Prices *and* Quantities in a Climate Policy Setting

Svante Mandell[†]

vti – Swedish National Road and Transport Research Institute

Abstract

The EU aspires on being a forerunner in adopting a stringent climate policy. The present paper takes its departure in two observations from the current EU policy. First, the EU has adopted a dual approach in that there is a trading scheme, the EU ETS, covering CO_2 emissions from the energy intensive industry, while the remaining emitters, *e.g.*, housing and transportation, are subject to emission taxes. Second, the targets are quantitatively phrased, *i.e.*, there is an upper limit on the total amount of CO_2 emissions. Both these observations are of interest under the realistic assumption of abatement costs being uncertain. Then the dual approach is likely not to be *ex post* cost effective. Furthermore, from earlier literature it is known that an emissions tax outperforms a tradable permit scheme due to the nature of the green house effect. Thus, quantitative targets, which are easier fulfilled by means of a trading scheme, stand in contrast to taxes being the preferable policy instrument. The present paper addresses, by the means of a stylized model, the two observations and shows when and why a dual approach is optimal from an efficiency point of view given an upper limit on total emissions. What determines the characteristics of the optimal solution, *e.g.*, size of the trading *vs.* taxed sector and tax levels, is studied in some detail.

JEL classifications; *H23, L51, Q28, Q58*

Keywords; *Regulation, Emissions tax, Emissions trading, Uncertainty*

[†] Financial support from CTS and the Swedish Energy Administration is gratefully acknowledged. The paper has benefited from suggestions by Björn Carlén. Comments are welcome to Svante.Mandell@vti.se

1. Introduction

In his famous paper, Weitzman (1974) analytically compares quantitative regulations to price based regulations under the presence of uncertainty. The analysis has been modified in several ways and applied to numerous fields, in particular regulations of emissions, and in latter years – due to a will to tackle climate change issues – not least to emissions of CO_2. For CO_2 it is widely understood that a price based instrument, *e.g.*, an emissions tax, is preferred to a quantitative approach, *e.g.*, a cap-and-trade regime, in the sense that the former results in a substantially smaller expected efficiency loss, Hoel and Karp (2001, 2002), Pizer (2002) and Karp and Zhang (2006). Nevertheless, existing policy measures geared towards greenhouse gases in general and CO_2 in particular are to a great extent quantity based. The closest we currently have to a global agreement on greenhouse gas emissions, the Kyoto Protocol, for instance stipulates a quantitative target. There are certain advantages in basing such agreements on quantitative targets. For instance, governments may prefer a quantitative constraint for the certainty of its effect on emissions, Pearce (1991). It is also the case that a cap-and-trade regime provides a simple mean to distribute the rents created by the emissions right in a flexible manner, Pizer (2002). Furthermore, when it comes to international agreements as the Kyoto Protocol, even if all nations managed to agree on a common level of an emissions tax, it is virtually impossible to avoid that some countries respond, *e.g.*, by calibrating other taxes, to counter adverse effects on domestic industry, Bohm (1997). For reasons such as these there may emerge a political constraint under which only agreements including an upper limit on emissions are feasible[1].

The primary purpose of this study is to analyse optimal regulation policy under the presence of a constraint that provides a predefined maximum amount of emissions as discussed above. The upper limit is referred to as a 'global cap', which should be interpreted as an overarching cap that serves as a restriction when designing policy measures[2]. It is global in that it

[1] It should be noted that there are studies arguing for using taxes that address these problems, *e.g.*, Nordhaus (2007) and Cooper (2008).

[2] An alternative application of the present paper is on emissions which are associated with a 'disaster level' at which the impact on the environment will be extremely high. Then the 'global cap' is given by a natural constraint rather than from political negotiations. As arguably the political application contains this application as a subset, the model is formulated with the former in mind.

stipulates an upper limit for aggregate emissions from all emitters participating in the particular agreement. The constraint suggests that a cap-and-trade regime would be preferable, since then total emissions are regulated directly. The question becomes interesting when the nature of the emissions is such that, in the absence of the global cap, an emissions tax would be preferable from an efficiency point of view.

Using quadratic functions and additive uncertainty, Weitzman (1974) shows that an emissions tax outperforms a quantity regime, *i.e.*, cap-and-trade, when the marginal abatement cost (*MAC*) function is steeper than the marginal abatement benefit (*MAB*) function. In a similar framework, Mandell (2008) develops a dual regulation system where only a subset of the emitters is subject to a cap-and-trade regime while the remaining emitters are subject to an emissions tax. It is shown that when the absolute slope of the *MAB* is larger than half that of the *MAC*, this is preferable to having all emitters subject to either cap-and-trade or emissions taxes. This is the case even when taking into account that the approach is not likely to be *ex post* cost effective, since the permit price will most likely differ from the emissions tax. However, if the slope of the *MAB* is sufficiently flat relative to the slope of the *MAC*, the conclusions in Weitzman (1974) and Mandell (2008) are the same; the preferred regulation regime entails subjecting all emitters to a common emissions tax. The combination of a sufficiently flat *MAB*-function and the aforementioned political constraint, which is easiest fulfilled with a cap-and-trade system, creates a trade-off between the two policy instruments. Drawing on the model in Mandell (2008), and thus allowing the policy maker to subject a sub-set of the emitters to an emissions tax and the others to cap-and-trade, we study the optimal share of the emitters to tax and at what rate.

As discussed above, there are three important ingredients in the model; an upper limit on emissions, a flat *MAB* and a possibility to decide on a policy that subject different emitters to different policy instruments. As noted, it is arguably the case that the problem of curbing CO_2 emissions provides a viable point of reference to the analysis. In particular, climate policies adopted by the EU, who actively strives to act as a forerunner in these matters, show many similarities to the subsequent model and results. Not least since the EU has adopted an approach under which roughly half of the CO_2 emissions – those originating from energy intensive industry – is subject to a tradable trading scheme, the EU ETS, while the remaining emissions are handled by the means of emission taxes. However, the analysis is applicable to a wider set of problems than CO_2 alone and, thus, the paper is written in a general manner. We will briefly discuss the results from a CO_2 perspective in the concluding section.

The rest of the paper is structured as follows; section 2 presents the model, section 3 includes the analysis of optimal policies and section 4 concludes.

2. The model

The subsequent model focuses on the aggregate outcome, *i.e.*, we compare the actual aggregate emissions level with the efficient level. Working on an aggregate level, the model does not incorporate strategic issues, for instance associated with negotiations about the global cap.

If a subset of emitters is taxed and there is uncertainty surrounding the *MAC*, total emissions under a given period will be stochastic. Under the presence of the political constraint this implies that actual emissions may fall short of the allowed amount. In line with many emissions regulation systems found in practice, *e.g.*, the Kyoto protocol and the US SO_2 allowance trading program, the model includes a possibility to bank permits and use them at subsequent periods. However, borrowing is not allowed. The model has two periods, which is enough to capture the basic implications of banking.

There are initially two policy variables in the model: what share of the economy to tax and the tax level. At a later stage in the analysis also the size of the cap is treated as a policy variable but initially it is assumed as exogenously given. The only variable that may be calibrated between the two periods, *i.e.*, after the outcome of period 1, is the tax level. This is not very restrictive. We may, for instance, model the case where the size of the taxed sector is to be calibrated after the outcome in period 1 is known by using two subsequent one-period models. Having different caps in period 1 and period 2 will not affect the qualitative results as long as both the period 1 and period 2 caps are declared prior to the first period. However, allowing the policy maker to alter the cap for period 2 after the outcome in period 1 has become known will have crucial impact on the result. It then seems likely that the policy maker would set a low cap if banking occurred as high. The emitters, knowing this, would probably adjust their behaviour in period 1 such that emissions would be higher. As the model does not capture strategic behaviour of this kind, such a setting will not be analysed in the present paper. For similar reasons, the policy maker is not allowed to destroy emissions allowances even if it from an efficiency point of view would be optimal to do so.

It is assumed that the market is perfectly competitive and that transaction costs and income effects are negligible. Furthermore, it is assumed that all emitters are identical and that the firms subject to cap-and-trade emit as much as they are allowed to each period. Probable implications from relaxing the latter assumptions are briefly discussed after the analysis. The policy variables have to be decided upon before the uncertainty has been resolved and the policy maker's objective is to minimize expected efficiency losses.

For simplicity, we look at the extreme case where the damage function is linear, which implies that the *MAB* is constant with respect to abatements, *i.e.*, has a slope equal to zero. This simplifies the mathematical presentation but it also ensures that, as long as the *MAC* increases in abatements, a full emissions tax solution would be preferable in the absence of the political constraint. Algebraically the *MAB* is written as

$$MAB_j = f + \delta_j \qquad j=1,2 \tag{1}$$

where f is a non-negative constant and δ_j is an independent stochastic variable symmetrically distributed around zero with a variance of σ^2. Subindex j denotes which period, 1 or 2, the variables refer to. The economy consists of N emitters, each with a *MAC* of the following linear shape

$$MAC_{ij} = K - Le_{ij} + \varepsilon_j \qquad i=1,2\ldots N;\ \ j=1,2 \tag{2}$$

Where K and L are non-negative constants, e_{ij} is emitter i's emissions during period j and ε_j is an independent stochastic variable. It is assumed that ε_j is uniformly distributed over a given range of $-a$ to a. The reason for using a uniform distribution is that the upper and lower bound are known and finite. Thereby we can design a regulation policy such that emitters are always in compliance. Alternatively, a penalty system could be incorporated in the model, but this would complicate the presentation without adding to the understanding of the question under study. Consequences of using other distributions, but with the same range, are briefly discussed in a subsequent section. Summing over all N emitters yields the aggregate *MAC* as

$$MAC_j^{tot} = K - \frac{L}{N} e_j + \varepsilon_j \tag{3}$$

Equating (1) and (3) yields the efficient emissions volume, e^*, in a given period as

$$e_j^* = \frac{N\left(K - f - \delta_j + \varepsilon_j\right)}{L} \tag{4}$$

In a standard setting the policymaker may only choose between subjecting all emitters to cap-and-trade, the optimal cap is then equal to the expected efficient emissions level given by (4) with ε and δ equal to zero, or taxing all emitters, the optimal tax would then equal *f*. As previously noted, the latter alternative yields a lower expected efficiency loss, but is generally not applicable when the political constraint is present.

The present paper deviates from the standard setting due to the presence of the political constraint but also due to the possibility of taxing only a subset of the emitters. Thus, let n emitters be members of the taxed sector 1, denoted by subscript $s1$. Clearly, n may not exceed N and it may not be negative. Total emissions made by these n emitters in a given period depend on the tax, T, in the following manner

$$e_{s1,j} = n\frac{K - T_j + \varepsilon_j}{L} \tag{5}$$

2.1 Period 1

The aforementioned political constraint introduces a global cap, Q, which establishes a total emissions volume that may never be exceeded[3]. As the taxed sector's emissions depend on the stochastic variable, ε, the room for emissions made by the trading sector, denoted $s2$, is also stochastic. Since the cap is strict and cannot be calibrated after the realization of ε, the emissions made by the trading sector may not exceed the global cap minus the highest possible emitted amount from the taxed sector. That is, for the first period

$$e_{s2} = Q - \max(e_{s1}) = Q - n\frac{K - T + a}{L} \tag{6}$$

where e_{s2} is total emissions from the trading sector, assuming that the trading sector will emit as much as it is allowed to. Total emissions, from both sectors, during period 1 is given by

$$e_1^{tot} = e_{s1} + e_{s2} = Q - n\frac{a - \varepsilon}{L} \tag{7}$$

[3] As will be shown, if banking occurs during period 1 the global cap in period 2 will be Q plus the amount of banked permits. When there is risk of confusion Q is referred to as the 'underlying global cap'.

which states that at the highest possible realization of ε, *i.e.* $\varepsilon = a$, total emissions will amount to the global cap, Q. Lower values of ε yield lower total emissions. Generally, e^{tot} will differ from e^* – a discrepancy referred to as a 'volume error' – giving rise to an efficiency loss. This may be calculated by the following integral

$$\int_{e_j^{tot}}^{e_j^*} \left(MAC_j^{tot} - MAB_j\right) de \tag{8}$$

or, using (1), (3), (4) and (7) and, thus, looking only on period 1

$$\int_{Q-n\frac{a-\varepsilon}{L}}^{\frac{N(K-f-\delta+\varepsilon)}{L}} \left(K - \frac{L}{N}e + \varepsilon - (f+\delta)\right) de \tag{9}$$

Solving (9) yields an expression for the dead-weight loss following from the volume error in period 1, denoted DWL_{VE1},

$$DWL_{VE1} = \frac{\left(LQ - an + N(f - K + \delta_1 - \varepsilon_1) + n\varepsilon_1\right)^2}{2LN} \tag{10}$$

As noted, the dual regulation approach does not equate marginal abatement costs among the emitters, *i.e.*, it does not allocate abatement efforts efficiently throughout the economy. This is referred to as an 'allocation error'. To calculate the efficiency loss following from the allocation error the total abatement cost (*TAC*) of achieving the outcome under a dual regulation approach is compared to the lowest possible total abatement cost of achieving the same outcome. *TAC* is reached by integrating the relevant *MAC* function from the business as usual (*BAU*) emissions level, calculated as the emissions level at a price or tax equal to zero, to the level valid under the regulation. *TAC* for the taxed sector, *s1*, is

$$TAC_{s1} = \int_{e_{s1}}^{e_{s1}(BAU)} MAC_{s1} de_{s1} = \int_{n\frac{K-T+\varepsilon}{L}}^{n\frac{K+\varepsilon}{L}} \left(K - \frac{L}{n}e_{s1} + \varepsilon\right) de_{s1} = \frac{nT^2}{2L} \tag{11}$$

Where $e_{s1}(BAU)$ is the taxed sector's emissions under *BAU*, derived using (5) with a T=0. The corresponding operation for the trading sector, *s2*, is

$$TAC_{s2} = \int_{e_{s2}}^{e_{s2}(BAU)} MAC_{s2} de_{s2} = \int_{Q-n\frac{K-T+a}{L}}^{(N-n)\frac{K+\varepsilon}{L}} \left(K - \frac{L}{N-n} e_{s2} + \varepsilon \right) de_{s2} = \frac{\left(an - LQ + N(K+\varepsilon) - n(T+\varepsilon)\right)^2}{2L(N-n)} \tag{12}$$

That TAC_{s2} depends on the stochastic variable ε, while TAC_{s1} does not, is not surprising given the assumption of linear *MAC* functions. Perhaps more surprising is that TAC_{s2} depends on the tax. This follows from that when the tax is high emissions from the taxed sector will be lower, which leaves a greater room for emissions from the trading sector and, hence, a lower *TAC* for this sector.

Total abatement costs following from the dual regulation approach, given ε, is thus $TAC_{s1+s2} = TAC_{s1} + TAC_{s2}$, which is to be compared to the lowest possible *TAC* that achieves the same outcome in terms of aggregate emissions. This may be calculated by integrating the aggregated *MAC* from the emissions level given by the dual regulation outcome to the *BAU* level.

$$TAC_{\min} = \int_{e^{tot}}^{e^{tot}(BAU)} MAC^{tot} de = \int_{Q-n\frac{a-\varepsilon}{L}}^{N\frac{K+\varepsilon}{L}} \left(K - \frac{L}{N} e + \varepsilon \right) de = \frac{\left(an - LQ - n\varepsilon + N(K+\varepsilon)\right)^2}{2LN} \tag{13}$$

The efficiency loss following from the allocation error during period 1, denoted DWL_{AE1}, is the difference between TAC_{s1+s2} and TAC_{min} for that period,

$$DWL_{AE1} = TAC_{s1+s2} - TAC_{\min} = \frac{n\left(an - LQ - n\varepsilon_1 + N(K - T_1 + \varepsilon_1)\right)^2}{2L(N-n)N} \tag{14}$$

Note that DWL_{AE1} equals zero if $n = 0$, *i.e.*, when all emitters are in the trading sector, which is the expected result. However, this should also be the case when all emitters are taxed, *i.e.*, when $n = N$. Casual inspection of (14) does not suggest this to be the case. However, a solution where all emitters are taxed, implying $e_{s2} = 0$, can only coexist with the political constraint, Q, if the tax is such that e_{s1} is less than or equal to Q at the highest possible realization of the *MAC*, *i.e.*, at $\varepsilon_1 = a$. Assuming this constraint is binding, which subsequently will be shown to be the case, this tax amounts to $a + K - LQ/n$. Inserting this into (14) yields $(N - n)(LQ + n\varepsilon_1 - an)^2/2LnN$, which approaches zero when n approaches N. Thus, taking account of the political constraint, there is no efficiency loss due to the allocation error neither in the situation where all emitters trade nor when they are all subject to an emissions tax.

2.2 Period 2

For the second period the same calculations as above apply with two modifications. First, the global cap will differ. As noted, this is because total emissions during period 1 are likely to be less than the global cap creating a surplus that may be utilized during period 2. Second, the tax applied to the taxed sector may differ between periods. The number of taxed emitters, n, and the underlying global cap, Q, however, may not. Denote the global cap applicable for period 2 by Q_2. This may be calculated by

$$Q_2 = Q + \left(Q - e_1^{tot}\right) = Q + \left(Q - \left(Q - n\frac{a-\varepsilon_1}{L}\right)\right) = Q + n\frac{a-\varepsilon_1}{L} \tag{15}$$

Q_2 thus depends on the outcome of MAC in the first period, through ε_1, such that the lower the MAC the higher the surplus that may be utilized in period 2 and, consequently, the higher the Q_2. The efficiency loss due to the volume error in period 2, DWL_{VE2}, is calculated in the same manner as DWL_{VE1}. A shortcut is thus to first substitute ε_1 and δ_1 in DWL_{VE1}, from (10), with ε_2 and δ_2 respectively and then substitute Q in the resulting expression with (15). This yields

$$DWL_{VE2} = \frac{\left(fN - KN + LQ + N\delta_2 - n\varepsilon_1 - (N-n)\varepsilon_2\right)^2}{2LN} \tag{16}$$

Similarly, the efficiency loss due to the allocation error in period 2 is reached by the same substitutions, together with substituting T_1 with T_2, applied to DWL_{AE1}, from (14). This yields

$$DWL_{AE2} = \frac{n\left(KN - LQ - NT_2 + n\varepsilon_1 + (N-n)\varepsilon_2\right)^2}{2L(N-n)N} \tag{17}$$

2.3 Total Expected Efficiency Loss

The total present value of the efficiency losses occurring in both periods is

$$DWL_{tot} = DWL_{VE1} + DWL_{AE1} + \rho\left(DWL_{VE2} + DWL_{AE2}\right) \tag{18}$$

where ρ is a discount factor taking on values from 0 to 1. A $\rho = 0$ implies that the effects occurring in period 2 are ignored and the model, in effect, becomes a one-period model. In expectation terms (18) becomes

$$
\begin{aligned}
E\{DWL_{tot}\} = {} & \frac{a^2\left(4n^2 - 2nN + N^2\right) - 6an\left(fN - KN + LQ\right) + 3\left(fn - KN + LQ\right)^2 + 3N^2\sigma^2}{6LN} + \\
& \frac{n\left(a^2\left(4n^2 - 2nN + N^2\right) - 6an\left(LQ + NT_1 - KN\right) + 3\left(LQ + NT_1 - KN\right)^2\right)}{6LN\left(N - n\right)} + \\
& \rho\left(\frac{a^2\left(2n^2 - 2nN + N^2\right) + 3\left(fn - KN + LQ\right)^2 + 3N^2\sigma^2}{6LN} + \right. \\
& \left. \frac{n\left(a^2\left(2n^2 - 2nN + N^2\right) + 3\left(KN - LQ\right)^2 + 3NT_2\left(2LQ + NT_2 - 2KN\right)\right)}{6LN\left(N - n\right)}\right)
\end{aligned}
\tag{19}
$$

Where the first row is $E\{DWL_{VE1}\}$, the second is $E\{DWL_{AE1}\}$, the third is the discount factor times $E\{DWL_{VE2}\}$ and the fourth is $E\{DWL_{AE2}\}$.

3. The Analysis

The aim for the policy maker is to set the policy variables n (the number of emitters in the taxed sector), T_1 and T_2 (the tax in each period respectively) in such a way that the expected efficiency loss given by (19) is minimized. Prior to period 1 only n and T_1 may be set as T_2 is set after period 1 when the value of ε_1 is known. As noted, Q is initially assumed to follow from a political constraint and is treated as an exogenously given constant. Later in the analysis this assumption is relaxed and the case where also Q is a policy variable is studied. For now, let the (underlying) global cap be given by

$$
Q = \frac{N\left(K - f\right)}{L} + \Omega\frac{aN}{L} \tag{20}
$$

where the first term is the expected efficient emissions level, $E\{e^*\}$, and Ω is a constant capturing how far from $E\{e^*\}$ the Q is. Ω may take on values between –1 and 1. The lower (higher) limit corresponds to a Q that equates the expected *MAB* with the lowest (highest) possible *MAC*.

In the following it is assumed that N is large enough to justify an approximation of each emitter being marginal and thereby making it possible to differentiate the expected efficiency loss with respect to n.[4]

3.1 Optimal Taxes

Substituting Q in (19) with (20) and calculating the first order condition applicable for T_1^* – the optimal tax to apply during period 1 – yields

$$T_1^* = f + a\frac{n}{N} - a\Omega \tag{21}$$

As noted T_2^* is set after period 1, when the value of ε_1 is known. At that point in time the outcome of period 1 is 'sunk' and the aim of the policy maker is to set T_2 such that the expected efficiency loss in period 2 is minimized, *i.e.*, $min(DWL_{VE2}+DWL_{AE2})$ treating ε_1 as a known constant. This yields

$$T_2^* = f + \varepsilon_1\frac{n}{N} - a\Omega \tag{22}$$

Equation (21) and (22) provide some interesting insights. First, T_1^* is the lowest possible tax such that the total emissions level in period 1 never exceeds Q, *i.e.*, a lower tax is not feasible. Consequently, if ε_1 takes on its highest possible value, *i.e.*, a, there is no surplus generated in period 1 to be utilized in period 2 and T_2^* will equal T_1^*. All other realizations of ε_1 result in a T_2^* strictly less than T_1^*, since such realizations result in a greater global cap during period 2. Second, if $\Omega = 0$ and all emitters are taxed then T^*_1 equals $f + a$. If n is lowered from N, the tax in period 1 is less than $f + a$ such that when n approaches zero, T^*_1 approaches f. If Q is set at its highest possible level, *i.e.*, $\Omega = 1$, then T^*_1 will equal $f - a(1 - n / N)$, which implies that, if all emitters are taxed, T^*_1 will equal f. This is of certain interest as it corresponds to the optimal solution in the standard emissions tax case (without the political constraint). Third, decreasing Ω increases both T_1^* and T_2^*.

[4] An alternative approach would be to normalize N to 1 and letting it be perfectly devisable. The result will not differ between these approaches.

3.2 Optimal share of the economy to include in the taxed sector

An expression for the optimal amount of emitters in the taxed sector, n^*, is reached in a similar manner as above; using (20) in (19) to derive a first order condition that, as its only real root, yields

$$n^* = \frac{N}{18}\left(13 + 2\rho + 6\Omega - \frac{(2\rho + 6\Omega - 5)^2}{\varphi} - \varphi\right) \tag{23}$$

where

$$\varphi = \Big(-8\rho^3 + \rho^2(60 - 72\Omega) - 24\rho(3\Omega - 7)(2 + 3\Omega) - (6\Omega - 5)^3 +$$
$$18\sqrt{3\rho\left(-8\rho^3 + \rho^2(60 - 72\Omega) - 24\rho(3\Omega - 7)(2 + 3\Omega) - (6\Omega - 5)^3 - 243\rho\right)}\Big)^{1/3}$$

Equation (23) is well behaved for all valid values of ρ and Ω such that $0 \leq n^* \leq N$.

Before moving on to investigate in more detail how n^* depends on the chosen global cap, let us provide some intuition for the result in (23). Consider the first period under a global cap such that $\Omega = 0$. Then, for the lower half of the possible realizations of ε the efficient emissions level is below the global cap. Introducing a taxed sector will, for these realizations, result in an outcome closer to the efficient emissions level compared to a case without such a sector. This is because at low realizations of ε the emissions from the taxed sector will be relatively small[5]. Thus, for these realizations the presence of a taxed sector decreases the volume error and may, subject to the adverse effect of the allocation error, decrease the expected efficiency loss. Important to note is, however, that for the upper half of possible realizations the opposite applies. Only at the highest possible value of ε, *i.e.*, $\varepsilon = a$, will aggregate emissions equal the global cap. As long as $n > 0$ all other realizations yield aggregate emissions strictly below the global cap. That is, when $\Omega = 0$, all realizations of ε between zero and a result in outcomes further from the efficient emissions level when $n > 0$, compared to when $n = 0$. An optimal solution where $n^* > 0$ thus implies that the efficiency

---

[5] The emissions volume from the trading sector does not depend on the realized value of ε, as this only affects the price.

gain from a dual regulation regime when ε turns out to be low outweighs the efficiency loss at the corresponding high outcome. From this line of reasoning it seems intuitive that if the global cap is increased, the size of the taxed sector should also increase. This is because fewer realizations are such that the efficient emissions level lies above the global cap.

That this intuition is correct is seen in figure 1, which plots the optimal proportion of emitters in the taxed sector, n^*/N, as a function of Ω, using (23), for three different values of the discount factor; $\rho = 0$, $\rho = 0.5$ and $\rho = 1$. A point on the left hand side of figure 1 implies a 'strict' global cap, *i.e.*, a cap below the expected efficient emissions level. Similarly, points on the right hand side of figure 1 imply 'lenient' global caps, *i.e.*, that global emissions are allowed to exceed the expected efficient level.

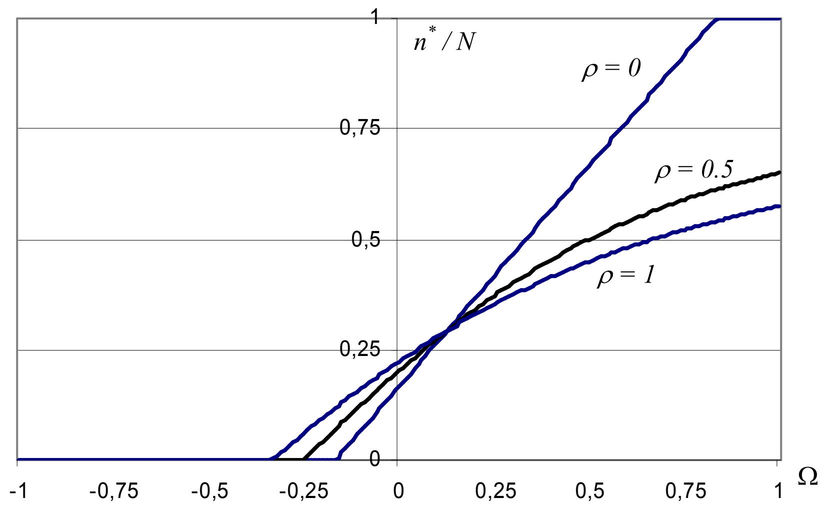

Figure 1, *optimal proportion of emitters in the taxed sector as a function of* Ω *for three different values on the discount factor.*

First note that for most, but not all, values of Ω the optimal proportion of emitters to include in the taxed sector is larger than zero and less than 1, *i.e.*, a dual regulation approach is in these cases better from an efficiency point of view than a full cap-and-trade system or an emissions tax with full coverage. Second, n^* / N increases in Ω, *i.e.*, under a stricter global cap the optimal taxed sector is smaller. This follows from the intuition outlined above. For low values of Ω, n^* / N equals zero. This also follows from the intuition discussed above since for sufficiently strict global caps the majority of realizations of ε will result in outcomes further from the efficient ones if there is a taxed sector present. Thus it is, in these cases, better from an efficiency point of view to let all emitters belong to the trading sector.

Looking at the single period case, *i.e.* when ρ equals zero, it is evident that for some very high values of Ω the optimal strategy entails including all emitters in the taxed sector. Note that, in order for the global cap not to be exceeded, the tax must then be higher than f for all values of Ω less than 1. The logic behind this is similar to the previous one; if Ω is high the global cap is 'lenient' and the possibility that the efficient emissions level is below this cap is high. For a majority of the possible realizations of ε an inclusion of a taxed sector then implies an actual outcome closer to the efficient emissions level. However, this also implies a surplus to be utilized in period 2. As the global cap is already in period 1 above the expected efficient emissions level this is likely to result in a global cap during period 2 far above the efficient emissions level and, thus, a high expected efficiency loss during that period. When ρ equals zero, this has no impact on the optimal policy as future outcomes then is given zero weight. Giving more weight to efficiency losses occurring in period 2, *i.e.*, applying a higher ρ, results in that n^* / N decreases for high values of Ω. Subjecting a smaller proportion of the emitters to the tax decreases the expected surplus and consequently the expected efficiency loss in period 2. However, this happens on the expense of an increased expected efficiency loss in period 1.

The very same argument explains why an increase in the discount factor yields a higher optimal proportion of emitters to tax for values of Ω around zero. Increasing the taxed proportion of the economy increases the expected global cap during period 2. For Ω around zero this decreases the expected efficiency loss during period 2, as the actual emissions volume is closer to the efficient one for a larger share of the possible outcomes. Thus, putting more weight on the efficiency loss in period 2 through a higher discount factor favours a larger proportion of emitters to be taxed. This also explains why, for some values of Ω, it is optimal from a one-period perspective to adopt a cap-and-trade approach with full coverage while, if also future outcomes are taken into account, the optimal approach is to tax a subset of the emitters.

As discussed above, banking plays a crucial role in the optimal policy design through its influence on the global cap in period 2. It may be interesting to examine the outcome in a situation where banking is not allowed. As there is then no interaction between periods this, as noted in the previous section, corresponds to studying period 1 only, by applying a ρ equal to zero. The optimal design in period 2 will be identical to that in period 1, since there is no longer a link between the two. For clarity, not allowing for banking does not necessarily imply that future outcomes are of no interest, which is the literal implication of setting ρ to zero, but in terms of the present model it will be captured through $\rho = 0$.

As seen in (23) the use of Ω makes the optimal taxed proportion independent of the stochastic range surrounding the *MAC*, *a*, as well as the slope of the *MAC* functions, *L*. However, for a given global cap, *i.e.*, a given value of *Q*, the value of Ω depends on both *a* and *L*, seen from rearranging (20) as

$$\Omega = \frac{1}{a}\left(\frac{QL}{N} - (K - f)\right) \tag{26}$$

Differentiating (26) with respect to *a* yields

$$\frac{\partial \Omega}{\partial a} = \frac{L}{Na^2}\left(\frac{N(K-f)}{L} - Q\right) = \frac{L}{Na^2}\left(E\{e^*\} - Q\right) \tag{27}$$

The sign of (27) depends on the relationship between the expected efficient emissions level and the global cap. If the stochastic range increases, the value of Ω must decrease (increase) to keep *Q* constant at a value above (below) the expected efficient emissions level. This, from figure 1, implies that the optimal proportion to tax decreases for lenient global caps and increases for strict ones if the stochastic range is increased. Similarly, differentiating (26) with respect to *L* yields

$$\frac{\partial \Omega}{\partial L} = \frac{Q}{aN} \tag{28}$$

which is positive. Thus, if the slope of the *MAC* functions is increased, Ω must increase to keep *Q* constant. This implies that the optimal taxed proportion must increase, as seen in figure 1. It should be noted that a change in *L* influences not only the term which Ω is multiplied with in (20), as in the case with a change in *a*, but also the expected efficient emissions level itself. This explains why (27) changes sign depending on the value of *Q* while (28) does not.

3.3 Optimal Global Cap

Thus far it has been assumed that the global cap is given by the political constraint. This line of reasoning perhaps fit better with some real life situations than a scenario where the global cap is optimally set from an efficiency point of view. Nevertheless, it is interesting to study the optimal level of the global cap, *i.e.*, the Q that minimizes the total expected efficiency loss.

Using (20), *i.e.*, Ω is studied rather than Q directly, in (19) the Ω that solves the resulting first order condition is given by

$$\Omega^* = \frac{n\left(a + f - T_1 + \rho\left(f - T_2\right)\right)}{aN\left(1+\rho\right)} \tag{29}$$

Inserting the expressions for T_1 and expected T_2 from (21) and (22) into (29) yields

$$\Omega^* = \frac{n}{N + N\rho} \tag{30}$$

Which has the, in the light of the previous discussion, expected characteristic that if n is large the optimal global cap must also be large. This has to be the case since otherwise would the global cap be exceeded for some realizations of ε. Also in line with the previous discussion, the optimal global cap decreases the more weight is put on future outcomes. Things get more interesting when the expression for optimal number of emitters in the taxed sector, from (23), which in itself is a function of Ω is inserted in (30). This yields a complex expression in which Ω^* is a function of the discount factor, ρ, only. Figure 2 plots Ω^* for ρ from 0 to 1.

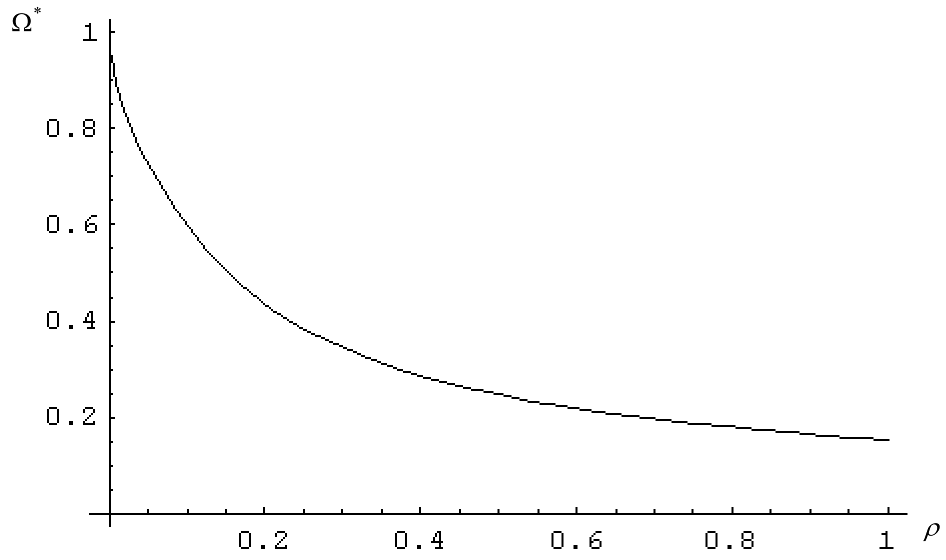

Figure 2, *Ω* as a function of ρ*

If $\rho = 0$, *i.e.*, no weight is put on efficiency losses in period 2, the optimal Ω equals 1. This implies that in a one-period setting the optimal policy entails setting the global cap at its highest possible level, such that it never binds. Then, from (23) or seen in figure 1, all emitters should be taxed and, from (21), the tax should equal f. This thus corresponds exactly to the optimal solution in a setting as the one described in Weitzman (1974). In the light of the previous discussion, this is not surprising. As noted, a $\rho = 0$ implies putting no weight on future outcomes, but it also corresponds to the case where banking is not allowed.

If ρ is increased, matters will be different due to the influences from policy choices in period 1 on the global cap in period 2. This, in turn, influences the expected efficiency losses accruing in period 2 and, for $\rho > 0$, these must be taken into account when deciding on policy design. From the discussion above, a lenient global cap results in a large expected surplus in period 1. If, as has been assumed, this surplus may be used in period 2, the expected efficiency loss in that period is large. Consequently, if ρ increases, thereby putting more weight on efficiency loss in period 2, Ω^* strictly decreases in a convex manner, as seen in figure 2.

3.4 Effects of relaxing the assumptions of homogenous emitters and uniform distribution

To facilitate the presentation a series of simplifying assumptions has been made. For instance, in the model it is assumed that all emitters are identical. If instead they could differ in respect to the slope of their individual marginal abatement cost functions it would be possible to address the question of which emitters to optimally put in the taxed sector. Drawing on the model in Mandell (2008) it is unlikely that it would matter whether emitters with steep or flat marginal cost functions are taxed. Rather, it is the slope of the aggregate marginal abatement cost function for the taxed sector that is of concern. The same outcome may thus be achieved by taxing many emitters with steep functions or few with flat ones. Therefore, this is likely to be a question of whether the two alternatives differ in some other respect.

Furthermore, an assumption about the uncertainty around the marginal abatement cost function, saying that it is uniformly distributed around zero, has been used. Using some other density function, but with the same upper and lower bound, will influence the result. The outcome from the dual regulation approach performs best, compared with a full cap-and-trade alternative, for realizations close to the lower bound. For realizations close to the upper bound the outcome from the two approaches is nearly identical. For some (high) intermediate realizations, however, the outcomes from the dual regulation approach are worse than those under cap-and-trade. Consequently, a distribution that put high probability on such intermediate realizations may favour a full cap-and-trade approach to the dual regulation one.

The ability of banking implies that the expected global cap in period 2 exceeds that in period 1, which in turn implies that the expected tax level in period 2 is below that in period 1. This might seem counterintuitive from a practical perspective where it may be more appropriate to gradually decrease emissions, and perhaps also increase taxes, over time. To a large extent this is driven by an assumption saying that the (underlying) global cap is the same in both periods. As noted, relaxing this assumption, *e.g.*, by applying a stricter underlying global cap in period 2, will have no influence on the qualitative results. However, for the present model to be applicable, both caps must be declared prior to period 1.

4. Conclusions

In centre of attention has been the combination of a marginal abatement function that is constant with respect to short run emissions with a constraint saying that the amount of emissions made during a specific period may not exceed a given amount. What makes this combination interesting is that efficiency concerns suggest the use of an emissions tax for the kind of benefit functions studied, while the constraint is easiest fulfilled with a cap-and-trade approach. The use of a two-period model makes it possible to examine the effects of banking, such that emissions permits not used in period 1 may be utilized during period 2.

The model allows for a share of the economy to be subject to an emissions tax while the rest is subject to a standard cap-and-trade regulation. The most striking conclusion from the model is that, if the amount of emissions allowed is not too small, the optimal solution entails taxing some emitters while the others trade. This is the case even though this approach does not equate marginal abatement costs throughout the economy and, thus, not is *ex post* cost effective. This approach is optimal due to the presence of a trade-off: On the one hand, taxing a share of the emitters will under some states yield an outcome closer to the efficient one in terms of aggregate emissions. On the other hand, the share of emitters that are taxed may not be too large, since then the tax would have to be inefficiently high in order not to exceed the given level of allowed emissions.

The case when the maximum allowed amount of emissions, referred to as the global cap, is a policy variable is also studied. It is shown that the optimal global cap is above the expected efficient emissions level. However, since in optimum some emitters will be taxed total emissions actually made will most likely be less than the global cap.

Both the optimal share of the economy to tax and the optimal global cap, but not the optimal tax levels, crucially depend on how much weight is put on future efficiency losses, *i.e.*, those accruing in period 2. It has been showed that the optimal global cap decreases the more weight is put on future outcomes. The optimal share to tax decreases (increases) the more weight is put on future outcomes when the global cap is high (low).

In the introduction two potential applications that may both serve as back-drop to the model were discussed. One entailed emissions associated with a distinct 'disaster level', which should never be exceeded. The other was climate change, and in particular regulating CO_2 emissions. Let us therefore conclude this paper by briefly elaborating on these topics.

The main difference between the two potential applications lies in why there exists an upper limit on emissions. In the climate change setting this is arguably due to political reasons, as discussed in the introduction, while in the other setting it is due to natural characteristics. Hence, in the latter case the discussion in Section 3.3 about how to optimally set the upper limit is hardly applicable. Whether or not to use subsequent series of one-shot periods without interaction or a two period approach depends on the characteristics of the emission at hand. In some situations it seems reasonable that the upper limit on emissions is independent on emission levels in earlier periods, thus calling for the one-shot approach. In other situations it may well be that low emissions in one period will increase the 'disaster level' in the following period, which corresponds to the banking scenario used in the two periods approach. As noted, the model is designed such that it easily captures both approaches.

Turning to the climate change context, both crucial circumstances underlying the model in this paper are present; 1) there is a political constraint resulting in the use of a quantitative target but 2) in the absence of this constraint, a common emissions tax would be preferred. Thus, the underlying logic and intuition of the model should apply. However, there are also several differences between the model and the current agreements regarding climate policy, which imply that the exact results derived may not be applicable. For instance, the global cap utilized in the model is fixed and may not be exceeded. Under the Kyoto protocol the cap may be calibrated through the use of the Clean Development Mechanism (CDM), which allows a party to increase emissions by conducting emissions decreasing projects in nations that do not have obligations under the Kyoto Protocol. Furthermore, if actual emissions exceed a party's allowed amount there is a penalty system, which enables some level of banking. These circumstances would suggest that the optimal size of the taxed sector is larger than suggested in the model, since there are possibilities to increase the cap if the uncertainty is resolved in such a way that this is needed.

A counteracting effect follows from that the model assumes that emitters in the trading sector always emit the maximum allowed amount each period. Under the Kyoto Protocol individual nations and firms may bank permits for use in future periods, thus contradicting this assumption. An option to bank permits is more likely to be used when marginal abatement costs are low, which coincides with when emissions from the taxed sector are low. That is, if the realized marginal abatement costs are low in period 1 the global cap applicable for period 2 is likely to be high when the trading sector is not allowed to bank permits but even higher when they are and, consequently, if this assumption is relaxed the taxed sector should in optimum be smaller.

That is, there are important differences between actual policies used in climate policy and the stylized model developed in this paper. Nevertheless, there are also many similarities and, thus, the model and the intuition it builds on would seem to provide some important input to policy design. In particular, the current mix – for instance illustrated in the EU by subjecting some emitters to a trading programme, the EU ETS, while the rest are handled mainly through emissions taxes – may be justified by efficiency concerns even though it under uncertainty obviously do not lead to a cost effective allocation of abatement efforts.

References

Bohm, Peter (1997) Designing Efficient Treaties to Protect the Global Environment, in Claude Jeanrenaud (ed.) *Environmental Policy Between Regulation and Market*, Birkhäuser Verlag, Basel, pp. 155-170

Cooper, Richard N. (2008) The case for charges on greenhouse gas emissions, Discussion paper 2008-10 *Harvard Project on International Climate Agreements*, Cambridge, Mass.

Hoel, Michael and Larry Karp (2001) Taxes and quotas for a stock pollutant with multiplicative uncertainty, *Journal of Public Economics*, **82**, pp. 91-114

Hoel, Michael and Larry Karp (2002) Taxes versus quotas for a stock pollutant, *Resource and Energy Economics*, **24**, pp. 367-384

Karp, Larry and Jiangfeng Zhang (2006) Regulation with anticipated learning about environmental damages, *Journal of Environmental Economics and Management*, **51**, pp. 259-279

Mandell, Svante (2008) Optimal Mix of Emissions Taxes and Cap-and-Trade, *Journal of Environmental Economics and Management*, **56**, 131-140

Nordhaus, William D. (2007) To tax or not to tax: alternative approaches to slowing global warming, *Review of Environmental Economics and Policy*, **1**, pp. 26-44

Pearce, David (1991) The role of carbon taxes in adjusting to global warming, *The Economic Journal*, **101**, pp. 938-948

Pizer, William A. (2002) Combining price and quantity control to mitigate global climate change, *Journal of Public Economics*, **85**, pp. 409-434

Weitzman, Martin L. (1974) Prices vs. Quantities, *Review of Economic Studies*, **41**, pp. 477-491