Cambridge Working Papers in Economics CWPE 0457

UNIVERSITY OF CAMBRIDGE
Department of Applied Economics

Tradable Permits with Incomplete Monitoring: Evidence from Santiago's Particulate Permits Program

Juan-Pablo Montero

The Cambridge-MIT Institute

*Massachusetts Institute of Technology
Center for Energy and
Environmental Policy Research*

CMI Working Paper 57

Cambridge Working Papers in Economics

UNIVERSITY OF
CAMBRIDGE
**Department of
Applied Economics**

Not to be quoted without permission

The
Cambridge-MIT
Institute

*Massachusetts Institute of Technology
Center for Energy and
Environmental Policy Research*

CMI Working Paper

Tradable permits with incomplete monitoring: Evidence from Santiago's particulate permits program

Juan-Pablo Montero*

Catholic University of Chile and MIT

August 23, 2004

Abstract

I explore the advantages of tradable emission permits over uniform emission standards when the regulator has incomplete information on firms' emissions and costs of production and abatement (e.g., air pollution in large cities). Because the regulator only observes each firm's abatement technology but neither its emissions nor its output, there are cases in which standards can lead to lower emissions and, hence, welfare dominate permits. I then empirically examine these issues using evidence from a particulate permits market in Santiago, Chile.

Keywords: asymmetric information, imperfect monitoring, pollution markets, permits

JEL classification: L51, Q28

*<jmontero@faceapuc.cl> Department of Economics, Catholic University of Chile, Vicuna Mackenna 4860, Santiago, Chile. The author is also Research Associate at the MIT Center for Energy and Environmental Policy Research (MIT-CEEPR). I wish to thank J.R. DeShazo, Charlie Kolstad (the Editor), two anonymous referees and participants at the UC-Santa Barbara Workshop on Economic Instruments for numerous comments.

1 Introduction

Attention to tradeable emission permits (or emissions trading) as an alternative to the traditional command-and-control (CAC) approach of setting uniform emission and technology standards has significantly increased in the last decade or so. A notable example is the 1990 U.S. Acid Rain program that implemented a nationwide market for electric utilities' sulfur dioxide (SO_2) emissions (Schmalensee et al., 1998; Ellerman et al., 2000). In order to have a precise estimate of the SO_2 emissions that are going to the atmosphere, the Acid Rain program requires each affected electric utility unit to install costly equipment that can continuously monitor emissions. Another example with similar monitoring requirements is the Southern California RECLAIM program that implemented separated markets for nitrogen oxide (NOx) and SO_2 emissions from power plants, refineries and other large stationary sources.[1]

These and other market experiences, which are also documented by Stavins (2004) and Tietenberg (2004) elsewhere in this book, suggest that conventional tradable permits programs are likely to be implemented in those cases where emissions can be closely monitored, which almost exclusively occurs in large stationary sources like electric power plants and refineries. At least this is consistent with the evidence that environmental authorities continue relying on CAC instruments to regulate emissions from smaller sources for which continuous monitoring is prohibitively costly (or technically unfeasible). In such cases, compliance with CAC instruments only requires the authority to ensure that the regulated source has installed the required abatement technology or that its emissions per unit of output are equal or lower than a certain emissions rate standard.

These observations raise the question as to why the flexibility of permit trading cannot

[1] It is worth noting that RECLAIM did not include a market for volatile organic compounds (VOC) in large part because of the difficulties with monitoring actual emissions from smaller and heterogeneous sources (Harrison, 1999).

be extended to the regulation of sources whose emissions can only be imperfectly measured through the observation of their abatement technologies or emission rates, as would be done under CAC regulation.[2] Since under such a (second-best) permit scheme sources would not be trading emissions but some proxy for emissions, one may conjecture that actual emissions can be higher or lower than under an alternative CAC regulation. One can argue, for example, that emissions are likely to be higher if the trade pattern is such that lower-output firms sell permits to higher-output firms.

In looking for an answer to the above question, it is interesting to observe that despite its limited information on each source's actual emissions (and costs), Santiago-Chile's environmental agency has already implemented a market to control total suspended particulate (TSP) emissions from a group of about 600 stationary sources (Montero et al., 2002).[3] Based on estimates from annual inspections for technology parameters such as source's size and fuel type, Santiago's environmental regulator approximates each source's actual emissions by the maximum amount of emissions that the source could potentially emit in a given year. In particular, the observable firm's emission rate (mg/m^3) is multiplied by its maximum possible output (m^3/year) to infer its maximum emissions (mg/year) for which the firm must buy permits.[4]

Since most of the literature on environmental regulation under asymmetric information deals with the case in which firms' costs are privately known but emissions are publicly observed (see Lewis (1996) for a survey), a closer examination of Santiago's TSP permits program represents a unique case study of issues of instrument choice and design that can arise in the practical

[2] It is also assumed that the firm's output or utilization is not observed by the regulator, so actual emissions can not be indirectly inferred.

[3] These 600 sources affected by the TSP program are responsible for only 5% of 2000 TSP emissions in Santiago. Remaining TSP sources are controlled through CAC regulation.

[4] As we shall see later, using the source's maximum emissions as a proxy does not prevent any adverse effects that the use of permits (instead of CAC regulation) could eventually have on aggregate emissions. The choice of proxy is an arbitrary matter because the number of permits being allocated can always be adjusted accordingly with no efficiency effects.

implementation of permits markets under imperfect monitoring of emissions (e.g., air pollution in large cities).[5] While there is some literature looking at the latter (e.g., Segerson, 1988; Fullerton and West, 2002; Cremer and Gahvari, 2002),[6] only Montero (2004) focus specifically on the effect of imperfect information about emissions and costs on the design and performance of a permits market.

In comparing permits versus standards, Montero (2004) identifies a trade-off between cost-savings and possible higher emissions.[7] On the one hand, the permits policy retains the well known cost-effectiveness property of conventional permits schemes (i.e., those based on actual emissions) that is that permit trading allows heterogenous firms to reduce their abatement and production costs. On the other hand, the permits policy can sometimes provide firms with incentives to choose combinations of output and abatement technology that may lead to higher aggregate emissions than under standards (i.e., CAC regulation). Thus, when (abatement and production) cost heterogeneity across firms is large, the permits policy is likely to work better. In contrast, as heterogeneity disappears, the advantage of permits reduces, and standards might work better provided that they lead to lower emissions.

In this paper, I extend the theoretical model of Montero (2004) and then apply it to the TSP program with the purpose of comparing the actual performance of this program with that of a hypothetically equivalent standards policy. In doing so, I first recover production

[5] Varios of the permit trading programs documented in USEPA (2001) face similar issues because these are programs that are not based on actual emissions (e.g., the averaging programs for mobile sources, the fireplace permit trading in Colorado, etc.).

[6] Segerson (1988) study the control of emissions from (few) non-point sources using a "moral hazard in teams" approach. Fullerton and West (2002) consider the control of vehicle emissions using a combination of taxes on cars and on gasoline as an alternative to an (unavailable) tax on emissions. Cremer and Gahvari (2002) look at output and emission taxation under costly (rather than imperfect) monitoring.

[7] The instrument choice problem studied by Montero (2004) paper is similar in spirit and approach to the instrument choice dilemma considered by Weitzman (1974). There are, however, important differences. Weitzman (1974) compares the relative advantage of a price instrument (taxes) over a quantity instrument (permits) when the regulator has imperfect information about the aggregate abatement cost curve (and possibly about the damage curve as well). Thus, cost heterogeneity across firms plays no role in Weitzman's analysis. Instead, Montero (2004) compares the performance of two quantity instruments and focus on the effect of cost heterogeneity on instrument performance.

and abatement cost characteristics of affected sources and the regulator's perception about environmental damages. Based on these estimates, I find that permits have provided large cost-savings but also lead to higher emissions; about 6% higher than what would have been observed under an equivalent standards policy. However, the welfare loss from higher emissions is only 8% of the welfare gain from lower abatement and production costs.

The theoretical and empirical results of this paper make a strong case for the wider use of pollution permits even in those situations in which emissions are imperfectly observed. Furthermore, because permits are always less costly than standards and may or may not lead to higher emissions, I would add that permits should be adopted as a default, unless the available cost and pollution damage information indicates the opposite. In other words, the burden of proof should lie with the CAC policy and not with the permits policy. Nevertheless, in many cases it may be welfare improving to combine permits with some optimally chosen standard rather than just use permits (Montero, 2004).[8]

The rest of the paper is organized as follows. In Section 2, I present the model. I consider a competitive market for an homogeneous good supplied by a continuum of firms whose pollution is going to be regulated by either permits or a uniform emission rate standard. In Section 3, I first derive the optimal design for these two regulations and then compare the welfare difference between the two optimal designs. In Section 4, I apply the theoretical model to data from the Santiago's TSP permits program. Concluding remarks are in Section 5.

2 The model

Consider a competitive market for an homogeneous good supplied by a continuum of firms of mass 1. Each firm produces output q and emissions e of a uniform flow pollutant. To

[8] There are very few cases, which are unlikely to hold in practice, in which it may be optimal to just rely on standards and totally abstract from permits. See Montero (2004) for more.

simplify notation, I assume that when the firm does not utilize any pollution abatement device $e = (1+\alpha)q$, so $1+\alpha$ is the firm's emissions rate in the absence of regulation, which the evidence shows that greatly vary across firms (it does not have any implications whether the regulator knows α or not).

A firm can abate pollution at a positive cost by installing technology x, which reduces emissions from $(1+\alpha)q$ to $e = (1+\alpha-x)q$. Hence, the firm's emission rate is $r \equiv e/q = 1+\alpha-x$, which is observed by the regulator during his inspection visits. I assume that there is full compliance with either regulatory instrument. In addition to α, each firm is represented by a pair of cost parameters (β, γ). A firm of type (α, β, γ) has a cost function $C(q, x, \beta, \gamma)$ where β and γ are firm's private information. To keep the model mathematically tractable, I assume that the cost function has the following quadratic form in the relevant output-abatement range[9]

$$C(q, x, \beta, \gamma) = \frac{c}{2}q^2 + \beta q + \frac{k}{2}x^2 + \gamma x + vxq \tag{1}$$

where c, k and v are publicly known parameters common to all firms and $c > 0$, $k > 0$ and $v \gtreqless 0$.[10] Although α does not directly enter into the cost function, it can be indirectly related to costs through its correlation with β and γ, capturing, for example, that a firm with high counterfactual emissions (i.e., high α) is likely to find it cheaper to reduce emissions (i.e., low γ).

Function (1) incorporates two key cost parameters that are essential to understand firms' behavior under permits and standards regulation. One of these cost parameters is the correlation between β and γ (that we shall denote by $\rho_{\beta\gamma}$), which captures whether firms with higher output ex-ante (i.e., before the regulation) are more or less likely to install more abatement

[9] This approach was first introduced by Weitzman (1974).

[10] The parameter v can be negative, for example, if switching to a cleaner fuel saves on fuel costs but involves such a large retrofitting cost (i.e., high k) that no firm switches to the cleaner and cheaper fuel unless regulated.

x. The other cost parameter is v, which captures the effect of abatement on output ex-post (note that we have constrained v to be the same for all firms, thus, a negative value of v would indicate that, on average, the larger the x the larger the increase in q ex-post). As we shall see, the values of the cost parameters v and $\rho_{\beta\gamma}$ play a fundamental role in the design and choice of policy instruments when emissions are not closely monitored.

Although the regulator does not observe firms' individual values for α, β and γ (but observes r), I assume that he knows that they are distributed according to the cumulative joint distribution $F(\alpha, \beta, \gamma)$ on $\alpha \in [\underline{\alpha}, \overline{\alpha}]$, $\beta \in [\underline{\beta}, \overline{\beta}]$ and $\gamma \in [\underline{\gamma}, \overline{\gamma}]$.[11] To simplify notation further and without any loss of generality I let Exp[α] =Exp[β] =Exp[γ] = 0, where Exp[$\cdot$] is the expected value operator.[12]

Market (inverse) demand is totally elastic and given by $P(Q) = P$, where Q is total output. Total damage from pollution is a linear function given by $D(E) = hE$, where E are total emissions and $h > v$. Functions $P(Q)$ and $D(E)$ are known to the regulator. Firms behave competitively, taking the output clearing price P as given. Hence, in the absence of any environmental regulation, each firm will produce to the point where its marginal production cost equals the product price (i.e., $C_q(q, x, \beta, \gamma) = P$), and install no abatement technology (i.e., $x = 0$). Because production involves some pollution, this market equilibrium is not socially optimal. The regulator's problem is then to design a regulation that maximizes social welfare.

[11]Note that we can easily add aggregate uncertainty to this formulation by simply letting $\beta^i = \beta^i + \theta$ and $\gamma^i = \gamma^i + \eta$, where θ and η are random variables common to all firms.

[12]Note that because β and γ are negative for some firms, one can argue that marginal costs can take negative values. This possibility is eliminated by assuming parameter values (including those in the demand and damage functions) that lead to interior solutions for q and x in which $\partial C/\partial q > 0$ and $\partial C/\partial x > 0$ for all β and γ. Furthermore, since these interior solution are assumed to fall within the range in which (1) is valid, what happens beyond this range is not relevant for the analysis of instrument design and choice that follows. Alternatively, one can let $\beta \in [0, \overline{\beta}]$ and $\gamma \in [0, \overline{\gamma}]$ with some further notation in the optimal designs but no change in the welfare comparisons.

I let the benevolent regulator's social welfare function be

$$W = PQ - \int_{\underline{\alpha}}^{\overline{\alpha}} \int_{\underline{\beta}}^{\overline{\beta}} \int_{\underline{\gamma}}^{\overline{\gamma}} C(q, x, \beta, \gamma) F_{\gamma\beta\alpha} d\gamma d\beta d\alpha - hE \tag{2}$$

where $Q = \int_\alpha \int_\beta \int_\gamma q(\beta, \gamma) F_{\gamma\beta\alpha} d\gamma d\beta d\alpha$ is total output and $E = \int_\alpha \int_\beta \int_\gamma r(\alpha, \beta, \gamma) q(\beta, \gamma) F_{\gamma\beta\alpha} d\gamma d\beta d\alpha$ is total emissions (with $r(\alpha, \beta, \gamma) = 1 + \alpha - x(\beta, \gamma)$). In this welfare function, the regulator does not differentiate between consumer and producer surplus and transfers from or to firms are lump-sum transfers between consumers and firms with no welfare effects.[13]

The regulator's problem then becomes to maximize (2) subject to different information constraints and to the restriction that he can use one of two regulatory instruments: standards or permits.[14] It should be mentioned that I focus on these two (second-best) policies and not on more optimal ones not only because the latter include the use of nonlinear instruments and transfers to firms which has not been used in practice (Stavins, 2003; Hahn et al., 2003), but more importantly, because I want to specifically explore whether permits can still provide an important welfare advantage over traditional CAC regulation when emissions are imperfectly monitored.

3 Instrument design and choice

The regulator faces a sequential instrument design and choice problem. Given the information that he has at hand, he must first derive the optimal designs for standards and permits and then determine which of the two optimal designs lead to higher welfare W. Proceeding in this same order, this section develops the solution to the regulators' overall problem.

[13]The model can be generalized by allowing the regulator to consider a weight $\mu \neq 1$ for firm profits and a shadow cost $\lambda > 0$ for public funds. However, this would not add much to our discussion.

[14]Montero (2004) derives the optimal hybrid policy that optimally combines permits and standards. In many cases, this optimal hybrid policy converges to the permits-alone policy and in others, although very few, to the standards-alone policy.

3.1 Standards design

The regulator's problem here is to find the emission rate standard r_s to be required to all firms that maximizes social welfare (subscript "s" denotes standards policy). The regulator knows that for any given r_s, firm (α, β, γ) will maximize $\pi(q, x_s, \alpha, \beta, \gamma) = Pq - C(q, x_s, \beta, \gamma)$ subject to $r = 1+\alpha-x \leq r_s$. Assuming interior solutions throughout, we know that no firm will reduce emissions beyond the standard r_s, so

$$x_s(r_s) \equiv x_s = 1 + \alpha - r_s \tag{3}$$

In turn, firm's (α, β, γ) output decision will solve the first-order condition

$$P - cq - \beta - vx_s = 0$$

which provides the regulator with firm's output q as a function of the standard r_s

$$q_s(r_s) \equiv q_s = \frac{P - \beta - v \cdot (1 + \alpha - r_s)}{c} \tag{4}$$

Using the welfare function (2), the regulator now solves

$$\max_{r_s} \int_{\underline{\alpha}}^{\overline{\alpha}} \int_{\underline{\beta}}^{\overline{\beta}} \int_{\underline{\gamma}}^{\overline{\gamma}} [Pq_s(r_s) - C(q_s(r_s), x_s(r_s)) - r_s q_s(r_s)h] \, F_{\gamma\beta\alpha} d\gamma d\beta d\alpha$$

where $x_s(\cdot)$ and $q_s(\cdot)$ are given by (3) and (4), respectively. By the envelope theorem, the regulator's first-order condition is

$$\int_{\underline{\alpha}}^{\overline{\alpha}} \int_{\underline{\beta}}^{\overline{\beta}} \int_{\underline{\gamma}}^{\overline{\gamma}} \left[(-kx_s - \gamma - vq_s) \frac{\partial x_s}{\partial r_s} - r_s h \frac{\partial q_s}{\partial r_s} - q_s h \right] F_{\gamma\beta\alpha} d\gamma d\beta d\alpha = 0 \tag{5}$$

By replacing $\partial x_s/\partial r_s = -1$, (4) and $\partial q_s/\partial r_s = v/c$ into (5), the first-order condition (5) reduces to

$$ck \cdot (1 - r_s) + v \cdot (P - v + vr_s) - hr_s v - h \cdot (P - v + vr_s) = 0$$

which leads to the optimal standard[15]

$$r_s = \frac{\Lambda + hv - P \cdot (h - v)}{\Lambda + 2hv} = 1 - \frac{P \cdot (h - v) + hv}{\Lambda + 2hv} < 1 + \underline{\alpha} \tag{6}$$

where $\Lambda \equiv ck - v^2 > 0$. Comparative statics can be easily illustrated for $v = 0$, in which case $r_s = 1 - Ph/ck$. As expected, the optimal standard becomes tighter (i.e., lower) as marginal damages increase (i.e., higher h) and loosen as marginal (production and abatement) costs shift up (i.e., higher c and k). It is perhaps less obvious that the optimal standard decreases with the output price P. The reason is that an increase in P stimulates more output and higher emissions, which makes it optimal to tighten the standard all else equal.

3.2 Permits design

Since the regulator only observes the firm's emissions rate r, the permits scheme is not based on actual emissions e but on some proxy for emissions that we denote by $\widetilde{e}$. The regulator's problem is then to find the total number permits $\widetilde{e}_0$ to be distributed among firms that maximizes social welfare. Let R denote the equilibrium price of permits, which will be determined shortly.[16] The regulator knows that firm (α, β, γ) will take R as given and solve

$$\max_{q,x} \pi(q, x, \beta, \gamma) = Pq - C(q, x, \beta, \gamma) - R \cdot (\widetilde{e} - \widetilde{e}_0)$$

[15] Note that the second-order condition imposes $-ck + v^2 - 2hv < 0$.

[16] Note that under a tax policy, the optimal price R will be the tax. If we add aggregate uncertainty to the model, the two policies will not be equivalent from an efficiency standpoint (Weitzman, 1974).

where $\widetilde{e} = (1+\alpha-x)\widetilde{q}$ are firm's proxied emissions and $\widetilde{q}$ is some arbitrarily output or utilization level that is common to all firms. For example, $\widetilde{q}$ could be set equal to the maximum possible output that could ever be observed, which would occur when $x = 0$ and $\beta = \underline{\beta}$. As we shall see later, the exact value of $\widetilde{q}$ turns out to be irrelevant because it simply works as a scaling factor. Note that if $\widetilde{e} < \widetilde{e}_0$ the firm will be a seller of permits.

From firms' first-order conditions

$$x: \quad -kx - \gamma - vq + R\widetilde{q} = 0 \tag{7}$$

$$q: \quad P - cq - \beta - vx = 0 \tag{8}$$

we have that firm's (α, β, γ) optimal abatement and output responses to R and $\widetilde{q}$ (or, more precisely, to $R\widetilde{q}$) are

$$x_p = \frac{R\widetilde{q}c - \gamma c - (P-\beta)v}{\Lambda} \tag{9}$$

$$q_p = \frac{P - \beta - vx_p}{c} \tag{10}$$

where the subscript "p" denotes permits policy (recall that the firm's rate will be $r_p = 1+\alpha-x_p$).

We can now solve the regulator's problem of finding the optimal $\widetilde{e}_0$. Since the market clearing condition is

$$\int_{\underline{\alpha}}^{\overline{\alpha}} \int_{\underline{\beta}}^{\overline{\beta}} \int_{\underline{\gamma}}^{\overline{\gamma}} \widetilde{e} F_{\gamma\beta\alpha} d\gamma d\beta d\alpha = \int_{\underline{\alpha}}^{\overline{\alpha}} \int_{\underline{\beta}}^{\overline{\beta}} \int_{\underline{\gamma}}^{\overline{\gamma}} (1 + \alpha - x_p)\widetilde{q} F_{\gamma\beta\alpha} d\gamma d\beta d\alpha = \widetilde{e}_0 \tag{11}$$

and x_p is a function of $R\widetilde{q}$ as indicated by (9), it is irrelevant whether we solve for $R\widetilde{q}$ or $\widetilde{e}_0/\widetilde{q}$. Hence, we let the regulator to find $R\widetilde{q}$ as to maximize (permits purchases and sales are transfers with no net welfare effects)

$$\int_{\underline{\alpha}}^{\overline{\alpha}} \int_{\underline{\beta}}^{\overline{\beta}} \int_{\underline{\gamma}}^{\overline{\gamma}} \left[P q_p(x_p(R\widetilde{q})) - C(q_p(x_p(R\widetilde{q})), x_p(R\widetilde{q})) - (1+\alpha - x_p(R\widetilde{q})) q_p(x_p(R\widetilde{q})) h \right] F_{\gamma\beta\alpha} d\gamma d\beta d\alpha$$

By the envelope theorem, the first-order condition is

$$\int_{\underline{\alpha}}^{\overline{\alpha}} \int_{\underline{\beta}}^{\overline{\beta}} \int_{\underline{\gamma}}^{\overline{\gamma}} \left[-(1+\alpha - x_p) h \frac{\partial q_p}{\partial (R\widetilde{q})} + q_p h \frac{\partial x_p}{\partial (R\widetilde{q})} - R\widetilde{q} \frac{\partial x_p}{\partial (R\widetilde{q})} \right] F_{\gamma\beta\alpha} d\gamma d\beta d\alpha = 0 \qquad (12)$$

By plugging $\partial q_p / \partial (R\widetilde{q}) = [\partial q_p / \partial x_p][\partial x_p / \partial (R\widetilde{q})]$, $\partial q_p / \partial x_p = -v/c$, (9) and (10) into (12), the first-order condition can be rearranged to obtain the optimal permits price

$$R\widetilde{q} = \frac{Ph(kc + v^2) + hv\Lambda}{(\Lambda + 2hv)c} \qquad (13)$$

which, in turn, allows us to obtain the optimal permits allocation $\widetilde{e}_0 / \widetilde{q}$ by simply replacing (13) in (9) and that in (11).

We can now replace $R\widetilde{q}$ in (9) and (10) to obtain expressions for x_p, r_p and q_p that are more readily comparable to x_s, r_s and q_s (see eqs. (4) and (6)). After some algebra, the following expressions are obtained

$$x_p = 1 - r_s + \frac{v\beta - c\gamma}{\Lambda} = x_s - \alpha + \frac{v\beta - c\gamma}{\Lambda} \qquad (14)$$

$$r_p = r_s + \alpha - \frac{v\beta - c\gamma}{\Lambda} \qquad (15)$$

$$q_p = \frac{P - v \cdot (1 - r_s)}{c} - \frac{k\beta - v\gamma}{\Lambda} = q_s + \frac{v\alpha}{c} - \frac{v^2\beta - cv\gamma}{c\Lambda} \qquad (16)$$

where r_s is the (constant) optimal standard defined by (6). If firms are homogeneous (i.e., $\alpha = \beta = \gamma = 0$ for all firms), it is not surprising that $x_p = x_s$, $r_p = r_s$ and $q_p = q_s$ and that both regulations provide the same welfare. As firms become heterogenous, x, r and q move in different magnitude and sometimes direction depending on the policy choice, which will ultimately affect the welfare comparison between the two policies. Suppose, for example, that $v > 0$. As firms differ on their abatement costs γ, emission rates r increase with γ under the permits regulation while they remain constant under CAC regulation. Thus, permits appear more efficient in accommodating abatement cost heterogeneity to abatement decisions. As firms differ on their production costs β, however, emission rates r decrease with β under permits. Hence, standards appear more efficient in accommodating production cost heterogeneity to abatement decisions. A similar pattern can be found from analyzing firms' production decisions to changes in β and γ. I study the implication on instrument choice of these and related issues more formally in the next section.

3.3 The choice between permits and standards

For a regulator that is limited to use permits or standards,[17] the difference in the social welfare between the optimal permits policy and the optimal standards policy is

$$\Delta_{ps} = W_p(\widetilde{e}_0/\widetilde{q}) - W_s(r_s) \tag{17}$$

where $\widetilde{e}_0$ is the optimal number of permits normalized by some $\widetilde{q}$ and r_s is the optimal standard. The normative implication of (17) is that if $\Delta_{ps} > 0$, the regulator should implement the permits policy.

[17] Although by construction a hybrid policy cannot be welfare dominated by either single-instrument policy, Montero (2004) shows that in many cases the hybrid policy converges to a single-instrument policy.

To explore under which conditions this is the case, we write (17) as

$$\Delta_{ps} = \int_{\underline{\alpha}}^{\overline{\alpha}} \int_{\underline{\beta}}^{\overline{\beta}} \int_{\underline{\gamma}}^{\overline{\gamma}} \left[Pq_p - C(q_p, x_p) - r_p q_p h - Pq_s + C(q_s, x_s) + r_s q_s h\right] F_{\gamma\beta\alpha} d\gamma d\beta d\alpha \tag{18}$$

where q_p, x_p (or r_p), q_s and x_s (or r_s) can be expressed according to (14)–(16). Since $Q_p = Q_s$ $= (P - v(1 - r_s))/c$, eq. (18) can be re-written as

$$\Delta_{ps} = \int_{\underline{\alpha}}^{\overline{\alpha}} \int_{\underline{\beta}}^{\overline{\beta}} \int_{\underline{\gamma}}^{\overline{\gamma}} \left[\{C(q_s, x_s) - C(q_p, x_p)\} + \{r_s q_s - r_p q_p\} h\right] F_{\gamma\beta\alpha} d\gamma d\beta d\alpha \tag{19}$$

Recalling that $e = rq$, the first curly bracket of the right hand side of (19) is the difference in costs between the two policies, whereas the second curly bracket is the difference in emissions that multiplied by h gives the difference in pollution damages.

If we plug (14)–(16) into (19), after some algebra, (19) becomes

$$\begin{aligned}\Delta_{ps} = &\frac{\Lambda^2\sigma_\alpha^2 + v^2\sigma_\beta^2 + c^2\sigma_\gamma^2 - 2v\Lambda\rho_{\alpha\beta}\sigma_\alpha\sigma_\beta + 2c\Lambda\rho_{\alpha\gamma}\sigma_\alpha\sigma_\gamma - 2cv\rho_{\beta\gamma}\sigma_\beta\sigma_\gamma}{2c\Lambda} \\ &- h \cdot \frac{vck\sigma_\beta^2 + vc^2\sigma_\gamma^2 - v^2\Lambda\rho_{\alpha\beta}\sigma_\alpha\sigma_\beta + vc\Lambda\rho_{\alpha\gamma}\sigma_\alpha\sigma_\gamma - (kc + v^2)c\rho_{\beta\gamma}\sigma_\beta\sigma_\gamma}{c\Lambda^2}\end{aligned} \tag{20}$$

where σ_i^2 is the variance of i $(= \alpha, \beta, \gamma)$ and $\rho_{ij}\sigma_i\sigma_j$ is the covariance between i and j. Eq. (20) is a long expression whose sign is not readily seen.

To grasp the intuition behind this expression, it is useful to develop an equivalent but simplified version for it. From (15), the variance of firms' emission rates under permits, $r_p(\alpha, \beta, \gamma)$, is

$$\mathrm{Var}[r_p] = \frac{1}{\Lambda^2}(\Lambda^2\sigma_\alpha^2 + v^2\sigma_\beta^2 + c^2\sigma_\gamma^2 - 2v\Lambda\rho_{\alpha\beta}\sigma_\alpha\sigma_\beta + 2c\Lambda\rho_{\alpha\gamma}\sigma_\alpha\sigma_\gamma - 2cv\rho_{\beta\gamma}\sigma_\beta\sigma_\gamma) \tag{21}$$

and from (15) and (16) that the covariance between firms' output and emission rates under permits, $q_p(\alpha, \beta, \gamma)$ and $r_p(\alpha, \beta, \gamma)$, is

$$\begin{aligned} \text{Cov}[q_p, r_p] &\equiv \text{Exp}[q_p r_p] - \text{Exp}[q_p]\text{Exp}[r_p] \\ &= \frac{1}{c\Lambda^2}(ckv\sigma_\beta^2 + cv\sigma_\gamma^2 - \Lambda v^2 \rho_{\alpha\beta}\sigma_\alpha\sigma_\beta + vc\Lambda\rho_{\alpha\gamma}\sigma_\alpha\sigma_\gamma - (kc + v^2)c\rho_{\beta\gamma}\sigma_\beta\sigma_\gamma) \end{aligned} \quad (22)$$

Using (21) and (22), the welfare difference between permits and standards can be conveniently written as

$$\Delta_{ps} = \frac{\Lambda}{2c}\text{Var}[r_p] - h\text{Cov}[q_p, r_p] \quad (23)$$

As in (20), the first term in (23) is the difference in total costs between standards and permits while the second term is the difference in environmental damages (note that $\text{Cov}[q_p, r_p]$ is the difference in aggregate emissions between the two policies, that is, $E_p - E_s = \text{Cov}[q_p, r_p]$). As we shall see in the next section, expression (23) greatly simplifies the empirical comparison of the two policies because this exercise can be just based on data from the existing permits policy with no need for an explicit construction of a hypothetical standards policy.

Expression (23) also facilitates our understanding of the conditions under which the permits policy dominates the standards policy. Unlike standards, permits allow emission rates r_p to vary across firms. Because of this flexibility, firms will, on average, always find it cheaper to comply with permits than with standards. Depending on the degree of heterogeneity across firms (i.e., differences along α, β and γ), this flexibility can result in substantial (production and abatement) cost savings, as indicated by the first term of (23).

The second term of (23) shows, on the other hand, that the same flexibility that allows firms to save on production and abatement costs can sometimes provide these same firms with incentives to choose combinations of output and abatement levels that may lead to higher aggregate emissions than under standards. Since the actual firm's emissions are the product of q_p and r_p, (23) indicates that permits will lead to higher (lower) emissions when the cost structure of firms is such that the permits policy result in a positive (negative) relationship between output q_p and emission rates r_p. A positive relationship, for example, means that those more utilized firms are, on average, doing less abatement. In other words, permits are flowing from low-utilized firms to high-utilized firms.

These results suggest that when the cost heterogeneity across firms is large (such that emission rates r_p vary greatly across firms), the permits policy is likely to work better than the standards policy. In contrast, as heterogeneity disappears, the advantage of permits reduces, and standards might work better provided that they lead to lower emissions. The possibility that permits can result in higher emissions depends to a large extent on the values of two parameters of the cost function: $\rho_{\beta\gamma}$ (the correlation between production costs and abatement costs) and v (the interaction between production and abatement).

In fact, if $v = \rho_{\beta\gamma} = 0$, the second term in (20) vanishes, i.e., aggregate emissions are the same under either instrument. Provided that firms' output are, on average, the same under either policy (see eq. (16)), when there is neither correlation nor interaction between production and abatement, any given firm is equally likely to emit as much as under permits than under standards. Conversely, if $v = 0$ but $\rho_{\beta\gamma} \neq 0$, the second term in (20) reduces to $h\rho_{\beta\gamma}\sigma_\beta\sigma_\gamma$. In particular, when there is a negative correlation between production and abatement costs (i.e., $\rho_{\beta\gamma} < 0$), aggregate emissions are higher under permits because permits induce primarily low-output firms to install abatement technologies while standards force all firms to invest in abatement more or less equally.

Similarly, if $\rho_{\beta\gamma} = 0$ but $v \neq 0$, the second term in (20) is likely to be different from zero. In particular, if $\rho_{\alpha\beta} = \rho_{\alpha\gamma} = \rho_{\beta\gamma} = 0$ and $v > 0$, emissions will be larger under permits. In this case, when firms doing more abatement find it optimal to reduce output ex-post (i.e., $v > 0$), the permits policy has the disadvantage of reducing the output of firms doing more abatement relative to the output of those doing less abatement. This problem is less significant under standards because all firms are required to install similar abatement technologies.

As the different parameters values are likely to vary from case to case, there will be cases in which standards are the correct policy choice and others in which permits are the correct choice.[18] It could be argued, however, that because permits are always less costly than standards and may or may not lead to higher emissions, permits should be adopted as a default, unless the available cost and pollution damage information indicates the opposite. In other words, the burden of proof should lie with the CAC policy and not with the permits policy.

4 An empirical evaluation

The theoretical analysis indicates that whether permits welfare dominate standards when emissions are imperfectly observed is ultimately an empirical question. In this section, I use the experience from Santiago's total suspended particulate (TSP) permits program to evaluate the advantages, if any, of using permits for regulating the emissions of the group sources affected by the TSP program. Because firms are not required to provide the regulator with information on production and abatement costs, I apply the theoretical framework previously developed to infer the cost structure of the firms affected by the TSP program and other parameters. These estimates are then used to compare the actual performance of the TSP permits program with the performance of a hypothetically equivalent standards policy.

The empirical evaluation is carried out under several assumptions that deserve explanation. First, I retain the exact structure of the theoretical model that includes, among other things, constant output prices P and constant marginal pollution damages h. While the TSP program

[18] Yet, in other cases the correct choice is to optimally combine permits and standards (Montero, 2004).

is relatively small to affect output prices and total emissions, I retain these assumptions because otherwise I would not be able to estimate the parameters of the model in a relatively simple way. In other words, I use the model as a useful interpretative guide of the data but this does not exclude other alternative interpretations of the data. Second, in recovering key parameters of the model such as P and h and comparing policies, I impose some consistency in the regulator's behavior in that the equivalent standards policy is constructed under the assumption that if the regulator had to introduce a standard he will do it optimally using the same value of h (together with the other parameters) that he implicitly used in implementing the permits policy.[19] This does not imply, however, that the regulator is necessarily implementing a policy based on a value of h supported by scientific evidence.

4.1 The TSP permits program

The city of Santiago has constantly presented air pollution problems since the early 1980s. The TSP trading program, established in March of 1992, was designed to curb TSP emissions from the largest stationary sources in Santiago (industrial boilers, industrial ovens, and large residential and commercial heaters) whose emissions are discharged through a duct or stack at a flow rate greater than or equal to 1,000 m^3/hr. Because sources were too small to require sophisticated monitoring procedures, the authority did not design the program based on sources' actual emissions but on a proxy variable equal to the maximum emissions that a source could emit in a given period of time if it operates without interruption.

The proxy for emissions (expressed in kg of TSP per day) used by the authority in this particular program was defined as the product of emissions concentration (in mg/m^3) and flow rate (in m^3/hrs) of the gas exiting the source's stack (multiplied by 24 hrs and 10^{-6}

[19] Allowing for a regulator with objective functions and parameter values that dependen on the instrument under consideration introduces new elements to the policy analysis that go well beyond the scope of this paper.

kg/mg to obtain kg/day).[20] Although the regulatory authority monitors each affected source's concentration and flow rate once a year,[21] emissions $\widetilde{e}$ and permits $\widetilde{e}_0$ are expressed in daily terms to be compatible with the daily TSP air quality standards. Thus, a source that holds one permit has the right to emit a maximum of 1 kg of TSP per day indefinitely over the lifetime of the program.

Sources registered and operating by March 1992 were designated as existing sources and received grandfathered permits equal to the product of an emissions rate of 56 mg/m^3 and their flow rate at the moment of registration. New sources, on the other hand, receive no permits, so must cover all their emissions with permits bought from existing sources. The total number of permits distributed (i.e., the emissions cap) was 64% of aggregate (proxied) emissions from existing sources prior to the program. After each annual inspection, the authority proceeds to reconcile the estimated emissions with the number of permits held by each source (all permits are traded at a 1:1 ratio). Note that despite the fact that permits are expressed in daily terms, the monitoring frequency restricts sources to trade permits only on an annual or permanent basis.[22]

4.2 The data

The data for the study were obtained from PROCEFF's databases for the years 1993 through 1999.[23] Each database includes information on the number of sources and their dates of registration, flow rates, fuel types, emission rates and utilization (i.e., days and hours of operation

[20]In terms of our model, this is equivalent as to make $\widetilde{q}$ equal to the maximum possible output, which in our case is $(P - \underline{\beta})/c$. But note that the program would have worked equally well with an either higher or lower $\widetilde{q}$. The use of a different $\widetilde{q}$ only requires to adjust the number of quasi-permits $\widetilde{e}_0$ to be distributed such that $R\widetilde{q}$ remains at its optimal level.

[21]There are also random inspections to enforce compliance throughout the year.

[22]In addition, the authority introduced an emission rate standard of 112 mg/m^3 for all stationary sources. It seems, however, that this either was only enforced by 1998 or became non-binding after the arrival of natural gas.

[23]PROCEFF is the government office responsible for enforcing the TSP program.

during the year). While information on flow rates, fuel types and emission rates is directly obtained by the authority during its annual inspections, information on utilization is obtained from firms' voluntary reports.[24] The 1993 database contains all the information, including the flow rate used to calculate each source's allocation of permits, before the program became effective in 1994. Table 1 presents a summary of the data. The first two rows show the proportion of existing and new sources.[25]

[INSERT TABLE 1 HERE OR BELOW]

The next rows of Table 1 provide information on the evolution of flow rates, emission rates and utilization. The large standard deviations indicate that these three variables vary widely across sources in all years.[26] In order to comply with the TSP trading program, affected sources can hold permits, reduce emissions or do both. They can reduce emissions either by decreasing their size (i.e., flow rate) or by decreasing their emission rates. The latter can be done through either fuel switching (for example, from wood, coal, or heavy oil to light oil, liquid gas, or natural gas) or the installation of end-of-pipe technology (e.g., filters, electrostatic precipitators, cyclones, and scrubbers).[27] Sources do not gain anything, in terms of emissions reduction, by changing their utilization level (i.e., days and hours of operation), because by definition it is assumed to be at 100%. Given that the authority controls for the size of the source (i.e., flow rate) at the moment of permits allocation and monitoring, emission rates and utilization are captured, respectively, by r_p and q_p in our theoretical model.

[24] Since utilization has no effect at all on the source's compliance status, there is no reason to beleive that firms have incentives to misreport their true utilization. For the same reason, this information is available for most but not all sources.

[25] It is interesting to point out that by 1999, 36% of the affected sources were new sources despite the fact they did not receive any permits.

[26] It may seem strange to observe some flow rates below the 1,000 (m^3/hr) mark. In general, these are existing sources for which flow rates were wrongly estimated to be above 1,000 (m^3/h) at the time of registration. Nevertheless, these sources chose to remain in the program to keep the permits they had already received.

[27] Note that for most sources, flow rates do not change over time.

The last two rows of Table 1 show data on total emissions and permits.[28] Although 1994 was in principle the first year of compliance with the program, trading activity did not occur until 1996 when compliance was more effectively enforced (Montero et al, 2002). The emissions goal of the TSP program was only achieved by 1997 (total emissions below total permits).[29] This is the year after in which natural gas became available from Argentina at unexpectedly attractive prices, such that many affected sources switched to this cleaner fuel, leaving the cap of 4,087.5 permits largely non-binding.[30] Consequently, the empirical evaluation that follows is mainly based on the 1997 data and to a lesser extent on the 1998 data.

4.3 Estimation of parameters and Δ_{ps}

Based on (23), the sign of Δ_{ps} can be first explored by looking at the covariance matrix for the emission rate (r_p) and utilization (q_p). Using the flow rate as a weight to control for size differences across sources, the weighted statistics for the 1997 data (499 observations) are $\text{Var}[q_p] = 0.112$, $\text{Var}[r_p] = 0.211$ and $\text{Cov}[r_p, q_p] = 0.026$ (to work with dimensionless variables hereafter, emission rates are divided by their 1993 mean value of 94.9 mg/m^3)[31] and for the 1998 data (543 observations) the weighted statistics are, respectively, 0.111, 0.056, and 0.005. Although these figures do not allows us to sign Δ_{ps} yet, they indicate that emissions have been somewhat larger than what would have been under an equivalent standards policy. Since $E_p = \text{Exp}[r_p q_p]$ and the weighted value of $\text{Exp}[r_p q_p]$ in 1997 is 0.445, emissions would have been

[28] A few permits were retired from the market in 1997 as the authority revised the eligibility of some sources to receiving permits (Montero et al., 2002).

[29] The fact that total emissions in 1997 are somewhat below the cap should not be interpreted as either overcompliance or non-binding regulation. One explanation is that firms tend to hold a few extra permits as an insurance against some measurement uncertainty (inherent to a monitoring precedure of this sort). A second explanation is the uncertainty associated with revision of the initial allocation of permits carried out by the authority after the beginning of the program. The 1997 allocation drop is, in fact, the result of such a revision.

[30] This is consistent with the fact that inter-firm trading activity stopped by mid 1998. Obviously, intra-firm trading activity has continued as new sources come into operation.

[31] The unweighted statistics are, respectively, 0.101, 0.221, and 0.004.

0.419 under the equivalent standard of 0.663 (the latter is the weighted value of $\text{Exp}[r_p]$ in 1997).

The 1997 figures also show that $\text{Var}[r_p]$ is more than eight times larger than $\text{Cov}[r_p, q_p]$, raising the possibility that the higher emissions may be more than offset by cost savings. To test for this possibility, however, more information on various parameters is required.

A more precise estimate of Δ_{ps} requires then values of v, c, k, h and P. This information is to be recovered from the data described in Table 1 (no detailed information on production and abatement costs is available elsewhere; at least to my knowledge). I start with the estimation of v. Based on first order conditions (7) and (8), v is obtained by estimating the following simultaneous-equation system

$$\begin{aligned} REDUC_i = a_0 + a_1 UTIL_i + a_2 FLOW93_i + a_3 EMRTE93_i \\ + a_4 ENDPIPE_i + a_5 INDUST_i + a_6 STATE_i + \varepsilon_i \end{aligned} \quad (24)$$

$$\begin{aligned} UTIL_i = b_0 + b_1 REDUC_i + b_2 UTIL93_i + b_3 FLOW93_i \\ + b_4 INDUST_i + b_5 STATE_i + u_i \end{aligned} \quad (25)$$

where i indexes sources, ε^i and u^i are error terms whose characteristics will be discussed shortly, and the different variables relate to those in (7)–(8) as follows. $REDUC$ corresponds to x_p, i.e., the level of reduction under the permits policy. $REDUC$ is calculated as the difference between the source's counterfactual emission rate $(1 + \alpha)$ and its actual emission rate (r_p).[32] I use the 1993 as the counterfactual year,[33] so $EMRTE93$ is the counterfactual emissions rate.

[32] Recall that emission rates are normalized by the 1993 mean.

[33] Results do not qualitatively change when I use 1995 as the counterfactual year (the year in which I have a few more data points).

The variable $UTIL$ corresponds to q_p, i.e., the level of utilization or output. As in the theoretical model, the TSP program's authority does not observe $UTIL$, and therefore, he cannot use it for monitoring and enforcement purposes. To put it differently, because the regulator only observes a source's flow rate and emissions rate, he only has control over changes in emissions due to changes in the source's size (i.e., flow rate) and emission rates but not over changes in emissions due to changes in utilization.

The variables $FLOW93$, $EMRTE93$, $ENDPIPE$, $INDUST$ and $STATE$ included in (24) are intended to capture differences in abatement costs across sources (i.e., γ).[34] $FLOW93$ is the source's flow rate in 1993. If there are any scale economies associated with pollution abatement, then we should expect more abatement from bigger sources (i.e., larger $FLOW93$), other things equal (I also use $FLOW93^2$ and $\ln FLOW93$).[35] Similarly, I expect a source with a high emissions rate before the TSP program (i.e., high $EMRTE93$) to face more abatement possibilities and hence lower costs. Conversely, I expect a source already equipped with some end-of-pipe abatement technology required by previous (and source specific) regulation to be less likely to reduce emissions. Hence, I introduce the dummy variable $ENDPIPE$ that equals 1 if the source has any type of end-of-pipe abatement technology by 1993. I also introduce the dummy variables $INDUST$ and $STATE$ to see whether there is any difference in abatement costs (or abatement behavior) between industrial sources ($INDUST = 1$) and residential/commercial sources, or between state or municipality owned sources ($STATE = 1$) and privately owned sources.[36]

[34] Since sources' emissions were unregulated by 1993 (except for a very few sources that were required to install end-of-pipe abatement technology before 1993), there is no reason to beleive that a sources' utilization in the absence of emission control ($UTIL93$) could tell anything about how easy or difficult is for the source to reduce emissions.

[35] I use the 1993 flow rate instead of the actual flow rate to control for possible endogeneity problems. However, results are virtually the same when I use the actual flow rate. This is in part because the firm's flow rate barely change over time (the drop in average flow rates shown in Table 1 is mainly due to changes in one particular large firm).

[36] For example, $INDUST = 0$ and $STATE = 1$ for the boiler of the central heating system of a public hospital.

The variables $UTIL93$, $FLOWRTE93$, $INDUST$ and $STATE$ included in (25) are intended to capture differences in production costs across sources (i.e., β).[37] $UTIL93$ is the source's utilization in 1993 and serves as a proxy for the level of utilization that would have been observed in the absence of the TSP program and of changes in exogenous factors (e.g., input prices, demand, etc.).[38] Since, on average, utilization has been increasing over time, $FLOW93$ should capture whether expansion in larger units is relatively cheaper than in smaller units. For the same reason, I also include $INDUST$ and $STATE$.

An estimate of the sign (and relative value) of v can then be inferred from either $a_1 = -v/k$ or $b_1 = -v/c$. Since $UTIL$ and $REDUC$ enter as endogenous variables in (24)–(25), however, their correlations with the error terms ε_i and u_i would produce biased OLS estimators. Therefore, I employ a two-stage least squares (2SLS) estimation procedure to obtain unbiased estimates. 2SLS results for equations (24) and (25) are presented in Table 2 (first-stage results are omitted).

[INSERT TABLE 2 HERE OR BELOW]

The first three columns of Table 2 show the results for the 1997 data. Results in column (1) indicate that the coefficients of $UTIL$ and $REDUC$ (i.e., a_1 and b_1, respectively), although positive, are not significantly different from zero. Because our theoretical model assumes that all firms are expected to produce, on average, the same amount of output ($\text{Exp}[q_p] = (P - vx_s)/c$), however, one can argue that these coefficients may provide a biased estimation of v by not taking into account the fact that firms have different sizes. One could further argue that the true value of v may even be of different sign, because the coefficients of $FLOW93$ and $FLOW93^2$ in the

[37] Note that the variables $EMRTE93$ and $ENDPIPE$ are excluded from the utilization regression. There is nothing particular about $ENPIPE$ and $EMRTE93$ that can affect utilization beyond its effect, if any, in 1993, which is already captured by $UTIL93$.

[38] To work with a larger dataset I use the 1995 utilization for 66 sources. This should not baised the results in any particular way since the TSP program was not effectively enforced until 1996 (see Table 1).

reduction equation indicate that the amount of reduction decreases with size throughout the relevant range. To control for such possibility, I run a weighted 2SLS regression using the 1997 flow rate as weight. The new estimates, which are reported in columns (2) and (3), are not very different from the unweighted estimates, confirming that the interaction term v in equation (1) is not statistically different from zero.

The last three columns of Table 2 show the 2SLS results for the 1998 data [weighted estimates are in columns (5) and (6)]. In particular, we observe that the coefficients of $UTIL$ and $REDUC$ in column (5) are positive and significantly different from zero at the 10% level. This negative value of v can be attributed in large part to the arrival of natural gas at relatively low prices by the end of 1997.[39] Although the 1998 results must be carefully interpreted because of the apparently slack cap, they are useful to illustrate the estimation of Δ_{ps} when v is different from zero, as we shall see next.

We can finally use the estimated value of v to obtain an estimate for the remaining parameters of the model, and hence, for Δ_{ps}. Following the 1997 econometric results, let us consider first the case in which $v = 0$. When this is the case, we have that $h = cR\widetilde{q}/P$ from (13), $R\widetilde{q} = k\text{Exp}[x_p]$ from (9), and $P = c\text{Exp}[q_p]$ from (10). Replacing the 1997 (weighted) statistics for $\text{Exp}[x_p] = 0.203$ and $\text{Exp}[q_p] = 0.631$ in the expression for h, (23) reduces to

$$\Delta_{ps}|_{97,v=0} = \frac{k}{2}\text{Var}[r_p] - \frac{k\text{Exp}[x_p]}{\text{Exp}[q_p]}\text{Cov}[r_p, q_p] = (0.1055 - 0.0084)k = 0.097k > 0$$

These numbers not only indicate that the permits policy is welfare superior to an equivalent standards policy, but, more importantly, that the welfare loss from higher emissions is only 8% of the welfare gain from cost savings.

[39] In fact, 112 the 144 affected sources that switched to natural gas in 1998 increased or maintained their utilization relative to 1997.

Based on the 1998 results contained in column (5) of Table 2, let us now consider the case in which $v < 0$. From the coefficients of $UTIL$ and $REDUC$ we obtain, respectively, $k = -1.86v$ and $c = -15.87v$ (which in turn, yields $\Lambda = 28.52v^2$). In addition, by simultaneously solving (9) and (10) for P and $R\widetilde{q}$ with $\text{Exp}[x_p] = 0.466$ and $\text{Exp}[q_p] = 0.669$, we get $P = -10.15v$ and $R\widetilde{q} = -0.20v$ that replaced into (13) gives $h = -0.31v$. Plugging these numbers and the corresponding statistics into (23), we finally obtain $\Delta_{ps}|_{98,v<0} = (0.0503 - 0.0016)(-v) = -0.049v > 0$. This result, while qualitatively similar to the 1997 result, shows an even smaller welfare loss from higher emissions —only 3% of the welfare gain from cost savings.

5 Final remarks

I have developed a model to study the design and performance of pollution markets (i.e., tradable permits) when the regulator has imperfect information on firms' emissions and costs. A salient example is the control of air pollution in large cities where emissions come from many small (stationary and mobile) sources for which continuous monitoring is prohibitively costly. In such a case the well known superiority of permits over the traditional command and control approach of setting technology and emission standards is no longer evident. Since the regulator only observes a firm's abatement technology but neither its emissions nor its output (utilization), permits could result in higher emissions if firms doing more abatement are at the same time reducing output relative to other firms and/or if more highly utilized firms find it optimal to abate relatively less. I then used emissions and output data from Santiago-Chile's TSP permits program to explore the implications of the theoretical model. I found that the production and abatement cost characteristics of the sources affected by the TSP program are such that the permits policy is welfare superior. The estimated cost savings are only partially offset (about 8%) by a moderate increase in emissions relative to what would have been observed under an equivalent standards policy.

Since sources under the TSP program are currently responsible for less than 5% of total TSP emissions in Santiago, the model developed here can be used to study how to expand the TSP program to other sources of TSP that today are subject to command and control regulation. A good candidate would be powered-diesel buses which are responsible for 36.7% of total TSP emissions. According to Cifuentes (1999), buses that abate emissions by switching to natural gas are likely to reduce utilization relative to buses that stay on diesel and that older, less-utilized buses are more likely to switch to natural gas. Since switching to natural gas is a major abatement alternative, both of these observations would suggest that the optimal way to integrate buses into the TSP program is by imposing, in addition to the allocation of permits, an emission standard specific to buses. It may also be optimal to use different utilization factors ($\widetilde{q}$) for each type of source (see Falk, 2003). These and related design issues deserve further research.

References

[1] Cifuentes, L. (1999), Costos y beneficios de introducir gas natural en el transporte público en Santiago, mimeo, Catholic University of Chile.

[2] Cremer, H., and F. Gahvari (2002), Imperfect observability of emissions and second-best emission and output taxes, *Journal of Public Economics* 85, 385-407.

[3] Ellerman, A.D, P. Joskow, R. Schmalensee, J.-P. Montero and E.M. Bailey (2000), *Markets for Clean Air: The U.S. Acid Rain Program*, Cambridge University Press, Cambridge, UK.

[4] Falk, D. (2003), *Permisos Transable con Monitoreo Imperfecto*, Master thesis, Economics Department, Catholic University of Chile.

[5] Fullerton, D. and S.E. West (2002), Can taxes on cars and on gasoline mimic an unavailable tax on emissions?, *Journal of Environmental Economics and Management* 43, 135-157.

[6] Hahn, R.W., S.M Olmstead and R.N. Stavins (2003), Environmental regulation in the 1990s: A retrospective analysis, *Harvard Environmental Law Review* 27, 377-415.

[7] Harrison, David Jr. (1999), Turning theory into practice for emissions trading in the Los Angeles air basin, in Steve Sorrell and Jim Skea (eds), *Pollution for Sale: Emissions Trading and Joint Implementation*, Edward Elgar, Cheltenham, UK.

[8] Lewis, T. (1996), Protecting the environment when costs and benefits are privately known, *RAND Journal of Economics* 27, 819-847.

[9] Montero, J.-P. (2004), Pollution markets with imperfectly observed emissions, *RAND Journal of Economics*, forthcoming.

[10] Montero, J.-P., J.M. Sánchez, R. Katz (2002), A market-based environmental policy experiment in Chile, *Journal of Law and Economics* XLV, 267-287.

[11] Schmalensee, R, P. Joskow, D. Ellerman, J.P. Montero, and E. Bailey (1998), An interim evaluation of sulfur dioxide emissions trading, *Journal of Economic Perspectives* 12, 53-68.

[12] Segerson, K. (1988), Uncertainty and incentives for nonpoint pollution control, *Journal of Environmental Economics and Management* 15, 87-98.

[13] Stavins, R. (2004), survey paper in this book.

[14] Stavins, R. (2003), Experience with market-based environmental policy instruments, in *Handbook of Environmental Economics*, Karl-Göran Mäler and Jeffrey Vincent eds., Amsterdam: Elsevier Science.

[15] Tietenberg, T. (2004), survey paper in this book.

[16] U.S. Environmental Protection Agency (USEPA, 2001), *The United States Experience with Economic Incentives for Protecting the Environment*, EPA-20-R-01-001, Washington, D.C.

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

TABLE 1. Summary statistics for all affected sources: 1993–1999.

Variable	1993	1995	1996	1997	1998	1999
No. of sources						
Existing	635	578	504	430	365	365
New	45	112	127	146	221	208
Total Affected	680	690	631	576	566	573
Flow rate (m^3/h)						
Average	4,910.7	4,784.1	4,612.6	4,062.1	4,213.9	4,146.6
Standard dev.	15,058.8	14,908.0	15,490.9	9,498.6	13,091.0	11,793.5
Max.	261,383.9	261,304.7	261,304.7	182,843.0	207,110.6	183,739.5
Min.	499.2	204.3	204.3	493.3	216.9	165.6
Emission rate (mg/m^3)						
Average	94.9	83.1	78.5	54.7	31.1	27.8
Standard dev.	88.1	77.8	76.8	43.0	21.1	18.5
Max.	702.0	698.2	674.0	330.7	110.0	108.2
Min.	1.5	1.5	3.4	3.6	2.9	4.6
Utilization (%)*						
Average	39.4	48.0	47.1	49.2	51.7	53.7
Standard dev.	30.3	31.5	31.7	31.8	32.0	32.3
Max.	100	100	100	100	100	100
Min.	0	0	0	0	0	0
No. of observations	278	463	457	499	543	542
Total emissions (kg/day)	7,051.9	6,320.9	5,094.4	3,535.0	1,975.3	1,665.0
Total permits (kg/day)	4,604.1	4,604.1	4,604.1	4,087.5	4,087.5	4,087.5

Source: Elaborated from PROCEFF's databases.

* An utilization of 100% corresponds to 24 hrs of operation during 365 days a year. As indicated by the No. of observations, utilization figures are not based on all sources (recall that information on utilization is not required for monitoring and enforcement purposes).

TABLE 2. 2SLS estimates for the reduction and utilization equations

Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)
Reduction Equation						
UTIL	0.078	0.137	0.087	0.256*	0.539*	0.308
	(0.153)	(0.175)	(0.175)	(0.132)	(0.309)	(0.322)
FLOW93	-0.789**	-0.788***		-0.937***	-1.090***	
	(0.330)	(0.275)		(0.345)	(0.422)	
$FLOW93^2$	0.270**	0.271**		0.346***	0.373**	
	(0.131)	(0.111)		(0.129)	(0.151)	
ln(FLOW93)			-0.088***			-0.093***
			(0.032)			(0.031)
EMRTE93	0.741***	0.717***	0.698***	0.987***	0.944***	0.940***
	(0.094)	(0.115)	(0.116)	(0.019)	(0.039)	(0.035)
ENDPIPE	-0.058	-0.191	-0.027	-0.128	-0.032	0.182
	(0.198)	(0.251)	(0.140)	(0.083)	(0.100)	(0.129)
INDUST	-0.008	0.079	0.120	0.014	-0.031	0.023
	(0.077)	(0.149)	(0.153)	(0.042)	(0.061)	(0.056)
STATE	-0.137	-0.193**	-0.193**	-0.105**	-0.083	-0.118
	(0.106)	(0.084)	(0.082)	(0.050)	(0.077)	(0.074)
Constant	-0.390***	-0.474***	0.217	-0.420***	-0.512***	0.272
	(0.075)	(0.116)	(0.201)	(0.050)	(0.115)	(0.211)
Utilization Equation						
REDUC	-0.003	0.005	-0.017	0.012	0.063*	0.054
	(0.030)	(0.032)	(0.039)	(0.029)	(0.035)	(0.039)
UTIL93	0.560***	0.567***	0.532***	0.364***	0.313***	0.275***
	(0.055)	(0.064)	(0.061)	(0.064)	(0.093)	(0.089)
FLOW93	0.401***	0.384***		0.416	0.417***	
	(0.130)	(0.096)		(0.267)	(0.129)	
$FLOW93^2$	-0.095**	-0.087***		-0.101	-0.098**	
	(0.048)	(0.033)		(0.101)	(0.045)	
ln(FLOW93)			0.078***			0.090***
			(0.012)			(0.011)
INDUST	0.069*	0.038	0.022	0.141***	0.077	0.044
	(0.037)	(0.045)	(0.049)	(0.044)	(0.057)	(0.057)
STATE	-0.077*	-0.042	-0.038	-0.039	-0.117**	-0.098*
	(0.045)	(0.045)	(0.043)	(0.058)	(0.051)	(0.055)
Constant	0.158***	0.182***	-0.407***	0.221***	0.293***	-0.376***
	(0.034)	(0.045)	(0.100)	(0.042)	(0.057)	(0.090)
No. Observations	344	344	344	288	288	288

Notes: First-stage results are omitted. White-corrected standard errors are in parenthesis. Columns (2), (3), (5) and (6) present weighted estimates (the 1997 flow rate is the weight in (2) and (3) and the1998 flow rate in (5) and (6)).
* significant at 10%, ** significant at 5%, *** significant at 1%