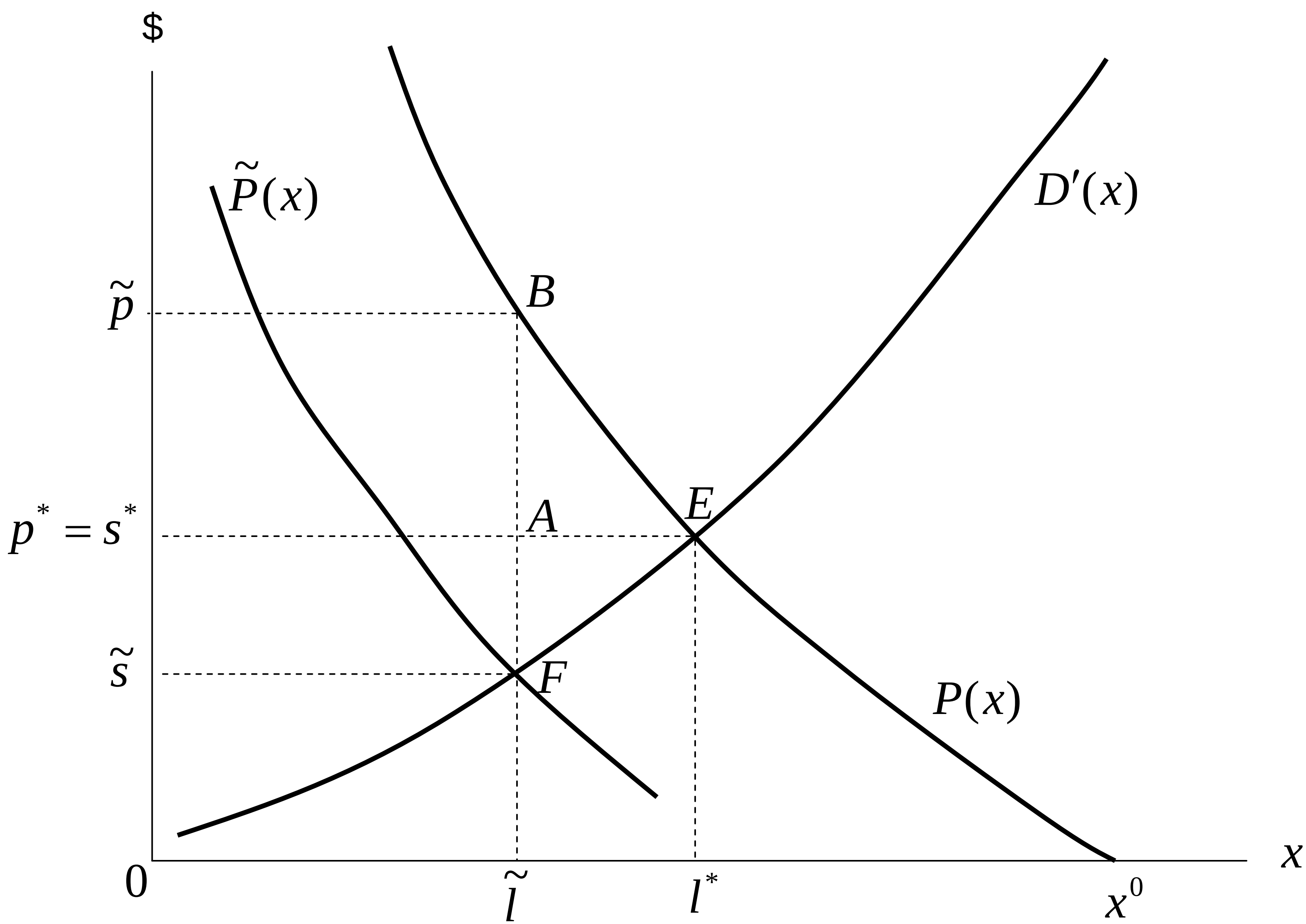

Figure 1b: Incentives to under-report

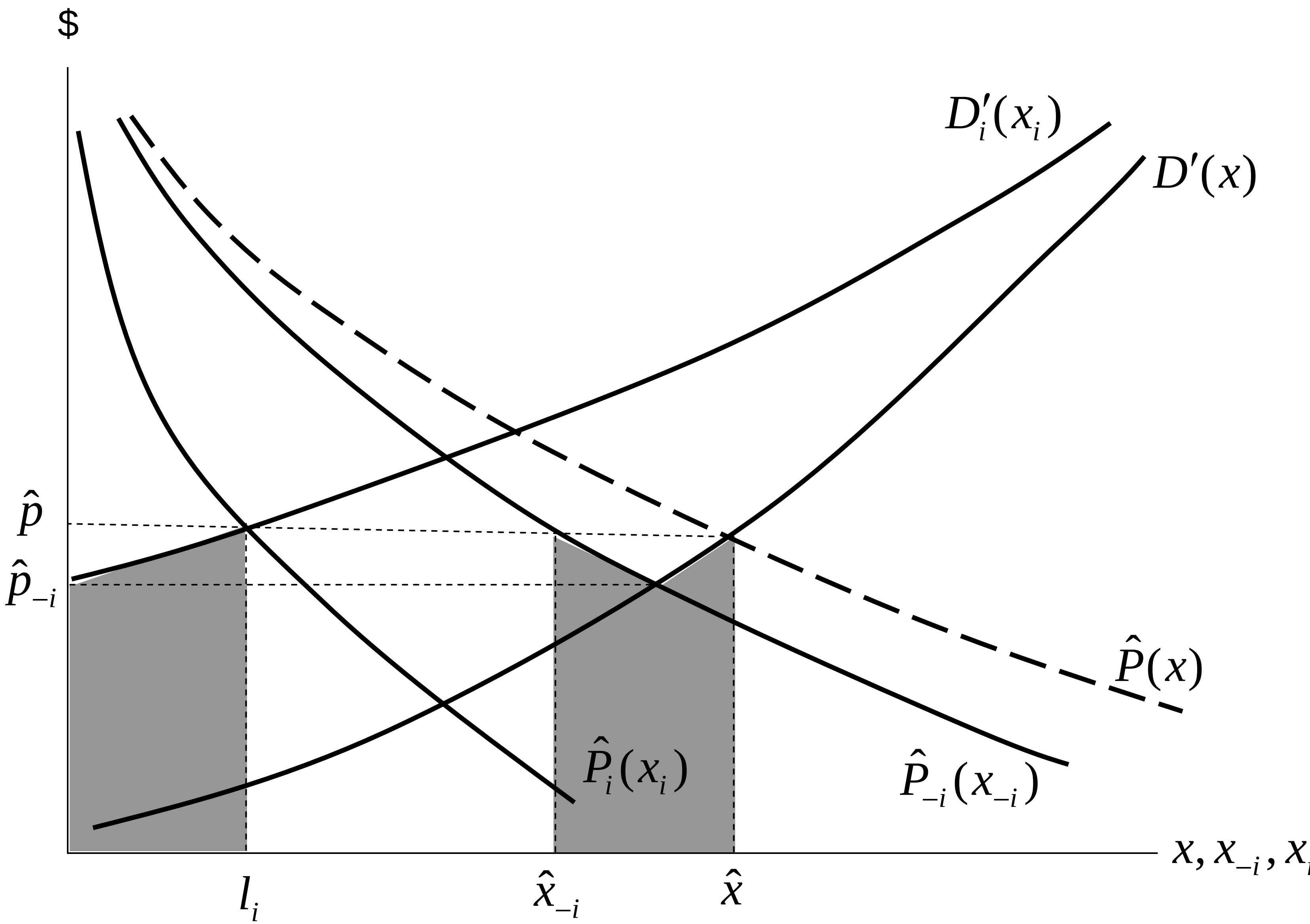

Figure 2: Residual supply (i.e., marginal damage) function