Prices vs. Quantities with Incomplete Enforcement

Juan-Pablo Montero*
Catholic University of Chile, and
Massachusetts Institute of Technology

This draft: July 7, 1999

Abstract

This paper extends Weitzman's (1974) "Prices vs. Quantities" to allow for incomplete enforcement. Whether the regulator uses prices (e.g., taxes) or quantities (e.g., tradeable quotas), a first-best design is always inefficient in the presence of incomplete enforcement. A second-best design that incorporates incomplete enforcement, and where cost and benefit curves are known with certainty, can be implemented equally well with either instrument. If benefit and cost curves are uncertain, however, a quantity instrument performs better than a price instrument. In fact, if the slopes of the marginal cost and marginal benefit curves are equal, quantities are always preferred over prices. Results are consistent to alternative enforcement policies.

Keywords: regulation, environment, incomplete enforcement, uncertainty

JEL Classification: L5, Q2, D8

---

* Assistant Professor of Industrial Economics at the Department of Industrial Engineering of the Catholic University of Chile (CUC), and Research Associate at the MIT Center for Energy and Environmental Policy Research (MIT-CEEPR). Send any correspondence to: Catholic University of Chile, Casilla 306, Correo 22, Santiago, Chile. Phone: 56-2-686-4613, fax: 56-2-552-1608; and e-mail: jpmonter@ing.puc.cl. I want to thank Denny Ellerman, Larry Karp, David Reiner and participants in a seminar at CUC for comments and discussions, and to MIT-CEEPR and Fondecyt-Chile (Grant # 1990100) for financial support.

Prices vs. Quantities with Incomplete Enforcement

1. Introduction

In a seminal paper, Weitzman (1974) explored the question of whether it is better to implement control over an economic variable using a price instrument (e.g., taxes) or a quantity instrument (e.g., tradeable quotas). He demonstrated that uncertainty concerning the marginal cost of control affects the choice between these two regulatory instruments. When the planner's objective is to maximize the expected net benefits of controlling that economic variable, a quantity instrument ought to be preferred if the marginal benefit curve is steeper than the marginal cost curve, otherwise, a price instrument ought to be preferred. The basic reasoning is that as the marginal benefit curve becomes relatively steeper than the marginal cost curve, having certainty on the amount of control becomes more important than having certainty on the cost of control.[1]

Weitzman's motivating example, that of instrument choice for pollution control, continues to be of great concern in today's environmental policy-making (Fisher et al., 1996). In this context, we find a growing literature that has extended Weitzman's basic framework to compare Pigouvian taxes and systems of tradable quotas in different directions (e.g., Roberts and Spence, 1976; Baldursson and von der Fehr, 1998; Hoel and Karp, 1998).[2] No work, however, has extended Weitzman's framework to the case in which the environmental agency (or any other agency) has limited monitoring and enforcement capabilities to ensure full compliance with the regulation.

Imperfect monitoring and incomplete enforcement have proved to be central in understanding environmental policy in practice (Russell, 1990),[3] and consequently, it is important to ask whether they have any policy implications in the choice of regulatory instruments. Early theoretical works used simple enforcement policies to study the implications of incomplete enforcement on instrument performance and design for the case of taxes and pollution standards (Harford, 1978; Viscusi and Zeckhauser, 1979) and tradeable quotas (Malik, 1990).

[1] For a complete discussion on Weitzman's results see Baumol and Oates (1988).

[2] Extensions in more general contexts are in Yohe (1978) and Finkelshtain and Kislev (1997).

[3] Particularly in less developed countries where compliance rates are substantially lower than in industrialized countries.

Empirical observation of compliance higher than anything predicted by these simple models motivated the development of richer models for the case of pollution-standards regulation. Harrington (1988) modeled incomplete enforcement as a dynamic repeated game between the agency and the firm. Firms that are detected to be in violation today are subject to more frequent inspections and higher fines tomorrow. In a recent paper, Livernois and McKenna (1999) offered a different explanation by adding self-reporting requirements and enforcement power to their model. Firms were required to monitor their own pollution and report their compliance status to the environmental agency, which had the enforcement power to bring into compliance any firm eventually found to be submitting a false report. Both models produce compliance in cases in which the expected penalty for noncompliance is insufficient to prevent violations in the earlier models. Unfortunately, no such models have been developed for the case of taxes and tradeable quotas.

In this paper I extend Weitzman's work to allow for incomplete enforcement. I frame the discussion within the context of environmental policy, but there is nothing specific in the model to indicate that the results do not apply to other contexts. I retain Weitzman's one-period framework and additive uncertainty, and model incomplete enforcement incorporating the concepts of self-reporting and enforcement power from Livernois and McKenna (1999). Since the focus is on instrument choice, I abstract from considerations concerning optimal enforcement policy and assume an exogenous enforcement structure, which can be considered optimal.[4]

Results indicate that a first-best design that neglects incomplete enforcement is always inefficient. A second-best design that incorporates incomplete enforcement, and where cost and benefit curves are known with certainty, can be implemented equally well with either instrument. If benefit and cost curves are uncertain, however, a quantity instrument performs better than a price instrument.

The rationale for the latter result is that with incomplete enforcement the effective (or observed) amount of control under a quantity instrument is not longer fixed. Instead, it adapts to the possible shocks affecting the marginal cost curve. Indeed, if the marginal costs curve turns out to be higher than expected, some firms would choose not to comply with the regulation, and

[4] It shall become clear that there is no reason to believe that enforcement policies should differ among instruments.

consequently, the effective amount and cost of control would be lower. Because a quantity instrument would now have this "flexibility" for the cost of control to adapt to unexpected shocks, the advantage of prices over quantities is necessarily reduced. Put differently, the marginal cost curve becomes relatively flatter under incomplete enforcement. A more intuitive explanation would be that under incomplete enforcement the regulator wants to pay closer attention to the amount of control (i.e., emissions reduction) than to the cost of control *ceteris paribus*; a task in which a quantity instrument is more effective.

The rest of the paper is organized as follows. Section 2 presents the model and explains compliance under prices (taxes) and quantities (tradeable quotas). Section 3 starts with the planning problem of a regulator maximizing social welfare (expected net benefits), and then illustrates the effect of incomplete enforcement on welfare for both taxes and tradeable quotas. Section 4 explores the second-best design under incomplete enforcement for two cases. First, for the case where marginal cost and marginal benefit curves are known with certainty, and then, for the case where these curves are subject to uncertainty. Section 5 answers the central question of instrument choice under benefit and cost uncertainty and incomplete enforcement. Section 6 explores the same question for an alternative enforcement policy. Concluding remarks are offered in Section 7.

2. The model

Consider the following one-period model. There is a continuum of polluting firms of mass 1 subject to an environmental regulation that can take the form of either taxes (price instrument) or tradeable quotas (quantity instrument). In the absence of regulation, each firm emits one unit of a uniform pollutant, which can be abated at a constant marginal cost c. Firms differ according to their control cost, c. The distribution function of c, $g(c)$, defined over the interval $[\underline{c},\overline{c}]$, is continuous and commonly known by firms and the welfare-maximizing regulator, and $G(c)$ is the cumulative distribution function. The regulator does not know the control cost of any particular firm, but can derive the aggregate abatement cost curve for the industry, $C(q)$, where $0 \leq q \leq 1$ is the aggregate quantity of emissions reduction.[5] The regulator also knows the social benefit curve from emissions reduction, $B(q)$. As usual, we assume that $B'(q) > 0$, $B''(q) \leq 0$, $C'(q) > 0$, $C''(q) \geq 0$, $B'(0) > C'(0)$, and $B'(q) < C'(q)$ for q sufficiently large.

An enforcement agency is responsible for enforcing compliance for individual firms under either regulatory regime. Firms are required to monitor their own emissions and submit a compliance status report to the enforcement agency (that for simplicity is the regulator).[6] Emissions are not easily observed by the regulator except during costly inspection visits, when they can be measured accurately. Thus, some firms may have incentive to report themselves as being under compliance when, in reality, they are not. The compliance report also includes either tax payments or details of quota transfers, which are assumed to be tracked at no cost by the regulator. As an example, someone submitting a report with one unit of pollution and no tax payment can be relatively easy identified. Similarly, a firm A submitting a report with one unit of pollution and a "false" quota transfer from firm B can be easily identified, since B would not be reporting a transfer for which does not get paid. To corroborate the truthfulness of reports received, however, the regulator must observed emissions, which is costly.

Since the regulator lacks sufficient resources to induce full compliance, in order to verify the truthfulness of reports, s/he randomly selects a fraction ϕ of firms reporting compliance to have their emissions and control devices inspected. Firms whose reports are found to be in violation are levied a fine F and brought immediately under the required compliance, indicated by their own reports.[7,8] Firms reporting noncompliance face the same treatment, so it is always in a firm's best economic interests to report compliance, even if that is not the case.[9] Thus, each firm faces a

---

[5] The aggregate cost curve is given by $C(q) = \int_{\underline{c}}^{y} c \mathrm{d}G$, where $y = G^{-1}(q)$. Note that $C'(q) = y$, $C'(0) = \underline{c}$, and $C''(q) = 1/g(y)$.

[6] Self-reporting is an increasingly common feature of enforcement, not only in the U.S., but also in some LDCs.

[7] An effective enforcement policy should not only fine the violator, but also brings the violator under compliance (Russell, 1990).

[8] In Section 6, we relax this assumption and consider an enforcement policy that permits violators to freely choose a strategy (even one that differs from the strategy set out in their compliance report,) to come under compliance.

[9] Noncompliance and truth-telling could be a feasible strategy under other circumstances (see Livernois and McKenna, 1999).

probability ϕ of being inspected.[10] We now describe optimal compliance for each regulatory regime (i.e., taxes and tradable quotas) when enforcement is incomplete.[11]

2.1 Compliance with taxes (prices)

Given a Pigouvian tax τ, inspection probability ϕ, fine F, and marginal abatement cost c, each firm seeks to minimize its expected total compliance costs. Let us first consider the case where $c < \tau$. Such a firm will never consider paying the tax as part of its compliance strategy. It will choose to reduce one unit of emissions at cost c and submit a truthful compliance report if

$$c < \phi(F + c) + (1 - \phi)0 \tag{1}$$

where $F + c$ would be the cost incurred if the firm is found submitting a false report. Eq. (1) gives the following "cut-off point" for a truthful compliance report when $c < \tau$

$$\tilde{c} = \frac{\phi}{1 - \phi} F\ . \tag{2}$$

If $\underline{c} \le c < \tilde{c}$, the firm reduces pollution and submits a truthfully compliance report. Conversely, if $\tilde{c} < c < \tau$, the non-compliant firm does not reduce pollution and submits a false report indicating that it is reducing pollution.

Let us now consider the case where $c \ge \tau$. In this case, a firm will never consider reducing emissions at cost c as part of its compliance strategy. It will choose to pay taxes for one unit of pollution and submit a truthful compliance report as long as

---

[10] Our one-period model assumes that the regulator does not alter its policy of random inspections in response to information acquired about firms' types.

[11] A two-period model of incomplete enforcement that would explicitly differentiate between "before" and "after" inspections would not alter any of our results.

$$\tau < \phi(F + c) + (1 - \phi)0 \tag{3}$$

Note that if the firm is caught with a false report, it must not only pay the fine *F* but also reduce a one unit of pollution to be in compliance as indicated in its report.[12] The cut-off point for a truthfully compliance report when $c \geq \tau$ is then

$$\hat{c} = \frac{\tau}{\phi} - F\,. \tag{4}$$

Thus, if $\bar{c} \geq c > \hat{c}$, the firm does not reduce pollution, but rather pays taxes and submits a truthful compliance report. Conversely, if $\tau \leq c < \hat{c}$, the firm does not reduce pollution, does not pay taxes and submits a false compliance report.

Compliant firms are those with either low or high costs. Low-cost firms find it cheaper to reduce one unit of pollution than to face the expected penalty and control costs. High-cost firms, on the other hand, find it cheaper to pay the tax than the expected penalty and control costs. Notice also that if c is uniformly distributed the fraction of noncompliant firms "before inspection" is $(\hat{c} - \tilde{c})/(\bar{c} - \underline{c})$, and "after inspection," this fraction drops to $(1-\phi)(\hat{c} - \tilde{c})/(\bar{c} - \underline{c})$.

A simple comparative static can be done with the aid of **Figure 1**. Individual control costs and the probability of being inspected are on the vertical and horizontal axes, respectively. For a given τ and *F*, it is drawn $\tilde{c} = \tilde{c}(\phi)$ and $\hat{c} = \hat{c}(\phi)$. As ϕ increases, the "before-inspection" fraction of noncompliant firms, $(\hat{c} - \tilde{c})/(\bar{c} - \underline{c})$, reduces. At a point such as $\phi_f < 1$ there would be full compliance and no need for inspection.(Note that at $\phi = \phi_f$, $\tau = \tilde{c} = \hat{c}$). Because enforcement is costly, however, the actual inspection probability ϕ_a is lower than ϕ_f, which inevitably leads to a fraction $(1-\phi_a)(\hat{c}_a - \tilde{c}_a)/(\bar{c} - \underline{c})$ of "after-inspection" noncompliant firms. The fine *F* also affect

---

[12] The alternative enforcement policy explored in Section 6 let the firm to comply by either reducing one unit of pollution or paying the tax.

compliance. If F increases $\tilde{c}(\phi)$ shifts upward, $\hat{c}(\phi)$ shifts downward and the non-compliance area shrinks.

Finally, because of enforcement power, the total reduction after inspection and (effective) enforcement is given by

$$q_e = G(\tilde{c}) + \phi[G(\hat{c}) - G(\tilde{c})] \tag{5}$$

which represents reductions from low-cost compliant firms and a fraction ϕ of non-compliant firms that return to compliance by reducing one unit of pollution. The second term of Eq. (5) shows that enforcement power yields much higher compliance rates than what can simply be attributed to F; in the spirit of Livernois and McKenna (1999).[13] In addition, control cost after inspection and enforcement is given by

$$C_e(q_e) = \int_{\underline{c}}^{\tilde{c}} c\mathrm{d}G + \phi \int_{\tilde{c}}^{\hat{c}} c\mathrm{d}G \tag{6}$$

Because some relatively high cost non-compliant firms found in violation must reduce pollution, $C_e(q_e)$ is not cost-effective. In other words, q_e is not achieved at the lowest possible cost. As compliance increases, however, $C_e(q_e)$ approaches $C(q)$, and when $\phi = \phi_f$ (i.e., full compliance), $C_e(q_e) \equiv C(q)$.

2.2 *Compliance with tradable quotas (quantities)*

Compliance with tradeable quotas is similar to compliance with taxes, but for a market clearing condition. As we shall show, it makes no difference whether the regulator distributes the tradeable quotas for free or auctions them off. Without loss of generality, consider the regulator distributes gratis a total amount of x pollution quotas, so each firm receives x. Assume for the

---

[13] If only F is used, the total reduction would be $G(\phi F)$

moment that the market equilibrium price of quotas is p. Let us first consider the case where $c < p$. In this case, a firm would never consider buying quotas as part of its compliance strategy. It would instead choose to reduce one unit of emissions at cost c, sell x quotas at price $p > 0$, and submit a truthful compliance report as long as

$$c - xp < \phi(F + c - xp) + (1 - \phi)(-xp) \tag{7}$$

Therefore, the cut-off point for a truthful compliance report when $c < p$ is, as before

$$\tilde{c} = \frac{\phi}{1 - \phi} F \; . \tag{8}$$

If $\underline{c} \leq c < \tilde{c}$, the firm reduces pollution, sells its quotas at a market price p, and submits a truthful compliance report. Conversely, if $\tilde{c} < c < p$, the firm does not reduce pollution, sells all its quotas, and submits a false report indicating that it is reducing all pollution. Note that the latter is an optimal non-compliance strategy.[14]

Let us now consider the case where $c \geq p$. In this case, a firm would never consider reducing emissions at cost c as part of its compliance strategy. It would choose to buy $1–x$ quotas, and submit a truthfully compliance report as long as

$$(1 - x)\,p < \phi(F + c - xp) + (1 - \phi)(-xp) \tag{9}$$

Therefore, the cut-off point for a truthful compliance report when $c \geq p$ is

[14] The alternative non-compliance strategy would be to keep a positive amount y (greater or lower than x) of quotas, not reduce pollution, and submit a false report for the difference $1–y$ of non-compliance. It is not difficult to show that the lowest expected cost is for $y = 0$, which is the previous non-compliance strategy.

$$\hat{c} = \frac{p}{\phi} - F\,. \tag{10}$$

If $\bar{c} \geq c > \hat{c}$, the firm does not reduce pollution, but buys quotas and submits a truthful compliance report. Conversely, if $p \leq c < \hat{c}$, the firm does not reduce pollution, sells its quotas, and submits a false report. Again, the latter is an optimal non-compliance strategy.

The fraction of "before-inspection" non-compliance, $(\tilde{c} - \hat{c})/(\bar{c} - \underline{c})$, as a function of ϕ could also be illustrated in **Figure 1** by simply substituting p for τ. This suggests an apparent "compliance equivalence" between taxes and tradeable quotas. However, p is endogenously determined, so compliance would be the same as long as the allocation x yields an equilibrium price $p = \tau$.

To find the market equilibrium price of quotas p (or alternatively, the auction clearing price), we impose the market clearing condition that sales equal purchases[15]

$$x\int_{\underline{c}}^{\hat{c}} \mathrm{d}G = (1 - x)\int_{\hat{c}}^{\bar{c}} \mathrm{d}G \tag{11}$$

Expression (11) indicates both that noncompliant firms sell all their quotas in the market and that no firm found in violation goes back to the market to buy quotas. The latter is because the optimal strategy for a noncompliant firm is to sell all its quotas and eventually reduce pollution if found in violation.[16]

Replacing (10) into (11), we obtain

---

[15] If x quotas are to be auctioned off, the market clearing condition would be $x = \int_{\hat{c}}^{\bar{c}} \mathrm{d}G$, which is eq. (9) rearranged. This and the fact that eqs. (5) and (7) are independent of the individual allocation of quotas demonstrate that in this particular context, it is irrelevant whether quotas are distributed gratis or auctioned off.

[16] The alternative enforcement policy explored in Section 6 let the firm to comply by either reducing one unit of pollution or buying quotas.

$$p = \phi G^{-1}(1-x) + \phi F \tag{12}$$

where $G^{-1}(1-x)$ is the marginal cost c just after $(1-x)$ units of pollution have been reduced, or alternatively, the price that we would observe under full compliance p_f (e.g. if $\phi \geq \phi_f$). Note that for a uniform distribution of $g(c)$, $G^{-1}(1-x) = \overline{c} - x(\overline{c} - \underline{c})$ and $p = \phi[\overline{c} - x(\overline{c} - \underline{c}) + F]$.[17]

It is not immediately apparent from eq. (12) whether the equilibrium price would be higher or lower than the price under full compliance p_f, but it is not difficult to demonstrate that

Lemma 1. *The equilibrium price with incomplete enforcement (p) is always lower than the price under full compliance* (p_f).

Proof. Replacing $G^{-1}(1-x)$ for p_f in eq. (12), using (10) and the fact that $p < \hat{c}$ by construction, yields $p < p_f$. Alternatively, observe that $p < p_f$ for $F = 0$ and that p grows linearly with F until full compliance, where $p = p_f$.

The reason for $p < p_f$ is that noncompliance and quotas are close substitutes, which depresses the net demand for quotas and therefore its price.[18]

Finally, emissions reduction after enforcement, q_e, and control cost after enforcement, $C_e(q_e)$, are given by expressions analogous to eqs. (5) and (6) above. Because there is additional reduction after firms have been found in violation, it is not so immediately clear whether q_e is lower or higher than $1-x$, which is the amount that was originally intended to be abated through tradeable quotas. We establish

---

[17] I introduce the uniform distribution here for later use. A uniform distribution implies that $C''(q) = (\overline{c} - \underline{c})$ over the interval $[\underline{c}, \overline{c}]$.

[18] Using a different enforcement policy, Malik (1990) found the same result when the inspection probability was the same across firms.

Lemma 2. *The total "after-enforcement" reduction (q_e) is always lower than the amount required by the original distribution of quotas (1 –x).*

Proof. From eqs. (12) and (10), we find that $\hat{c} = G^{-1}(1-x)$. This, together with eqs. (5) and (8), gives $q_e = (1-\phi)G[\phi F/(1-\phi)] + \phi(1-x)$. On the other hand, using eq. (12) and the fact that $p < p_f = G^{-1}(1-x)$ yields $G(\phi F/(1-\phi)) < (1-x)$. Combining both results finishes the proof that $q_e < (1-x)$.

3. The planning problem

Let us start with the regulator's original planning problem that is to choose the amount of control q that maximizes the social welfare function

$$W(q) = B(q) - C(q) \tag{13}$$

In a world of certainty and complete enforcement, the solution q^* must satisfy the first order condition

$$B'(q^*) = C'(q^*) \tag{14}$$

To implement the first-best outcome q^*, it makes no difference whether the regulator uses the optimal tax $\tau^* = B'(q^*) = C'(q^*)$ and has individual firms to reduce emissions accordingly, or whether s/he allocates x^* tradeable pollution quotas that produce a total reduction of $q^* = 1 - x^*$ (Weitzman, 1974; Baumol and Oates, 1988). The equilibrium price of quotas would be $p^* = \tau^*$; therefore, the regulator would show no preference for either prices (taxes) or quantities (tradeable permits). Eq. (14) can also be written as $G^{-1}(1-x^*) = B'(1-x^*) = \tau^*$, where the pair ($\tau^*$, x^*) is the first-best design that yields the first-best outcome if and only if full compliance can be enforced.

If, for some reason, the regulator does not acknowledge that enforcement is incomplete (i.e., a fraction of firms is noncompliant) and implements the first-best design either through taxes (τ*) or tradeable quotas (x*), instrument choice may have some welfare consequences.[19] We explore this matter by estimating the difference in social welfare Δ_{pq} between prices (taxes) and quantities (tradeable quotas). Using $C_e(q_e)$ instead of the "least-cost" cost curve, $C(q)$, we have

$$\Delta_{pq} \equiv W(\tau^*) - W(x^*) = B(q_e(\tau^*)) - C_e(q_e(\tau^*)) - [B(q_e(x^*)) - C_e(q_e(x^*))] \tag{15}$$

where $q_e(\tau^*)$, $C_e(q_e(\tau^*))$, $q_e(x^*)$ and $C_e(q_e(x^*))$ can be obtained from eqs. (5) and (6). In addition, these expressions require values of $\hat{c}(\tau^*)$ and $\hat{c}(x^*)$. From eq. (4), we know that $\hat{c}(\tau^*) = \tau^*/\phi - F$, and from (12) we know that $p = \phi G^{-1}(1-x^*) + \phi F = \phi\tau^* + \phi F$, which, when substituted into (10), yields $\hat{c}(x^*) = \tau^*$. Comparing $\hat{c}(\tau^*)$ and $\hat{c}(x^*)$, we can establish that

Lemma 3. *$\hat{c}(\tau^*) > \hat{c}(x^*)$, and therefore, $q_e(\tau^*) > q_e(x^*)$, $B(q_e(\tau^*)) > B(q_e(x^*))$ and $C_e(q_e(\tau^*)) > C_e(q_e(x^*))$.*

Proof. By construction we have that $\tau^* > \tilde{c} = \dfrac{\phi}{1-\phi}F$, and therefore, $\hat{c}(\tau^*) > \hat{c}(x^*)$. If this inequality holds, it is immediate from eqs. (5) and (6) that the rest of the lemma holds.

Lemma 3 indicates a welfare trade-off between prices and quantities when the regulator implements a first-best policy design in the presence of incomplete enforcement. Because $p(x^*) < p^* = \tau^*$, the number of non-compliant firms is higher under taxes, which implies higher after-enforcement reductions (q_e) and benefits (B), but also higher costs (C_e). To see under what circumstances the welfare trade-off favors prices (taxes), we develop expression (15) using linear approximations for the marginal benefit and marginal cost curves such as in Weitzman (1974) and Baumol and Oates (1988). For the benefit curve, let

[19] We leave the discussion of second-best design for the next section.

$$B(q) = \underline{b} \cdot q + \frac{B''}{2} \cdot q^2 \tag{16}$$

where $\underline{b} \equiv B'(0) > 0$ and $B'' < 0$ are fixed coefficients. We use this notation to be consistent with the cost curve, for which we simply assume that $g(c)$ follows a uniform distribution, so the slope of the marginal cost curve can be written as $C'' = \bar{c} - \underline{c}$. The first-best design would be

$$\tau^* = \frac{\underline{b}C'' - \underline{c}B''}{C'' - B''} \tag{17}$$

and

$$x^* = 1 - q^* = 1 - \frac{\underline{b} - \underline{c}}{C'' - B''} \tag{18}$$

With these simplifying assumptions and some algebraic manipulation, we obtain the following expression

$$\Delta_{\text{pq}} = -\frac{[(1-\phi)\tau^* - \phi F]^2}{2\phi(C'')^2}[C'' + \phi B''] \tag{19}$$

and state

Lemma 4. *If the regulator implements the first-best design under incomplete enforcement, the choice between prices (τ^*) vs. quantities (x^*) is not welfare neutral. In this particular enforcement policy, quantities provide higher social welfare as long as* $C'' > \phi|B''|$.

Note that under full compliance, $\tau^* = \tilde{c} = \hat{c}$ (see Figure 1) and $\Delta_{\text{pq}} = 0$, so instrument choice has no welfare consequences. Lemma 4 indicates, however, that if the absolute value of the slopes of the marginal cost and marginal benefit curves are the same, $C'' = |B''|$, and the regulator implements the first-best in the presence of incomplete enforcement, a quantity instrument always provides higher social welfare given that expression (19) is negative. These welfare differences arise because under incomplete enforcement both first-best designs (τ* and x*) are not longer optimal. Other things equal (i.e., $C'' = |B''|$), the quantity x* results to be closer to the second-best optimal design than the price τ*.

Lemma 4 has important policy implications. In many situations the task of choosing policy goals and instruments is independent from the task of policy implementation. If that is so, Lemma 4 suggests that choosing instruments under the assumption of complete enforcement may lead to the wrong choice of policy instrument.

4. Optimal design with incomplete enforcement

In the presence of incomplete enforcement, the first-best design is no longer optimal. A second-best design provides higher welfare. We first consider the case in which benefit and cost curves are known with certainty, and then turn to the uncertainty case.

4.1 Cost and benefit certainty

Let us first find the second-best tax τ**. The regulator maximizes social welfare

$$W(\tau) = B(q_e(\tau)) - C_e(q_e(\tau)) \tag{20}$$

where $C_e(q_e)$ is given by eq. (6), and $\tilde{c}$ and $\hat{c}$ are given by eqs. (2) and (4), respectively. Assuming the same linear approximations for $B(q)$ and g(c), the first-order condition can be written as

$$\underline{b}\frac{\partial q_e(\tau)}{\partial \tau} + B'' q_e(\tau)\frac{\partial q_e(\tau)}{\partial \tau} - \left(\frac{\tau}{\phi} - F\right)\frac{1}{C''} = 0 \tag{21}$$

where $q_e(\tau) = (\tau - \underline{c}) / C''$ and $\partial q_e(\tau) / \partial\tau = 1 / C''$. Thus, the solution to (21) is

$$\tau^{**} = \frac{(\underline{b}C'' - \underline{c}B'' + FC'')\phi}{C'' - \phi B''} \tag{22}$$

Note that under full compliance (for example, when ϕ =1 and F = 0), expression (22) becomes (17). In addition, we establish that

Lemma 5. $\tau^{**} < \tau^{*}$.

Proof. For linear approximations, the proof is straightforward. Develop τ^* and τ^{**} according to (17) and (22), and use, by construction, $\frac{\phi}{1-\phi} F = \tilde{c} < \tau^*$.[20]

When enforcement is incomplete, Lemma 6 indicates that for this particular enforcement structure the regulator is better off setting a less stringent policy.[21]

Let us now find the second-best quota allocation x^{**}. This can be found indirectly by estimating p^{**}, the equilibrium price that would prevail if x^{**} is implemented. If we replace τ by p in (20) and let the regulator choose p that maximizes social welfare, we obtain

$$p^{**} = \frac{(\underline{b}C'' - \underline{c}B'' + FC'')\phi}{C'' - \phi B''} = \tau^{**} \tag{23}$$

and, using eq. (12)

[20] Note that as ϕ decreases, τ^{**} approaches zero; and since $\partial\tau^{**}/\partial\phi > 0$ (from eq. (12)), the proof is also immediate.

[21] See Viscusi and Zeckhauser (1979) for a similar result.

$$x^{**} = 1 - G\left(\frac{p^{**}}{\phi} - F\right) = 1 + \frac{[(1-\phi)\underline{c} - \phi F]B'' - (\underline{b} - \underline{c})C''}{(C'' - \phi B'')C''} \tag{24}$$

Note that under full compliance (for example, when ϕ =1 and F = 0), expression (24) becomes (18). Before discussing eq. (23), we establish

Lemma 6. $x^{**} > x^*$.

Proof. Since $p^{**} = \tau^{**}$, from eq. (23), and $p^* = \tau^*$; employing Lemma 5, we find that $p^* > p^{**}$. Using eq. (12) and the fact that $G(\cdot)$ is an increasing function completes the proof.

Lemmas 5 and 6 and eq. (23) can be summarized in the following two propositions

Proposition 1. *A first-best design through either prices (τ^*) or quantities (x^*) is always inefficient (i.e., there is a design that yields higher welfare) in the presence of incomplete enforcement.*

Proposition 2. *When cost and benefit curves are known with certainty, it is irrelevant whether the regulator uses prices (τ^{**}) or quantities (x^{**}) to achieve the second-best outcome. Both instruments provide the same social welfare.*

Proposition 1 (its proof is immediate from Lemmas 5 and 6) has very important policy implications. If the task of choosing policy goals is done independently from policy implementation considerations, policy goals would be wrong. For this particular enforcement policy they would be too stringent.

Proposition 2 does not require a formal proof either. It suffices to realize that x^{**} yields p^{**} and that the compliance strategies are identical if $p = \tau$. Proposition 2 argues, as in the case of complete enforcement (e.g. Baumol and Oates, 1988), that when cost and benefit curves are known with certainty, incomplete enforcement does not affect instrument choice as long as the regulator implements the second-best design. In short, in a second-best context, quantities and prices are fully equivalent.

4.2 Cost and benefit uncertainty

It is well known that regulators' task of choosing policy goals and instruments must be carried out in the presence of significant uncertainty concerning $C(q)$ and $B(q)$. Following Weitzman (1974) and Baumol and Oates (1988), we model uncertainty as an additive stochastic error term. For the benefit curve, let $\partial B(q,\theta)/\partial q = B'(q) + \theta$, where θ is a random shock such that $E[\theta] = 0$ and $E[\theta^2] = \sigma_\theta^2$. For the cost curve, let $c(\eta) = c + \eta$, where η is another random shock such that $E[\eta] = 0$ and $E[\eta^2] = \sigma_\eta^2$. All individual costs face the same shock, which produces the desired "parallel" shift of the aggregate marginal cost curve, $C'(q)$, in the amount η. In other words, we have that $\partial C(q,\eta)/\partial q = C'(q) + \eta$.

The regulator's planning problem now is to choose τ (or x) that maximizes expected social welfare. In the case of taxes, the regulator maximizes

$$E[W(\tau,\theta,\eta)] = E\left[B(q_e(\tau,\eta),\theta) - C_e(q_e(\tau,\eta))\right] \tag{25}$$

where[22]

$$q_e(\tau,\eta) = \int_{\underline{c}+\eta}^{\tilde{c}} g(c-\eta)\mathrm{d}c + \phi\int_{\tilde{c}}^{\hat{c}} g(c-\eta)\mathrm{d}c \tag{26}$$

$$C_e(q_e,\eta) = \int_{\underline{c}+\eta}^{\tilde{c}} cg(c-\eta)\mathrm{d}c + \phi\int_{\tilde{c}}^{\hat{c}} cg(c-\eta)\mathrm{d}c \tag{27}$$

[22] We replace dG by gdc and $g(c)$ by $g(c\text{-}\eta)$, so we still have that $G(\underline{c}+\eta-\eta)=0$ and $G(\bar{c}+\eta-\eta)=1$.

and $\tilde{c}$ and $\hat{c}$ are given by eqs. (2) and (4). Substituting eqs. (26) and (27) into (25), and taking derivative with respect to τ, we obtain the first-order condition

$$E\left[\underline{b}\frac{\partial q_e(\tau,\eta)}{\partial\tau}+\theta\frac{\partial q_e(\tau,\eta)}{\partial\tau}+B''q_e(\tau,\eta)\frac{\partial q_e(\tau,\eta)}{\partial\tau}-\left(\frac{\tau}{\phi}-F\right)\frac{1}{C''}\right]=0 \tag{28}$$

where $q_e(\tau,\eta)=(\tau-\underline{c}-\eta)/C''$ and $\partial q_e(\tau,\eta)/\partial\tau=1/C''$. Taking expectation, we find the solution to eq. (25) to be τ**; the second-best tax under certainty.

Similarly, in the case of tradeable quotas, the regulator chooses x that maximizes

$$E[W(x,\theta,\eta)]=E\left[B(q_e(x,\eta),\theta)-C_e(q_e(x,\eta))\right] \tag{29}$$

Note that this formulation is slightly different from the certainty case, where we first obtain the optimal price of quotas, p**, and then the optimal allocation, x**. Here it is not immediate if we can proceed as before, because p is a random variable. In fact, using the market clearing condition (11), we have

$$1-x=\int_{\underline{c}+\eta}^{\hat{c}}\mathrm{d}G=G\left(\frac{p}{\phi}-F-\eta\right) \tag{30}$$

and therefore,

$$p(x,\eta)=\phi G^{-1}(1-x)+\phi F+\phi\eta=\overline{p}+\phi\eta \tag{31}$$

Eq. (31) shows that price fluctuations are lower under incomplete enforcement than under full enforcement, where the equilibrium price would be $\overline{p}+\eta$.

The formal solution to eq. (29) is left to the reader. We instead present an indirect method that yields the same result because of the linear approximation we are using. Consider $\overline{p}$ as the decision variable and, as before, replace it by τ in eq. (25). Not surprisingly, the solution is $\overline{p} = \tau^{**} = p^{**}$. Since the random variable η enters linearly in (31), the allocation x that solves $E[p] \equiv \overline{p} = p^{**}$ must be x^{**}. Therefore, the allocation of quotas x^{**} remains second-best optimal under uncertainty.

To summarize:

Proposition 3. *Cost and benefit uncertainty does not alter the second-best design neither for either prices (τ^{**}) or quantities (x^{**}) when marginal benefit and cost curves are assumed linear.*

Baumol and Oates (1998) obtained exactly the same result for Weitzman's (1974) problem, that uncertainty does not affect the second-best design when the marginal benefit and cost curves are assumed linear.

5. Instrument choice under uncertainty and incomplete enforcement

We now turn to the central question of this paper as to whether there should be any preference for prices over quantities (or vice versa) when cost and benefit curves are uncertain and enforcement is incomplete. We explore this question by estimating the difference in expected social welfare

$$\Delta_{pq} \equiv E[W(\tau^{**},\theta,\eta) - W(x^{**},\theta,\eta)] \tag{32}$$

Either design (τ^{**} or x^{**}) is *ex-ante* second-best optimum, but because of uncertainty neither one will be *ex post* second-best optimum. The relevant question here then becomes: which instrument comes closer to the *ex-post* second-best optimum? The normative implication of (32) is that if $\Delta_{pq} > 0$, prices provide higher expected welfare than do quantities, and accordingly, ought to be preferred as policy instrument. If $\Delta_{pq} < 0$, quantities ought to be preferred.

The development of (32) simplifies greatly if we use $p = \tau^{**} + \phi\eta$, as indicated by eq. (31), that together with linear approximations for marginal curves, leads to the following

$$q_e(\tau^{**},\eta) = \frac{\tau^{**} - \underline{c} - \eta}{C''} \tag{33a}$$

$$q_e(x^{**},\eta) = \frac{\tau^{**} + \phi\eta - \underline{c} - \eta}{C''} \tag{33b}$$

$$\hat{c}(\tau^{**},\eta) = \frac{\tau^{**}}{\phi} - F \tag{33c}$$

$$\hat{c}(x^{**},\eta) = \frac{\tau^{**} + \phi\eta}{\phi} - F = \frac{\tau^{**}}{\phi} - F + \eta \tag{33d}$$

Substituting these expressions into (32), Δ_{pq} becomes

$$\begin{aligned}\Delta_{\mathrm{pq}} \equiv E\Bigg[(\underline{b} + \theta)\left(\frac{-\phi\eta}{C''}\right) + \frac{B''}{2}\left(\frac{2\tau^{**} - 2\underline{c} - (2-\phi)\eta}{C''}\right)\left(\frac{-\phi\eta}{C''}\right)\ldots \\ - \frac{\phi}{2C''}\left(\frac{2\tau^{**}}{\phi} - 2F + \eta\right)(-\eta)\Bigg]\end{aligned} \tag{34}$$

Taking expectation, and assuming that $E[\theta\eta] = 0$, the right-hand side reduces to

$$\Delta_{\mathrm{pq}} = \frac{\phi\sigma_\eta^2}{2(C'')^2}[C'' + B'' + (1-\phi)B''] \tag{35}$$

Eq. (35) is the main result of this paper. Its implications can be better understood by recalling Weitzman's (1974) result, which can be seen by simply plugging $\phi = 1$ into eq. (35). Weitzman's result would suggest using prices as long as the marginal cost curve is steeper than the marginal benefit curve, that is to say, as long as $C'' > |B''|$. With incomplete enforcement, however, the advantage of prices over quotas reduces substantially given that $1-\phi > 0$. In fact, eq. (35) recommends to use prices if and only if $C'' > (2-\phi)|B''|$. Furthermore, if $C'' = |B''|$ quantities ought to be preferred as the policy instrument. We summarize the main result of the paper in the following proposition

Proposition 4. *Cost and benefit uncertainty does affect instrument choice. Under incomplete enforcement, the advantage of prices (τ^{**}) over quantities (x^{**}) is reduced substantially. Prices ought to be preferred if and only if $C'' > (2-\phi)/B''/$. Furthermore, if $C'' = /B''/$, a quantity instrument should be the preferable policy choice.*

To get the intuition behind Proposition 4, it is useful to recall Weitzman's (1974) basic reasoning of using prices over quantities as long as missing the *ex-post* optimum amount of control has lower welfare consequences than missing the *ex-post* optimum (marginal) cost of control, which happens when the marginal cost curve is steeper than the marginal benefit curve. In a quantity regime with full compliance the amount of control remains always fixed while the cost of control is the subject to large swings because of uncertainty. If the marginal cost curve is really steep, the (marginal) cost of control can be very well off the *ex-post* optimum, and that is when, a price instrument that fixes the marginal cost of control may be more appropriate in that situation.

With incomplete enforcement, however, the effective (or observed) amount of control under a quantity instrument is not longer fixed. Instead, it adapts to the possible shocks affecting the marginal cost curve. Indeed, if the marginal costs curve turns out to be higher than expected, some firms would choose not to comply with the regulation, and consequently, the effective amount and cost of control would be lower. Because a quantity instrument would now have this "flexibility" for the cost of control to adapt to unexpected shocks, the advantage of prices over quantities is necessarily reduced. Put differently, the marginal cost curve becomes relatively flatter under incomplete enforcement. A more intuitive explanation would be that under incomplete enforcement the regulator wants to pay closer attention to the amount of control (i.e., emissions reduction) than to the cost of control *ceteris paribus*; a task in which a quantity instrument is more effective.

Finally, note that if marginal cost and marginal benefit curves are positively correlated, that is $E[\theta\eta] > 0$, an additional negative term enters into (35), which increases the advantage of quantities over prices increases; otherwise benefit uncertainty does not intervene.[23]

6. Extensions

Results so far are based on a very specific enforcement policy. It is natural to think they may change in the context of alternative enforcement policies. Here we extend the model to consider one alternative enforcement policy and demonstrate, in particular, that our main result stated in Proposition 4 does not change.

The new enforcement policy still retains the elements of self-reporting and enforcement power from Livernois and Mackenna (1999). Instead of forcing a noncompliant firm to follow its compliance report in order to come under compliance, in this new enforcement policy, the regulator lets a noncompliant firm found to be in violation to choose any strategy to return to compliance. Consider, for example, the case of a firm with $c > \tau$ found to be in violation under a tax regulation (prices). For this firm, the cheapest strategy to come under compliance would be to pay taxes instead of reducing one unit of pollution. A similar situation occurs under a tradeable quota regime (quantities). A firm with $c > p$ found to be in violation would always prefer to buy quotas instead of reducing pollution as a way to come under compliance.

Under this enforcement policy the cut-off points for a truthful compliance report are $\tilde{c} = \phi F / (1-\phi)$ and $\hat{c} = \bar{c}$. For an interior solution we assume, as before, that τ (and p) > $\tilde{c} = \phi F / (1-\phi)$; otherwise there will be full compliance. Under this new enforcement policy, the submission of false reports increases because the expected penalty for high cost firms has decreased. Thus, we have

[23] See Stavins (1996) for a complete discussion on correlated uncertainty.

$$q_e(\tau) = \frac{(\tau + F)\phi - \underline{c}}{C''} \qquad (36a)$$

$$q_e(x) = \frac{(p + F)\phi - \underline{c}}{C''} \qquad (36b)$$

where $p = p(x)$ can be obtained from the following market clearing condition (assuming rational expectations)

$$x\int_{\underline{c}}^{\tilde{c}} \mathrm{d}G + x\int_{\tilde{c}}^{\bar{c}} \mathrm{d}G = \phi \cdot 1 \cdot \int_{p}^{\bar{c}} \mathrm{d}G \qquad (37)$$

where the first term corresponds to sales from compliant firms, the second term corresponds to sales from before-inspection noncompliant firms, and the term of the right hand side corresponds to purchases from noncompliant firms found to be in violation that go to the market to buy quotas to return to compliance. A uniform distribution of $g(c)$ (i.e., linear aggregate marginal costs) yields[24]

$$p(x) = \bar{c} - \frac{C''}{\phi} x \qquad (38)$$

and under uncertainty

$$p(x, \eta) = \bar{c} - \frac{C''}{\phi} x + \eta \qquad (39)$$

[24] The general form would be $p(x) = G^{-1}(1 - x/\phi)$.

It is important to mention that this new enforcement policy differs from the previous one in some important ways. While Lemmas 1 and 2 still hold, Lemma 3 does no longer apply because $\hat{c}(\tau^*) = \hat{c}(x^*) = \bar{c}$, Lemma 4 shifts in favor of prices, and Lemmas 5 and 6 are reverted (i.e., $\tau^{**} > \tau^*$ and $x^{**} < x^*$). The latter means that under incomplete enforcement, it is second-best optimal to set less stringent regulatory levels.

Finally, to see whether the main result of this paper—Proposition 4—continues to hold under this new enforcement policy (the proof of Propositions 1, 2 and 3 are left to the reader), we estimate Δ_{pq} by substituting the above expressions into the appropriate expected welfare function (see eq. (32)) and using, as before, the fact that

$$p(x^{**},\eta) = \tau^{**} + \eta \qquad (40)$$

the difference in expected welfare reduces to

$$\Delta_{\text{pq}} = \frac{\phi\sigma_\eta^2}{2(C'')^2}[C'' + B'' + (1-\phi)B''] \qquad (41)$$

which is the same expression we obtained for the previous enforcement policy (see eq. (35)). And as before, the reason is that under a quantity regime with incomplete enforcement, the amount of control does not remain fixed, but it rather adapts somewhat to unexpected cost shocks, which ultimately reduces the advantage of prices over quotas.

It would be useful to extend the analysis to other enforcement policies. A common approach used in earlier papers (e.g., Viscusi and Zeckhauser, 1977; and Malik, 1990) is to simply consider a fine F and no enforcement power. Although we cannot test such an approach here because the assumption of constant individual marginal costs would not yield interior solutions, there are no reasons to think that the results stated in Propositions 1-3 will no longer hold. Proposition 4 may be harder to prove unless one accepts that firms would respond to unexpected shocks in their control costs by changing their compliance strategies as they did in the two cases studied here.

6. Conclusions

There is no doubt that imperfect monitoring and incomplete enforcement are central in understanding regulatory performance in practice, and therefore, it is very relevant to ask whether they have any policy implications in the choice of regulatory instruments. In this paper, I extended Weitzman's (1974) "Prices vs. Quantities" to allow for incomplete enforcement. I found that whether the regulator uses prices (e.g. taxes) or quantities (e.g. tradeable quotas), a first-best design is inefficient in the presence of incomplete enforcement. A second-best design that incorporates incomplete enforcement and where cost and benefit curves are known with certainty, can be implemented equally with either instrument. If benefit and cost curves are uncertain, however, a quantity instrument performs better. In fact, if the slopes of the marginal cost and marginal benefit curves are equal, quantities are always preferred over prices.

The rationale for the latter result is that with incomplete enforcement the effective (or observed) amount of control under a quantity instrument is not longer fixed. Instead, it adapts to the possible shocks affecting the marginal cost curve. Indeed, if the marginal costs curve turns out to be higher than expected, some firms would choose not to comply with the regulation, and consequently, the effective amount and cost of control would be lower. Because a quantity instrument would now have this "flexibility" for the cost of control to adapt to unexpected shocks, the advantage of prices over quantities is necessarily reduced. Put differently, the marginal cost curve becomes relatively flatter under incomplete enforcement. A more intuitive explanation would be that under incomplete enforcement the regulator wants to pay closer attention to the amount of control (i.e., emissions reduction) than to the cost of control *ceteris paribus*; a task in which a quantity instrument is more effective.

A natural extension of the work presented here would be to include these results into a more general theory of regulatory instrument choice. I am sure that such a theoretical exercise would have to rely heavily on numerical simulations.

References

Baldursson, F.M., and N.-H. von der Fehr (1998), Prices vs. quantities: The irrelevance of irreversibility, working paper #9, Department of Economics, University of Oslo, Oslo.

Baumol, W.J., and W.E. Oates (1988), *The Theory of Environmental Policy*, Second edition, Cambridge University Press, Cambridge.

Finkelshtain, I. and Y. Kislev (1997), Prices versus quantities: The political perspective, *Journal of Political Economy* 105, 83-100.

Fisher, B.S., S. Barret, P. Bohm, M. Kuroda, J.K.E. Mubazi, A. Shah, R.N. Stavins (1996), An economic assessment of policy instruments for combating climate change, in J.P Bruce, H. Lee and E.F. Haites (eds.), Climate Change 1995: Economic and Social Dimensions of Climate Change, Cambridge University Press, Cambridge.

Harford, J.D. (1978), Firm behavior under imperfectly enforceable pollution standards and taxes, *Journal of Environmental Economics and Management* 5, 26-43.

Harrington, W. (1988), Enforcement leverage when penalties are restricted, *Journal of Public Economics* 37, 29-53.

Hoel, M., and L. Karp (1998), Taxes versus quotas for a stock pollutant, Working paper 29/98, Fondazione Eni Enrico Mattei, Milan.

Livernois, J., and C.J. McKenna (1999), Truth or consequences: Enforcing pollution standards with self-reporting, *Journal of Public Economics* 71, 415-440.

Malik, A. (1990), Markets for pollution control when firms are noncompliant, *Journal of Environmental Economics and Management* 18, 97-106.

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

Russell, C.S. (1990), Monitoring and enforcement, in *Public Policies for Environmental Protection*, P. Portney (ed.), Resources for the Future, Washington, D.C.

Stavins, R. (1996), Correlated uncertainty and policy instrument choice, *Journal of Environmental Economics and Management* 30, 218-232.

Viscusi, W.K., and R.J. Zeckhauser (1979), Optimal standards with incomplete enforcement, *Public Policy* 27, 437-456.

Weitzman, M. (1974), Prices vs. quantities, *Review of Economics Studies* 41, 477-491.

Yohe, G.W. (1978), Towards a general comparison of price controls and quantity controls under uncertainty, *Review of Economic Studies* 45, 229-238.

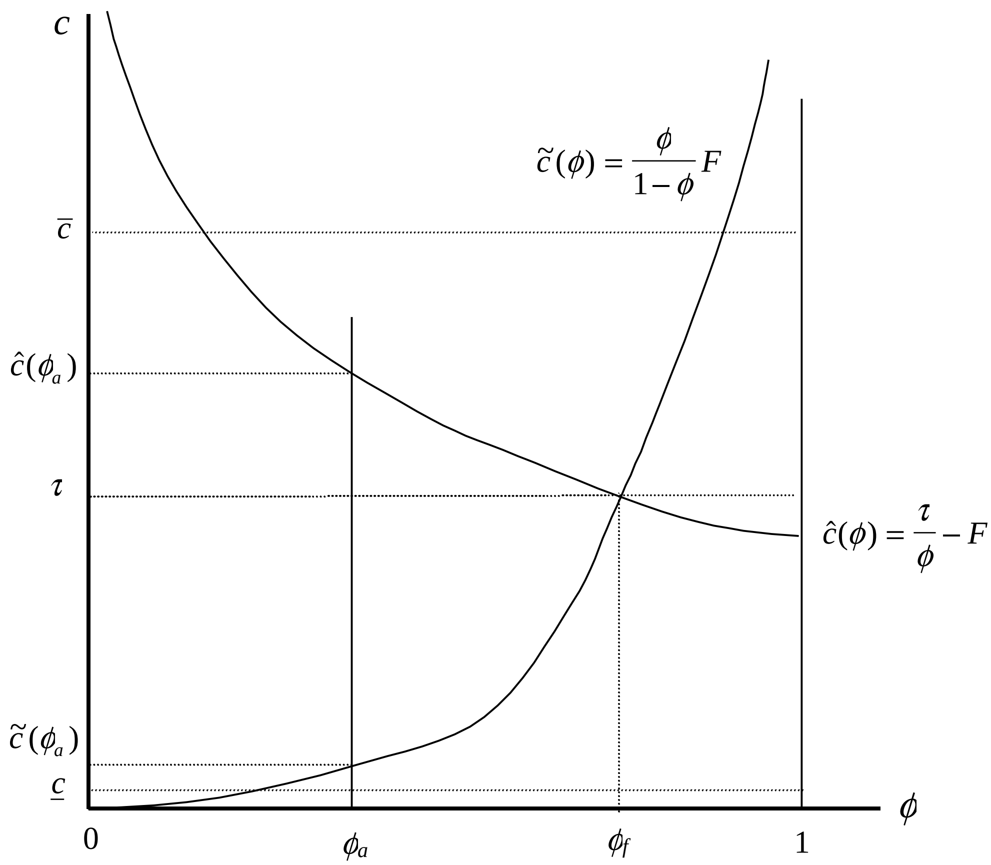

Figure 1. Compliance with prices (taxes)