NBER WORKING PAPER SERIES

REGULATION VERSUS TAXATION

Alberto F. Alesina
Francesco Passarelli

Working Paper 16413
http://www.nber.org/papers/w16413

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2010

We thank Pierpaolo Battigalli, Luigi de Paoli, Allan Drazen, Firouz Gahvari, Vincenzo Galasso, Erzo Luttmer, Eugenio Peluso, Jim Snyder, Pierre Yared and partcipants in the 2010 NBER Summer Institute and in a seminar at Bocconi University for useful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Regulation Versus Taxation
Alberto F. Alesina and Francesco Passarelli
NBER Working Paper No. 16413
September 2010
JEL No. D62,D72,D78,H23

ABSTRACT

We study which policy tool and at what level would be chosen by majority voting to reduce negative externalities, such as pollution. We consider three instruments: a rule, that sets an upper limit to the polluting activity, a quota that obliges to proportional reduction, and a tax on the activity. For all instruments the majority chooses too restrictive levels when pollution is mainly due to a small fraction of the population, and when costs for reducing activities or paying taxes are convex, and viceversa. Even though a tax is in general superior to the other two instruments, the majority may strategically choose a rule in order to charge the minority a larger share of the cost for the externality reduction.

Alberto F. Alesina
Department of Economics
Harvard University
Littauer Center 210
Cambridge, MA 02138
and IGIER
and also NBER
aalesina@harvard.edu

Francesco Passarelli
Department of Institutional Analysis and Managemen
via Rontgen, 1
20136 Milan, Italy
francesco.passarelli@unibocconi.it

1 Introduction

There are three ways usually discussed to reduce the amount of an activity generating negative externalities: a rule which sets an upper bound to this activity, a compulsory proportional reduction for everybody of the activity or a tax on it. Much of the analysis of the choice and the level of the instrument is normative, namely what would be the optimal policy tool to maximize social welfare. Generally speaking taxes are superior to quantitative limits because they allow individuals to optimize over the cost of paying the tax or reducing the activity and also because taxes generate revenues.

In this paper we adopt instead a "positive", approach by investigating which policy and at what level would be chosen by majority voting. The latter does not deliver the optimal policy for two reasons. First for given choice of policy instruments, majority voting does not deliver the optimal level of the instrument. Second, and perhaps more interestingly, when choosing amongst alternative instruments majority voting in general does not lead to the choice of the superior one (say taxes rather than rules). For instance, a majority may choose a rule instead of a proportional tax only because a rule concentrates on the minority the burden of the externality reduction. This may explain why sub optimal policy instruments often survive for a long time.

We can relate the nature of the departure from optimality to various features of the distributions of costs and benefits amongst individuals. One of our results is that all instruments tend to be too restrictive (i.e. not allow enough negative externalities) when those who generate them (which in short we sometime label "polluters") are a relatively small minority and the cost that they incur for reducing externalities grows at a fast rate. In particular, a rule is highly inefficient when the minority bears large costs compared to the benefits received by the majority. In fact, a rule is a "cheap" instrument in the hands of the majority for affecting minority's behavior, a sort of tyranny of the majority. Thus we should observe highly restrictive rules for activities concentrated in some specific sectors or for very large "polluters", while rules might be too lax for activities that a majority enjoys, say car pollution.[1]

We should make clear from the outset that we consider only proportional taxes on the polluting activities. By allowing any type of curvature on the tax schedule, including corners, one could reproduce patterns which approximate, say a rule, and are quite far from the allocation generated by a proportional tax. While this would not be a problem for a social planner, from the point of view of a "positive" politico economic model we need to worry about the existence of a Condorcet winner. While we can prove its

[1] A counter argument not considered in this paper is that specific and highly concentrated sectors might have a stronger lobbying capacity.

existence with a proportional tax, in general one cannot do that with any curvature of the tax schedule. Thus all of our positive results would be interpreted as comparing rules and quotas versus a proportional tax on the polluting activities. Realistically speaking these are the kind of policies routinely discussed in this area. Therefore from now on when we refer to a "tax" we mean a proportional tax on the polluting activity. We briefly return to this issue in the conclusions. On the spending side we examine both the case in which revenues are used to produce a public good (in the main body of the paper) and the case in which the revenues are redistributed lump sum (in Appendix). The results are qualitatively the same.[2] Also when thinking specifically of "pollution" there are important issues of intergenerational redistributions of costs and benefits of the externality and of the policies adopted to reduce it. We do not investigate these issues in this paper, so, strictly speaking, we should think of negative externalities which affect only the currently alive, say "noise".

We frame our model in the tradition of the political economy literature on fiscal policies;[3] but we focus upon rules and externalities rather than redistributive taxation. The small literature which introduces political economy considerations in this area is confined to environmental issues.[4] We share with this literature the idea that majorities may prefer regulation to taxes, even when the latter would be socially optimal, whenever regulation and taxes are available policy options. Buchanan and Tullock (1975) compare environmental taxes with a proportional reduction of polluting activities, that they call "regulation". There is no voting stage or any specification of the political process in their work. They offer several arguments in favor of taxes, but they claim that people are more likely to prefer proportional reduction. Congleton (1992) focuses on how political institutions affect the enactment of environmental regulations. Individuals belong to two classes: a high income elite, and the rest of the population. They have to choose the amount of a costly environmental standard. Since the standard reduces aggregate income, the elite wants lower control. Thus an authoritarian regime is less inclined to stringent standards than a democratic system. Boyer and Laffont (1999) look at the optimal level of flexibility that should be delegated to the majority. Different majorities have different stakes in the rents of a polluting monopolist, and there is asymmetry in information. Fluc-

[2]Some specific differences between the conditions which determine the extent of the inefficiency are described in the Appendix.

[3]Since the seminal works of Roberts (1977) and Meltzer and Richard (1981), this literature has extensively studied the conflict between majority and minority and how alternative institutional arrangements affect the efficiency of the government, its size, and the allocation of public spending. See Persson and Tabellini (2000). Note that in Meltzer and Richard's framework results only apply to linear tax schedules for the same reason discussed above in the context of the present paper.

[4]For a survey on arguments in favor or against environmental taxes and quantitative regulations, with some reference to political economy issues, see Hepburn (2006).

tuating majorities determine excessive fluctuation in environmental policy. Thus constitutional constraints may be desirable. Cremer, De Donder and Gahvari (2004) study the efficiency of majority voting on an environmental tax. The proceeds are used to refund income and capital taxes. If labor and capital taxes are rebated in the same proportion, the environmental tax chosen by the majority is too low. Efficiency increases by refunding a higher proportion of labor incomes. In a related work (Cremer, De Donder and Gahvari, 2008), they consider the role of militants and opportunists within political parties. When militants are powerful the outcome is a large environmental tax.

The present paper is organized as follows. In Section 2 we set up the basic model of the activity which produces negative externalities. In Section 3 we study the majority vote equilibrium when the policy instrument is a compulsory rule. In Section 4 we consider quotas, and in Section 5 taxation of the activity with the negative externality with public good provision. In Section 6 we study the choice of the policy instrument by majority rule. Section 7 concludes and illustrates several extensions of the model. All the proofs are in Appendix 8.1. Appendix 8.2 presents the model with proportional tax and uniform refunds of the proceeds.

2 The model

Consider a society with a continuum of individuals/voters of size one; each individual has an exogenously given location in the interval $[0, 1]$. Call those locations "types": t_i for individual i. t_i represents the behavior that i can assume at no costs. Instead behaving differently from t_i entails for i an "adjustment cost", which depends on the distance between type t_i and his behavior denoted by b_i. Types and behaviors are constrained in the unit interval: $t_i, b_i \in [0, 1]$. We can think of t_i as the level of the activity that maximizes profits (in case of a firm) or utility (in case of a consumer). For instance, the profit maximizing level of polluting emission for a firm, the ideal speed adopted on a highway by a driver, the ideal alcohol consumption, the effortless production of waste. A deviation from such optimal level entails a cost.

The adjustment cost function, c, is the same for all individuals:

$$c(|b_i - t_i|) \tag{1}$$

with $c(0) = 0$ and $c'(0) = 0$; $c(.) > 0$, $c'(.) > 0$, $c''(.) > 0$, for any $b_i \neq t_i$.

The externality produced by an individual with behavior b_i is $\varepsilon(b_i)$ with $\varepsilon'(b_i) < 0$ and $\varepsilon''(b_i) < 0$. The negative externality produced by an individual increases at the margin with his behavior. These assumptions on the externality function, together with those regarding costs, simplify the analysis and reduce the number of "special cases" in this non-parametric model.

More on this below. If we denote with $G(b)$ the cumulative distribution of behaviors, the total (per capita) externality produced in the society is:

$$\int_0^1 \varepsilon(b)dG(b) \tag{2}$$

A given behavior by someone generates the same externality as an equal behavior by anyone else, and the externality produced by an individual is not affected by the behavior of the others. For any behavioral profile $G(b)$ everyone receives the same externality. Changing behavior comes at the same cost for everyone, and it does not affect the externality produced by anyone else. The utility of individual i, U_i, is given by the difference between the total externality received and the private adjustment cost:

$$U_i = \int_0^1 \varepsilon(b)dG(b) - c(|b_i - t_i|) \tag{3}$$

Each individual is infinitely small. The externality change that he perceives from modifying his own behavior is infinitesimal. As a consequence, he is never willing to adopt a behavior that is different from his type, independently of the behavior of any other individual. Then if $F(t)$ is the cumulative distribution of types, free riding implies that $F(t)$ is also the equilibrium behavioral profile. Utility in equilibrium becomes:

$$U_i(F(t)) = \int_0^1 \varepsilon(t)dF(t)$$

There is scope for government intervention.

3 Voting on a rule

Consider the case of a rule, ρ, which fixes an upper bound to the behavior of all individuals. The timing is as follows: first, individuals compute their policy preferences regarding ρ; then they vote, selecting $\hat{\rho}$ in pair-wise voting; finally, they choose their behavior. The rule, $\hat{\rho}$, is fully enforced. All types higher than the rule have to adjusts and pay the costs; all types below $\hat{\rho}$ continue to behave just as their types. Any individual knows that, by voting for a rule ρ, he can affect the behavior of $1 - F(\rho)$ individuals whose types are above ρ, and can enjoy from the reduction of their negative externalities. However, if ρ is lower than his type, he has to bear a private adjustment costs. The individual preference function can be then written as

$$U_i(\rho) = \varepsilon(\rho) \cdot (1 - F(\rho)) + \int_0^{\rho} \varepsilon(t) \cdot f(t)dt - c(|\rho - t_i|) \tag{4}$$

where $f(t) = F'(t)$. The first term in the right-hand side of (4) is the externality received by i that is produced by all the affected individuals (i.e.

those with $t_i > \rho$); the second term is the externality received by i that is produced by the non-affected individuals below ρ; the third term is i's private compliance cost. If we assume that marginal gains from setting a lower rule are decreasing, then $U_i(\rho)$ is concave. In this case, i has a uniquely preferred rule. Formally, concavity is ensured when the following inequality is satisfied for any ρ:

$$\varepsilon''(\rho) \cdot (1 - F(\rho)) - \varepsilon'(\rho) \cdot f(\rho) < c''(|\rho - t_i|) \tag{5}$$

Note that convexity of costs and concavity of ε are not sufficient to satisfying (5). In fact, F needs to be "smooth" overall.[5] Under this condition, which we assume from now on, a Condorcet winner exists (Black, 1948).

When evaluating a rule an individual trades off his private compliance sacrifice against the reduction in externality due to affecting other people. Call ρ_i^* the most preferred rule, or i's bliss point. If $\rho_i^* \in (0, t_i)$, the FOC for maximizing $U_i(\rho)$ is satisfied:

$$(1 - F(\rho)) \cdot \varepsilon'(\rho) = c'(|\rho - t_i|) \tag{6}$$

Equation (6) shows that the most preferred rule is set where the marginal private benefit due to affecting $1 - F(\rho)$ individuals equals the marginal private cost due to complying with the rule.[6] Nobody would prefer a rule higher than his type; in fact that rule would be dominated by a rule equal to the individual's type, since he would reduce negative externalities and not suffer an adjustment cost. Due to our assumption that marginal cost in t_i is zero, an individual prefers a rule that is lower than his type; thus $\rho_i^* \in [0, t_i)$.[7]

The following Lemma states that lower types prefer lower rules.

Lemma 1 *For any two players i and j, if $t_i > t_j$, then $\rho_i^* \geq \rho_j^*$*

Call t_s the median type and let ρ_s^* be his bliss point. Under majority rule, the voting outcome, $\hat{\rho}_s$, is the bliss point of the median type:

$$\hat{\rho}_s = \rho_s^*$$

The socially optimal rule $\hat{\rho}^*$ in general differs from the voting outcome, $\hat{\rho}_s$. In fact $\hat{\rho}^*$ maximizes a "social" policy preference schedule, call it $W(\rho)$, that

[5] By "smoothness" we mean that the density of types must never be too high, otherwise, for some ρ, the marginal gains from reducing the rule might not be decreasing due to high density of new individuals affected.

[6] The reader should remind that both sides of (6) are negative. In particular, since $\rho < t_i$, the right-hand one is negative due to the presence of the absolute value operator in the argument of c.

[7] Of course, corner bliss points, with $\rho_i^* = 0$, will concern quite low types.

is the sum of all players' utilities subject to the fact that, once ρ has passed, all types above ρ lower their behaviors down to ρ:

$$W(\rho) = \varepsilon(\rho) \cdot (1 - F(\rho)) + \int_0^\rho \varepsilon(t) dF(t) - \int_\rho^1 c(|\rho - t|) dF(t)$$

If $\hat{\rho}^* \in (0,1)$, then it solves the following FOC:

$$(1 - F(\rho)) \cdot \varepsilon'(\rho) = ac'(\rho) \tag{7}$$

where $ac'(\rho) = \int_\rho^1 c'(|\rho - t|) f(t) dt$ represents the *average* marginal cost over the entire population.[8] Note how different the calculus of the social planner is from the calculus of the median voter. Both consider the per-capita marginal externality received from $1 - F(\rho)$ affected people. The former, however, is interested in the cost borne in average by any single individual in the society. The latter pays attention only to his own private marginal cost. Suppose that $c'(|\rho - t_s|)$ is, in absolute value, low compared to $ac'(\rho)$. In this case, the median voter has an incentive to fix a low rule. Since costs are convex, this case is more likely when the median voter is a low type, compared to the other affected people.

Say that, if $\hat{\rho}_s < \hat{\rho}^*$, the rule is *too restrictive*; if $\hat{\rho}_s > \hat{\rho}^*$, the rule is *too permissive*. In the case both $\hat{\rho}_s$ and $\hat{\rho}^*$ are internal, we have that:

Proposition 1 *i) Simple majority voting yields a too restrictive (too permissive) rule if and only if in equilibrium the ratio between the median voter's marginal cost and the average marginal cost of the affected population is lower (higher) than the share of the affected population.*
ii) Majority voting yields the socially optimal rule if and only if the ratio between those marginal costs equals the share of the affected population.

What makes an equilibrium rule too restrictive relative to the optimal rule? One factor is the nature of compliance costs: when cost convexity is high, the median voters' marginal cost may be substantially lower than the average. In this case compliance by high types can be socially very costly but the median voter does not care about it. In other words the median voter disregards "too much" relative to the social planner the high costs of high types. Another incentive to set a restrictive rule arises from the distribution of types. If the median voter has a concentration of types close but below his position lowering the rule is highly productive for him since he can affect the behavior of additional individuals with a relatively limited private adjustment cost. Broadly speaking, this situation is likely to occur when $F(t)$ is rather skewed towards high types, in the sense that the

[8] Notice that the average marginal cost *of the affected population* is $ac'(\rho)/(1 - F(\rho))$. This is the average considered in Proposition 1 below.

median type is substantially lower than the average type.[9] This tendency to disregard costs borne by others is more evident when externalities are low compared to adjustment costs. In this case the median chooses a rule that is very close to his type, forcing other people to substantial changes in their behavior. The benefits that the median enjoys are low, but his cost is low as well.

In summary a too restrictive rule is likely to come about when polluting activities are concentrated in a minority of high types. This implies that in a political equilibrium society is likely to allow "not enough" polluting activities when the latter are concentrated in some specific sectors or producers, and too much pollution for activities enjoyed by many like driving.

4 Voting on a quota

We now analyze a policy which requires a reduction of the activity by a proportion $\tau \in [0,1]$ which we call "quota". Once τ has been decided by the majority, any individual i has to lower his behavior from t_i to $b_i = (1-\tau)t_i$. The policy preference function for i is:

$$U_i(\tau) = \int_0^1 \varepsilon((1-\tau)t)dF(t) - c(\tau t_i) \tag{8}$$

The concavity of policy preferences is ensured by the concavity of ε and the convexity of c.[10] Each voter i has a preferred quota, call it τ_i^*, which, in case it is interior in $(0,1)$, solves the FOC below:

$$a\varepsilon'(\tau) = t_i c'(\tau t_i) \tag{9}$$

where $a\varepsilon'(\tau) = -\int_0^1 t\varepsilon'((1-\tau)t) \cdot f(t)dt$ is the (positive per capita) marginal externality produced, after the quota has been enforced. The most preferred quota equalizes private marginal compliance costs to the private marginal benefits due to behavior reductions by all individuals.

Lemma 2 shows how the most preferred quotas are related to types.

Lemma 2 *For any two individuals i and j, if $t_i > t_j$ then $\tau_i^* \leq \tau_j^*$.*

[9] Consider, however, that a rightward skewed $F(t)$ with a median lower than the average is only a favorable, but not a sufficient condition for a too restrictive rule emerging. Note the analogy with the Meltzer and Richard's (1981) taxation model of in which the key factor determining the tax level is the distance between the median income and its average.

[10] More precisely, the concavity condition is::

$$\int_0^1 t^2 \varepsilon''(.)dF(t) < t_i^2 c''(\tau t_i)$$

In words, consider that a quota forces different types to assume different behaviors. Therefore, the concavity of total private externalities must be evaluated as the average of the second derivative (left-hand side of the inequality above). Of course, since ε'' is negative in any point and c'' is always positive, inequality above is always satisfied.

Therefore, the types who produce larger negative externalities prefer lower quotas. The reason is that, for any quota, they get the same externality as any lower type, but, since they have to adjust more, they bear higher adjustment costs.[11]

Under majority rule, the pivot is the median type, and the voting outcome is his most preferred quota:

$$\hat{\tau}_s = \tau_s^*$$

Let us take the social perspective. We will say that a quota is too restrictive when it is higher than the socially optimal one, and vice versa for a too permissive quota. The social welfare function is:

$$W(\tau) = \int_0^1 \varepsilon((1-\tau)t)dF(t) - \int_0^1 c(\tau t)dF(t)$$

Due to the concavity of the $U_i(\tau)$'s, also $W(\tau)$ is concave. The optimal quota, $\hat{\tau}^*$, maximizes $W(\tau)$. If $\hat{\tau}^* \in (0,1)$, it has to satisfy the FOC:

$$a\varepsilon'(\tau) = ac'(\tau) \tag{10}$$

where, $a\varepsilon'(\tau)$ is the average marginal externality, and $ac'(\tau) = \int_0^1 tc'(\tau t) \cdot f(t)dt$ is the average cost due to a marginal increase in the quota. One could call them the *social per-capita* marginal benefits and costs of a quota.

The left-hand sides in (9) and (10) are the same. This means that the median voter and the social planner perceive marginal benefits from a quota in the same way. Hence, any differences in their choices resides in how they perceive marginal costs. Specifically, the median voter has an incentive to prefer higher quotas when his marginal costs are lower than the average. This is the point made in Proposition 2.

Proposition 2 *Simple majority voting yields a too restrictive (too permissive) quota if and only if in equilibrium the median voter's marginal cost is lower (higher) than the average marginal cost.*

The policy outcome is too restrictive when the median bears low marginal costs, compared to the average. This is more likely to occur when cost convexity is high or type distribution is slanted towards high types. The intuition is that, since a quota forces to adjustments that are proportional to types, a low median makes small adjustments. His marginal costs are relatively low, compared to the average. Then he is inclined to prefer a too high quota. Moreover, highly convex costs "push" the average upwards; therefore, the median's costs remain relatively low at the margin. Consider,

[11] There might be corner bliss points, $\tau_i^* = 1$, which are likely to concern low types, large externalities and low marginal costs.

however, that a rightward skewed $F(t)$ with a median lower than the average is only a favorable, but not sufficient, condition for a too restrictive rule emerging.[12]

In summary, in many cases both a rule and a quota can be too restrictive. But are there cases in which a majority that selects a too permissive quota would alternatively choose a too restrictive rule? The answer is "yes".[13] Interestingly, however, the vice versa is impossible.

Proposition 3 *When the majority selects a rule that is too permissive, then it selects a quota that is also too permissive. The vice versa is not true.*

To explain the intuition behind this result, recall that an instrument is too permissive when the median's marginal cost is higher than the average. Recall also that a rule concentrates costs on high types, whose marginal costs "push" the average upwards. Suppose the rule is too permissive. This means that the median's position is so high or cost convexity is so low that, despite cost concentration, the average marginal cost is lower than the median's marginal cost. By contrast, the quota shares costs more equally across population. This lowers the average which remains below the median's marginal cost. Thus the quota cannot be too restrictive. In synthesis, when the policy is a rule instead of a quota the risk of a too restrictive outcome is higher. In a sense, the risk of the tyranny of a majority of low types is always higher if a rule is adopted.

5 Voting on a tax

We now examine a proportional tax (tax for brevity) μ ($\mu \geq 0$) so that the tax burden for individual i is μb_i. The government uses the tax revenue to provide a non excludable public good, g. In Appendix 8.2 we consider the case in which tax proceeds are redistributed lump sum to population.[14] Call $\gamma(g)$ the individual utility from an amount g of public good. Let γ be increasing, concave and the same for all i. Call $d(\mu b_i)$ the cost that i bears from paying the tax when his behavior is b_i, with $d'(.) > 0$ and $d''(.) > 0$. Given a behavior profile $G(b)$, an amount g of the public good, and a tax μ, individual utility is:

[12] Suppose that average and median types coincide. A special case in which majority rule delivers the socially optimal quota is when marginal costs are linear. This special situation parallels Meltzer and Richard's (1981) model.

[13] Take for example linear marginal costs and a type distribution in which the median is above the average. In the case of a quota, the median's marginal cost is higher than the average (i.e., $t_s c'(\tau_s^* t_s)/ac'(\tau_s^*) > 1$). Thus the quota is too permissive. In the case of a rule, it may easily be that the median selcts a level such that his marginal cost is lower than the average cost of the affected people (i.e., $c'(|\rho_s^* - t_i|) \cdot (1 - F(\rho_s^*))/ac'(\rho_s^*) < 1$), the rule is too restrictive.

[14] The results are qualitatively similar. Some differences in the conditions leading to inefficiency of the policy choices are dicussed in Appendix.

$$U_i(g,\mu) = \int_0^1 \varepsilon(b)dG(b) + \gamma(g) - c(|b_i - t_i|) - d(\mu b_i)$$

Without a tax, non-atomic individuals do not contribute to the public good, unilaterally. The timing is, as always, that individuals compute first their preferences on μ; then they vote in pair-wise voting; then they choose their behavior, and pay taxes accordingly. The government provides the public good with a balanced budget:

$$g = \int_0^1 \mu \cdot b(t,\mu)dF(t) = \mu \cdot \bar{b}$$

where $\bar{b} = \bar{b}(\mu)$ is the "after-tax" average behavior in the society.

The relationship between behavior and tax derives from the individual cost optimization for given tax. The solution of that problem is given by the FOC:[15]

$$-c'(|b - t_i|) = \mu \cdot d'(\mu b) \tag{11}$$

If the tax increases, individuals reduce their behavior and for any tax, higher types will prefer higher behaviors. Thus, from the FOC we can derive in implicit form, the equilibrium behavior:[16]

$$b_i = b(\underset{-}{\mu}, \underset{+}{t_i}) \tag{12}$$

Because of taxes individuals lower their behaviors below their types. Let this effect be the "first dividend". The "second dividend" of taxation is the financing of the public good. Note that the second dividend does not simply add on to the first one. Paying taxes and adjusting behavior are substitutes: people pay taxes in order to avoid to adjust to zero. As a consequence, the public good provision substitutes further externality reductions. Note the difference with the usual Laffer curve. In that case a policy maker would never set an income tax beyond the level that maximizes revenues. In our model it could be optimal to do so because of the double dividend: beyond the maximum, at least for small increases, the loss of tax revenues can be offset by the gain due to the reduction in externality.

Individual' i's indirect utility is:

$$U_i(\mu) = \int_0^1 \varepsilon(b(t,\mu))dF(t) + \gamma(\mu\bar{b}(\mu)) - \omega\,(.,t_i) \tag{13}$$

with $\omega\,(.,t_i) = c(|b(t_i,\mu) - t_i|) + d(\mu b(t_i,\mu))$. The first term in the right-hand side of (13) represents the externality from the after-tax behaviors of

[15] Recall that $c'(|b - t_i|)$ is negative, thus $-c'(|b - t_i|)$ is positive.

[16] Function b is not indexed since, because of the symmetry assumptions on c and d, differences in the behaviors of the individuals result only from differences in their types. The convexity of both c and d takes care of the SOC.

all individuals; the second one is the benefit from the public good; the third term, $\omega\left(.,t_i\right)$, is the adjustment cost from modifying behavior plus the cost of paying taxes. Due to the concavity of ε and γ, and the convexity of c and d, $U_i(\mu)$ is concave. Thus, the bliss point μ_i^* is unique and, if different from zero, it solves the following FOC:

$$\varepsilon'_\mu + \gamma'_\mu = \omega'_\mu\left(.,t_i\right) \tag{14}$$

where $\varepsilon'_\mu = \int_0^1 \varepsilon'_b b'_\mu f(t)dt$ and $\gamma'_\mu = \gamma'_g \cdot \left(\bar{b} + \mu\bar{b}'_\mu\right)$ are the private marginal benefits from externality reduction and from the public good, respectively. The right-hand side of (14) represents the private marginal cost of taxation, where $\omega'_\mu\left(.,t_i\right) = c'(.)\cdot b'_\mu(.) + d'(.)\cdot\left(b(.) + \mu b'_\mu(.)\right)$. Since c and d are convex, then $\omega'_\mu\left(.,t_i\right)$ is increasing in t_i. Lemma 3 below is a tool to solve the voting stage.

Lemma 3 *For any two individuals i and j, if $t_i > t_j$ then $\mu_i^* \leq \mu_j^*$.*

Under the simple majority, the median type is the pivot, and the voting outcome is his most preferred tax: $\hat{\mu}_s = \mu_s^*$.

Let us consider the efficiency of the voting outcome. The policy benchmark maximizes the following social preference function:

$$W(\mu) = \int_0^1 \varepsilon(b(.))dF(t) + \gamma(\mu\bar{b}(\mu)) - \int_0^1 \omega\left(.,t\right)dF(t)$$

Concavity of the individual preferences ensures that also $W(\mu)$ is concave; thus the socially optimal tax, $\hat{\mu}^* \in (0,\infty)$, solves the following equation:

$$\varepsilon'_\mu + \gamma'_\mu = a\omega'_\mu \tag{15}$$

with $a\omega'_\mu = \left\{\int_0^1 \left[c'(.)\cdot b'_\mu(.) + d'(.)\cdot\left(b(.) + \mu b'_\mu(.)\right)\right]dF(t)\right\}$; ε'_μ and γ'_μ are the same as above. The social planner would choose a tax such that per-capita marginal benefits are equal to per-capita (or average) marginal costs. Observe that the per-capita social marginal benefit is the same as the private benefit of any individual voter. This means that any difference between the median voter's and the social planner's preferences about the externality tax is due to differences in the marginal costs.

Proposition 4 *Simple majority voting yields a too restrictive (too permissive) externality tax if and only if the median voter's marginal cost is lower than the average marginal cost.*[17]

Proposition 4 suggests that the efficiency of the voting outcome is related to the relative size of the median voter's marginal costs, like rules and quotas.

[17] The proof parallels the proof of Proposition 2 above. Thus we omit it.

Cost convexity and a relatively low position of the median tend to determine a too restrictive outcome.

We have seen above that a rule is more likely to be too restrictive with respect to a quota. What can we say about a tax? If the sacrifice of paying taxes is similar (or even equal) to the sacrifice of changing behavior (i.e. $d(.) = c(.)$) then individuals minimize costs by reducing their behaviors proportionally to their types. This is illustrated in the upper graph of Figure 1, where the marginal cost of adjusting behavior, $c'(.)$, and of paying taxes, $\mu d'(.)$, are symmetric and linear. Two individuals, a high type h, and a low type l, are represented. In this case, the impact of an externality tax is similar to a quota: both types reduce their behaviors proportionally, then they pay proportional taxes on the after-tax behaviors. The allocation of costs across population comes out of proportional adjustments, a mechanism that is similar to the quota. As a consequence also the ratio between median's marginal cost and average marginal cost behaves in the same way as for the quota. This means that a quota and a tax are either too permissive or too restrictive under the same conditions. For example, if both $c'(.)$, and $\mu d'(.)$ are linear (as in the figure) and the median is above the average, then both a tax and a quota are too permissive.

Suppose that the two marginal costs functions are not symmetric, and that the marginal cost of paying taxes grows at a faster rate. This is shown in the lower graph of Figure 1. High types' marginal costs grow more than lower types', pushing the average up. In this case a tax concentrates costs on high types, a mechanism that is similar to a rule. A majority of low types has more incentives to raise taxes opportunistically. The ratio between median's marginal cost and the average gets lower. Thus a tax is more likely to be too restrictive than a quota, and it is more similar to a rule. Vice versa, if the cost of paying taxes is less convex than the cost of adjusting behavior it happens the contrary: a risk of a too restrictive tax is smaller with respect to a quota.

Summing up, a majority is more likely to choose a too restrictive tax than a too restrictive quota when the cost of paying taxes grows faster than the sacrifice of reducing behavior. As for the rule, the majority's decision is always more likely to be too restrictive compared to the other two instruments.

In Appendix 8.2 we show that when tax proceeds are used to make transfers instead of providing a public good, the political distortion in the tax rate is lower since the majority chooses a level that is closer to the socially optimal one.

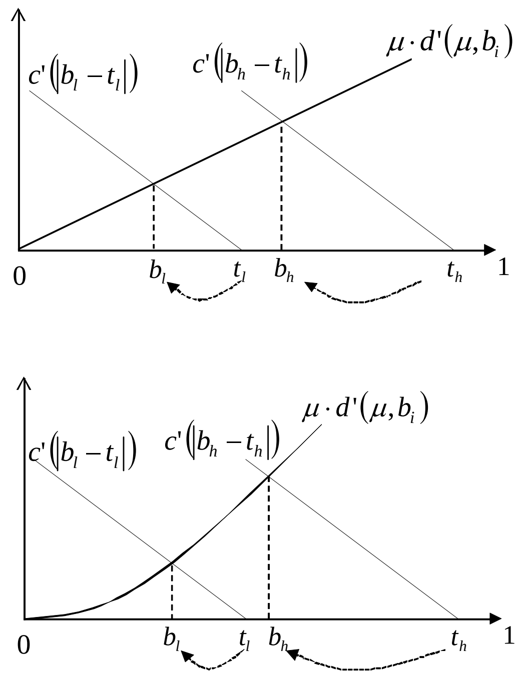

Figure 1: Adjustments and marginal cost equalization

6 The choice of the policy instrument

6.1 Preliminaries: Revenue equivalents

Consider the issue of voting on the policy instrument: now the majority determines not only the level of the policy but also which instrument to adopt. The policy issue becomes two-dimensional thus we have to take care of the existence of a Condorcet winner. Our main result in this Section is that under simple conditions the majority prefers a rule to a tax even though a social planner would choose a tax. This is more likely to happen when the median voter is a relatively low polluter. In this situation the risk of a too restrictive rule is quite high.

The benefits and costs of rules or quotas derive from "behaviors", whereas benefits and costs of externality taxes derive also from "tax revenues". Therefore, in order to compare benefits and costs we introduce the concept of *revenue equivalent* defined as the total amount of virtuous behavior or tax proceeds that it generates. As it will become clear later, we use this concept to represent within the same policy space the choice among the three different instruments. Behavior and proceeds are measured in the same unit. For example, reducing the behavior of a hundred people by 0.2 is revenue equivalent to collecting 20 units of taxes. Note that this is not restrictive since the benefits of taxes and polluting behavior are then evaluated by means of generically different utility and cost functions. Call $RE(.)$ the revenue equivalent of an instrument. For a rule, a quota and a tax we have, respectively:

$$RE(\rho) = \int_\rho^1 tdF(t) - \rho(1 - F(\rho))$$
$$RE(\tau) = \tau \int_0^1 tdF(t)$$
$$RE(\mu) = \int_0^1 [t - (1-\mu) \cdot b(\mu, t)]\, dF(t)$$

As for $RE(\rho)$, note that $\int_\rho^1 tdF(t)$ is the total pre-rule behavior of the affected people, and $\rho(1 - F(\rho))$ is their post-rule behavior. Thus the difference is the total amount of behavioral reduction (i.e., virtuous behavior) induced by the rule. In the second equation, $\int_0^1 tdF(t)$ is the total polluting behavior in the society. The reduction due to the quota is $\tau \int_0^1 tdF(t)$. In the third equation, the revenue equivalent is due to total behavioral reductions ($\int_0^1 [t - b(\mu, t)]\, dF(t)$) and total tax revenues on residual behaviors ($\int_0^1 \mu \cdot b(\mu, t) dF(t)$). Summing and rearranging yields the left-hand side.

Below we study benefits and costs of the three instruments, as a function of the same measure, RE. After, we look for the Condorcet winner, then we discuss efficiency. In Appendix 8.2 we analyze the simpler case in which tax revenues are redistributed lump sum and there is no public good. The results are similar, but we characterize the conditions that make the tax more attractive. A conceptually easy but analytically cumbersome extension would allow us to consider both public goods and transfers.

6.2 Private benefits from RE

A rule performs better when ε is rather concave: the main objective is restricting the behavior of top polluters, then a rule is the best way to do it. However benefits from a tax may be large because of the double dividend.[18] Call $BE(\rho)$ and $BE(\mu)$ the private benefit functions of a rule and a quota respectively. Figure 2 shows reasonable shapes for them: initially the rule performs better because the top polluters are affected; for large RE amounts the double dividend effect becomes more relevant.[19]

6.3 Private costs from RE

1. Let us start with a rule. How much does a given amount of RE costs individual i when the instrument is a rule? As long as $\rho \geq t_i$, there are no

[18] It is worth reemphasizing here that we are restricting our attention to proportional tax. Larger benefits for a majority of low types could be achieved with a progressive tax on polluting activities which would approximate a rule, but as discussed above existence of a Condorcet winner is unclear with generalized that schedules.

[19] We omit to represent $BE(\tau)$ in the graph since we show below that a quota is never preferred to a tax.

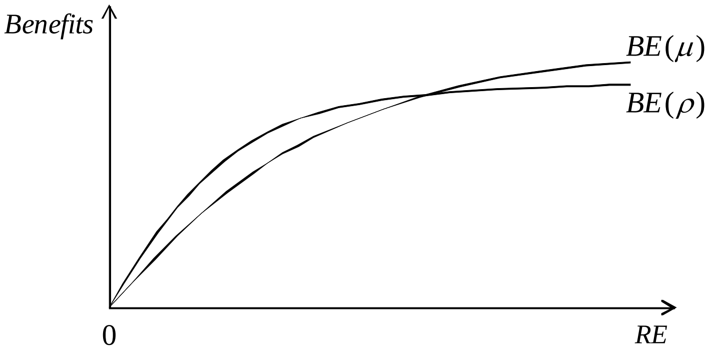

Figure 2: Benefits from rules and taxes

private costs. In other words, any i enjoys a "zero-cost area" up to $RE(\rho = t_i)$. Above that level, i has to reduce his behavior bearing a compliance cost. For a low type a rule is a cheap way to generate large amounts of RE: the zero-cost area is extended and the cost of additional RE is shared among many individuals. In fact, marginal costs are inversely proportional to the affected population.[20] Vice versa a rule can be quite costly for a high type. Denote with $CE(\rho, t_i)$ the cost function of type i. Figure 3 shows a possible shape when t_i is in an intermediate position.

2. Call $CE(\tau, t_i)$ the cost of RE when the instrument is a quota. There are no zero-cost areas in this case. The cost is shared more equally over the population, thus high types make smaller private sacrifices with respect to the rule, at least for sufficiently large amounts of RE. Low types pay more for a quota than for a rule due to the extended zero-cost area. Since revenue equivalents come from proportional adjustments, marginal cost of RE are proportional to types.[21] A possible shape is in Figure 3. When t_i increases the curve shifts upwards proportionally.

3. Let $CE(\mu, t_i)$ be the cost of RE generated by a tax. Since individuals

[20] Let us give some more details on this. Consider that, for any $\rho < t_i$,

$$\frac{\partial CE(\rho, t_i)}{\partial RE(\rho)} = \frac{\partial c(|\rho - t_i|)}{\partial \rho} \cdot \frac{\partial \rho}{\partial RE}$$

Since,

$$-\frac{\partial RE}{\partial \rho} = 1 - F(\rho)$$

then,

$$\frac{\partial CE(\rho, t_i)}{\partial RE(\rho)} = -\frac{1}{1 - F(\rho)} c'(|\rho - t_i|)$$

The marginal cost of RE is rather high when the affected population is low.

[21] More precisely, observe that $RE(\tau) = \tau \cdot \bar{t}$, where $\bar{t}$ is the average of $F(t)$. Since,

$$\frac{\partial CE(\tau, t_i)}{\partial RE(\tau)} = \frac{\partial c(\tau t_i)}{\partial \tau} \cdot \frac{\partial \tau}{\partial RE}$$

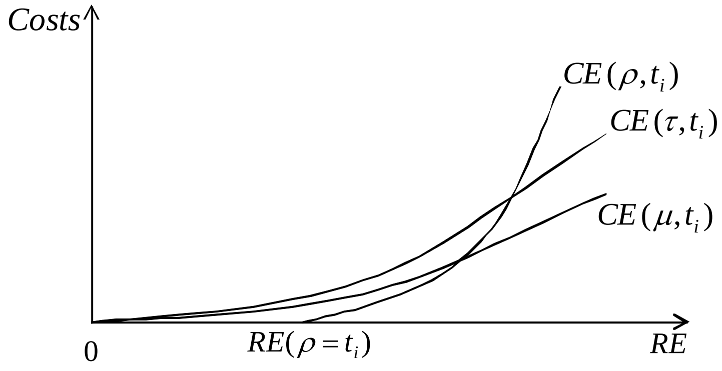

Figure 3: Costs from rules, quotas and taxes

minimize costs by balancing behavior and tax, for any RE and any type the private cost of a tax is always lower (and flatter) than the cost of a quota. A possible $CE(\mu, t_i)$ is in Figure 3. How is $CE(\mu, t_i)$ related to t_i? Let us assume that the sacrifice of paying a tax is comparable to the cost of reducing behavior, an assumption that we will keep hereafter. In this case

then,

$$\frac{\partial CE(\tau, t_i)}{\partial RE(\tau)} = \frac{t_i}{\bar{t}} c'(\tau t_i)$$

Any marginal increase in RE costs individual i an amount that is proportional to the ratio between his type and the average type.

marginal costs are proportional to types.[22]

We can now look at the choice of the policy instrument. Consider that a quota is always more costly than a tax and, due to the double dividend, it is also less beneficial. Thus a quota is never preferred to a tax. The instrument choice can be restricted to a binary comparison between a rule and a tax.

6.4 Existence of a Condorcet winner: simultaneous voting

Let us begin with the case in which voting takes place simultaneously on an issue that has two dimensions. The first is discrete and binary: either a tax or a rule. The second one, which is continuous, is the level of the instrument. Choosing the level implies choosing the amount of RE. Thus we will refer to the second dimension as the choice of RE. We can ensure the existence of a Condorcet winner if the equilibrium is a median "in both dimensions" (Davis, DeGroot and Hinich, 1972; Banks, Duggan and Le Breton, 2006). Observe that, as regards RE, preferences are single peaked and the median bliss points belong to the median type both in the case of a tax and in the case of a rule.

Thus an equilibrium exists as far as all individuals either above or below the median voter prefer, for any RE, the same instrument preferred by the median. For example, if the median prefers a rule because it is the best instrument to generate RE *and* all lower types prefer a rule too, the median

[22] In fact, $RE(\mu) = \bar{t} - (1-\mu) \cdot \bar{b}(\mu)$, where $\bar{t}$ is the average of $F(t)$, and $\bar{b}(\mu)$ is the average after-tax behavior. We can write

$$\frac{\partial RE}{\partial \mu} = \bar{b} - (1-\mu) \cdot \frac{\partial \bar{b}(\mu)}{\partial \mu}$$

Consider individual i. If paying taxes and changing behavior has the same impact on costs (i.e., $c(.) = d(.)$), then we can write:

$$\omega(., t_i) = c(t_i - b(\mu, t_i)) + c(\mu b(\mu, t_i))$$

Thus,

$$\frac{\partial \omega(., t_i)}{\partial \mu} = \left[b_i - (1-\mu) \cdot b_i'\right] \cdot c'(.)$$

where $b_i = b(\mu, t_i)$ comes out of (12). Since,

$$\frac{\partial CE(\mu, t_i)}{\partial RE(\mu)} = \frac{\partial \omega(., t_i)}{\partial \mu} \cdot \frac{\partial \mu}{\partial RE}$$

then

$$\frac{\partial CE(\mu, t_i)}{\partial RE(\mu)} = \frac{b_i - (1-\mu) \cdot b_i^{'}}{\bar{b} - (1-\mu) \cdot \bar{b}'} \cdot c'(.)$$

We can reasonably assume that the first term in the right-hand side of this equation is constant, and proportional to $\frac{t_i}{\bar{t}}$, thus

$$\frac{\partial CE(\mu, t_i)}{\partial RE(\mu)} = \frac{t_i}{\bar{t}} \cdot c'(.) \tag{16}$$

We will use this result to prove Lemma 4 below.

instrument is a rule and the median level is ρ_s^*. Thus the Condorcet winner is $RE(\rho_s^*)$.[23]

While it is true that low types are inclined towards rules and high types towards taxes, we do not necessarily expect that *all* voters under a given type prefer a rule and all those above it prefer the tax. Nonetheless, we can give a sufficient condition for it. The idea is simple: take the median type t_s. Assume that he prefers a rule; i.e., $BE(\rho_s^*, t_s) - CE(\rho_s^*, t_s) > BE(\mu_s^*, t_s) - CE(\mu_s^*, t_s)$. Consider lower types. They prefer more RE, thus they compare additional (i.e. marginal) benefits and costs for both instruments. Assume that the marginal benefits of the two instruments are the same.[24] Thus the choice is based on marginal costs of RE. We want that for all types under the median a rule is the cheapest way to generate more RE.[25] A sufficient condition is that for any lower type the marginal cost of a rule decreases more than the marginal cost of a tax. Consider that if the marginal cost of a rule is large then it is "more sensitive" to type decrease. Thus the idea is that if the median prefers a rule *and* his marginal cost is relatively large, lower types have stronger incentives to prefer a rule since in this case their costs decrease by a larger amount.

Our *sufficient* condition for monotonicity in both dimensions is: first, the median prefers a rule and the marginal cost of a rule is sufficiently large; second, this requisite on costs applies to all types below the median. Lemma below takes care of both parts of this condition.

Lemma 4 *a) If the median prefers a rule to a tax, if marginal benefits of rule and tax are the same and if*

$$1 - F(\rho_s^*) < \frac{\bar{t}}{t_s} \tag{17}$$

then a marginally lower type prefers a rule to a tax.
b) If for any $t_i < t_s$

$$1 - F(\rho_i^*) < \frac{\bar{t}}{t_i}$$

then all types below the median prefer a rule to a tax.

Inequality (17) may be explained with an example. Suppose that $(1 - F(\rho_s^*)) = 0.5$ and $\bar{t}/t_s = 1$. Recall that for the median the marginal cost of a rule is proportional to $1/(1 - F(\rho_s^*)) = 2$, and the marginal cost of a tax is proportional to $t_s/\bar{t} = 1$. Thus a marginal decrease in t_s will cause a

[23] We focus on this kind of equilibrium. The reader can easily infer the conditions such that the alternative equilibrium is $RE(\mu_s^*)$.

[24] The assumption of comparable marginal benefits is not unrealistic because, due to the concavity of ε, substantial differences in marginal benefits occur only for small RE.

[25] The reader may observe that a similar argument supports Grandmont's (1978) sufficient condition for the existence of a Condorcet winner in multidimesional policy problems.

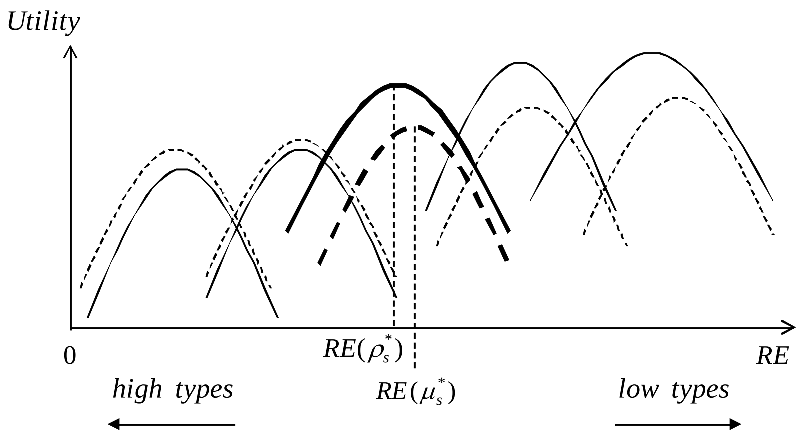

Figure 4: Monotonic preferences in both dimensions

reduction in marginal cost that is double in the case of a rule than in the case of a tax. Since marginal benefits for a rule and a tax are the same, then an individual that is marginally lower than the median chooses the rule. Observe that (17) introduces a "local" requisite that concerns only a marginal decrease in t_s. Part b) of Lemma 4 extends the requisite to all types below t_s.

The voting stage is in Figure 4. The x-axis is the second dimension of the policy space, the set of RE.[26] The first dimension is binary. Thus we have two kinds of policy preferences: the dashed curves are the preferences in case of a tax; the solid curves represent the rule preferences. The median's curves are bold. Lower types prefer more RE both under a rule and under a tax; i.e., bliss points are (inversely) monotonic in RE and the median bliss points belong to the median type in both cases. Under Lemma 4, the median prefers a rule. Moreover, for all types below him, a rule is the preferred instrument to generate additional RE: the rule preference functions are higher. It is easy to see that in this case the Condorcet winner is $RE(\rho_s^*)$.

What makes this equilibrium more likely to occur? Recall that a rule can be rather costly when the median is relatively high. Thus a rule is likely to emerge in equilibrium when the median is low. Moreover, if the double dividend is low and externalities are highly concave, the benefits of a rule compensate the costs. Thus a rule is likely when the externality problem is somehow more relevant than the provision of a public good, and the main concern is curbing top polluters' behaviors.

Proposition 5 *If the median is relatively low, the double dividend is low*

[26]Recall that people do not vote on RE, directly. However, by choosing the level of an instrument, they uivocally vote on RE.

and high behaviors produce large negative externalities, and conditions in Lemma 4 apply, then the majority prefers a rule and selects the median's most preferred one.

Interestingly, when the median is below the average and he prefers a rule, the Condorcet winner is "almost sure", in the sense that we have only to ensure that the second part of Lemma 4 applies.

Corollary 1 *If the median is below the average and Lemma 4-b) applies, then the majority always prefers a rule.*

Vice versa, when the median is in a relatively high position, the double dividend is important and consistent pollution is due also to the behavior of low types, then a tax is more likely to prevail.

Which is the composition of a majority that prefers a rule? Arguably there are extremely low types. Their instrument choice is determined only by costs concerns: with a rule they can induce at zero or very low cost virtuous behavior by others. There might also be intermediate types, who are available to pay a possibly large private cost because they look at the rule as a more efficient way to cope with the externality problem.

6.5 Existence of a Condorcet winner: sequential voting

We now assume that voting takes place sequentially: first, the majority selects the instrument; then it chooses its level. In this case, a Condorcet winner always exists.[27] Individuals know that, whatever the instrument, the level that will pass at the second stage is the one preferred by the median, either ρ_s^* or μ_s^*. The revenue equivalent outcome is either $RE(\rho_s^*)$ or $RE(\mu_s^*)$. All individuals compare their indirect utilities in those two cases, and choose which instrument to vote for. Since the choice at the first stage is binary, there will always be a majority that prefers either one of the instruments. No scope for strategic vote. This majority behaves as a Stackelberg leader: it selects the instrument and it lets another majority choose the level. Interestingly, the two majorities are possibly different since we do not impose any monotonicity condition at the first stage. The median's most preferred level always passes, but not always his most preferred instrument. This is the main difference with simultaneous voting.

What makes $RE(\rho_s^*)$ more ore less likely than $RE(\mu_s^*)$? The general idea is the same as for the simultaneous case: low types tend to prefer the rule and high types prefer the tax. Thus, a rule is more likely to pass when the distribution is slanted towards high types, externalities are quite concave and the double dividend is low.

[27] Usually, the equilibrium is sensitive to the voting sequence (De Donder, Le Breton and Peluso, 2010). In this case we do not have such a problem since the inverse sequence in which the instrument is decided *after* its level is unnatural.

6.6 Normative aspects

Consider now the policy benchmark. A proportional tax is socially preferable, except when externalities are very concave and the double dividend is low. In fact, a rule is always "socially" more costly than a tax. The first reason is cost convexity: a rule concentrates the burden of RE on higher types, who also have higher costs. The second reason is cost minimization: a tax allows individuals to choose if paying the tax or reducing behavior. As for benefits, a tax is not always the best instrument. It has the advantage of the double dividend, whereas a rule produces larger benefits when externalities are highly concave, since it affects top polluters. Thus, only in the case of quite concave externalities and rather low double dividend the social planner prefers a (socially costly) rule. In that case, however, the optimal rule would not be very restrictive, since its main task would be reducing the behavior of high types to moderate levels. In all other cases a proportional tax performs better: larger benefits from the double dividend and lower social costs. It is worth repeating that without any restrictions on the shape of the tax schedule a tax would always be superior to other policy instruments.

How is this normative conclusion related to the positive results above? A general idea is that the if the majority chooses a rule when the social planner would choose a tax, then a too restrictive rule is very likely. In fact, suppose that the socially optimal instrument is a tax. If the distribution is slanted towards high types, there will be sufficient support for selecting a rule and, under the same conditions, the rule is rather restrictive. Vice versa, if the socially optimal instrument is a rule, the majority selects the wrong instrument (a tax) only when the median is in a relatively high position. In this case, the tax level chosen by the majority is too low.

In Appendix 8.2 we show that this kind of distortions might be lower when tax revenues are redistributed lump sum.

7 Conclusions

We have examined the political economy of how to curb activities which generate negative externalities. We showed two things. First for given policy tool (rules, quota or taxes) the median voter almost never chooses the efficient level of the policy instrument. Both too lenient or too harsh policies against the activity with negative externalities are possible in a voting equilibrium, and we characterized under which circumstances one outcome or the other occurs. Too stringent restrictions (or taxes) occur when the activity with negative externality is concentrated on a relatively small fraction of the population and when negative externalities grow more than proportionally with the level of the polluting activity. In the opposite case instead regulation or taxation would be too lenient. Second we showed that when the majority can also choose the policy instrument it would not necessarily choose the most efficient one. For instance, taxing the activity with negative externality is in general superior to rules or quotas, but the majority may prefer a rule in order to charge the minority a larger share of the adjustment burden.

One could explore several extensions. First, some activities with negative externalities (but not all) impose cost on future generations who do not vote, at least not directly except for the intergenerational altruism of parents toward their children. One could discuss how these considerations would influence the median voter.

Second, one could try to extend the analysis to more sophisticated tax schedules allowing for some curvature in the tax rate. Our hunch is that when the population is concentrated on low types the majority would choose a more "progressive" tax than the social planner. In a sense even the outcome of a rule can be obtained by a zero tax rate up to the limit of the rule and a prohibitive tax rate above it. But presumably given the double dividend the majority could do even better perhaps by allowing a little bit of pollution but extracting a lot of tax revenues from the high types. Once again the solution of this problem is complex because the voting would take place on a multidimensional space including level and curvature of the tax rate.

Third, thus far we have imposed that rules and quotas are followed by all. This equilibrium is equivalent to assuming perfect monitoring (or imperfect monitoring with such a high fine if caught that nobody cheats in equilibrium). In reality, rules can be broken. In general the social choice involves a certain amount of investment in costly monitoring activities and the selection of a fine. The revenue from the fine could be used to finance monitoring and, if anything is left over, to provide public goods. With imperfect monitoring and a fine, individual polluters would choose how much to pollute and how much risk of being caught is worth taking. This would lead to a less sharp distinction between a rule (or a quota) and a tax. This distinction would become even less stark if tax avoidance or evasion is also allowed. An interesting result which would most likely hold is that for certain parameter values the median voter would not choose prohibitive fines for polluting activities if one considers the revenues obtained from collecting fines. This extension would imply adding several parameter values: the monitoring technology, risk aversion, expenditure of fine proceeds.

The fourth extension relates to voting rules. In our model any possible form of tyranny does not come from direct expropriation of the minority but rather from the fact that, within the political process, the majority ignores the costs incurred by the minority. This may result in decisions that are too costly from a social viewpoint. If for example the median's policy were too restrictive, efficiency would be enhanced by giving the minority of high types some amount of blocking power. This is frequently done by adopting super-majorities.[28] The problem is that a super-majority assigns blocking power not only to high types, but also to low types. In our model with single-peaked policy preferences even policies that are more restrictive and inefficient than the median's one may emerge in equilibrium (Black, 1948b). If the objective is avoiding that the median is the pivot, a super-majority does not make the job. A potential alternative is giving the minority more voting weight.[29] The idea is simple: when the median's policy is too restrictive we must "shift the pivot" towards a higher type, whose bliss point is at the socially efficient level. This can be done by keeping simple majority and assigning types above the desired pivot a mass of votes equal to the votes held by those who are below the pivot.[30] In this case the equilibrium is unique and it is socially efficient. The issue here is not equity: assigning more power to the most concerned individuals in order to counter balance the power of the least concerned ones improves efficiency. Implementation problems of such schemes are however extremely severe.

[28] Literature on super-majorities is vast and belongs to the normative analysis of constitutions. The focus is mainly on distributional issues (see Mueller (2003) for an extensive survey). Aghion and Bolton (2003) suggest that, when preferences are not single-peaked, higher super-majorities lower the risk of Condorcet cycles, but also lower the chance of circumventing ex-post vested interests; the solution of this trade-off yields the optimal majority threshold. Dixit, Grossman and Gul (2000) argue that super-majority rules may reduce compromise; as a consequence, the incidence of majority tyranny may increase. Aghion, Alesina and Trebbi (2004) analyze the constitutional choice about the level of supermajority needed to block policies of elected political leaders.

[29] The literature on weighted voting is possibly less developed, and mostly concerned with problems of equal representation in indirect democracies. Barbera and Jackson (2006) suggest a mixture of weights and super-majority that allows sticking with the status quo, unless at least a threshold of weighted votes is cast for change.

[30] Consider the case of a too restrictive rule. If the optimal level is ρ^* we must assign each individual whose bliss point is equal or larger than ρ^* additional weight, by multiplying their votes by the following factor:

$$1 + \frac{2F(\rho^*) - 1}{1 - F(\rho^*)}$$

For example, if the optimal rule corresponds to the individual in the third quartile position ($F(\rho^*) = .75$) we must give each higher type three votes instead of one.

8 Appendix

8.1 Proofs of Lemmas and Propositions

Lemma 1 *For any two individuals i and j, if $t_i > t_j$ then $\rho_i^* \geq \rho_j^*$.*

Proof. By implicit differentiation of equation (6) we get, for any i,

$$\frac{\partial \rho_i^*}{\partial t_i} = -\frac{-c''(|\rho - t_i|)}{\varepsilon''(\rho)\,(1 - F(\rho)) - \varepsilon'(\rho) f(\rho) - c''(|\rho - t_i|)} \tag{18}$$

The denominator in the right-hand of (18) is the second derivative of $U_i(\rho)$, which is negative by assumption. Thus, for any $\rho_i^* \in (0, t_i)$, the sign of $\frac{\partial \rho_i^*}{\partial t_i}$ is positive, since $c''(|\rho - t_i|) > 0$. The relationship between type and bliss point is weakly monotone in order to account for corner bliss points (i.e., $\rho_i^* = \rho_j^* = 0$, for some $t_i > t_j$). ■

Proposition 1 *i) Simple majority voting yields a too restrictive (too permissive) rule if and only if in equilibrium the ratio between the median voter's marginal cost and the average marginal cost of the affected population is lower (higher) than the share of the affected population.*
ii) Majority voting yields the socially optimal rule if and only if the ratio between those marginal costs equals the share of the affected population.

Proof. Let us prove part i) of the Proposition. Consider the case of a too restrictive rule. Recall that $\hat{\rho}^*$ solves (7), and that the $ac'(\rho)$ curve crosses the $\varepsilon'(\rho)$ curve 'from below'. Since $\hat{\rho}_s < \hat{\rho}^*$, we have

$$(1 - F(\hat{\rho}_s)) \cdot \varepsilon'(\hat{\rho}_s) > ac'(\hat{\rho}_s)$$

and

$$(1 - F(\hat{\rho}_s)) \cdot \varepsilon'(\hat{\rho}_s) = c'(|\hat{\rho}_s - t_s|)$$

Therefore,

$$c'(|\hat{\rho}_s - t_s|) > ac'(\hat{\rho}_s)$$

or, since $ac'(\hat{\rho}_s)$ is negative,

$$\frac{c'(|\hat{\rho}_s - t_s|)}{ac'(\hat{\rho}_s)/(1 - F(\hat{\rho}_s))} < 1 - F(\hat{\rho}_s)$$

where $ac'(\hat{\rho}_s)/(1 - F(\hat{\rho}_s))$ is the average marginal cost computed over the affected population.

Equivalently, the condition for a too permissive rule emerging is the following:

$$\frac{c'(\hat{\rho}_s - t_s)}{ac'(\hat{\rho}_s)/(1 - F(\hat{\rho}_s))} > 1 - F(\hat{\rho}_s)$$

The proof of the part ii) of the Proposition is straightforward. ■

Lemma 2 *For any two individuals i and j, if $t_i > t_j$ then $\tau_i^* \leq \tau_j^*$.*

Proof. The proof parallels the proof of Lemma 1. By implicit differentiation of (9) we get

$$\frac{\partial \tau_i^*}{\partial t_i} = -\frac{-c'(|-\tau t_i|) - \tau t_i c''(|-\tau t_i|)}{\frac{\partial a\varepsilon'(b)}{\partial \tau} + t_i^2 c''(|-\tau t_i|)} \tag{19}$$

The denominator in the right hand of (19) is the SOC, which is negative by assumption. The first term at the numerator is the impact of an increase in t_i on the marginal adjustment cost, which is negative; the second one is negative, by convexity of the cost function. Thus, $\frac{\partial \tau_i^*}{\partial y_i}$ is negative. Monotonicity is weak because of corner bliss points. ■

Proposition 2 *Simple majority voting yields a too restrictive (too permissive) quota if and only if in equilibrium the median voter's marginal cost is lower (higher) than the average marginal cost.*

Proof. Let us consider the case of a too restrictive quota, $\hat{\tau}_s > \hat{\tau}^*$. From (9) and from (10), we have:

$$a\varepsilon'(\hat{\tau}_s) < ac'(\hat{\tau}_s)$$

and

$$a\varepsilon'(\hat{\tau}_s) = t_s c'(|-\hat{\tau}_s t_s|)$$

Therefore,

$$\frac{t_s c'(|-\hat{\tau}_s t_s|)}{ac'(\hat{\tau}_s)} < 1$$

The vice versa holds for a too low quota. ■

Proposition 3 *When the majority selects a rule that is too permissive, then it selects a quota that is also too permissive. The vice versa is not true.*

Proof. If an interior rule is too permissive then

$$\frac{c'(|\hat{\rho}_s - t_s|)}{ac'(\hat{\rho}_s)} > 1$$

We want to show that, for any τ,

$$t_s c'(\tau t_s)/ac'(\tau) > 1$$

Consider that $ac'(\hat{\rho}_s)$ is an average in which the only non-zero elements are the $(1 - F(\hat{\rho}_s))$ marginal costs of the affected people above $F(\hat{\rho}_s)$; where $1 - F(\hat{\rho}_s) > 0.5$. Moreover 50% of these elements are larger than $c'(|\hat{\rho}_s - t_s|)$. Further consider that when a quota is adopted, marginal costs are more evenly distributed across the population. This means that, for any τ, 50% elements above the average and above the median's marginal cost enter

$ac'(\tau)$ with lower values, which continue to be above the median's marginal cost; moreover $F(\hat{\rho}_s)$ elements enter $ac'(\tau)$ with a non-zero value that is in any case below the median's marginal cost. Thus, it might be the case that $ac'(\tau) > ac'(\hat{\rho}_s)$, but $ac'(\tau)$ cannot be larger that the median's marginal cost, $t_s c'(\tau t_s)$. ■

Lemma 3 *For any two individuals i and j, if $t_i > t_j$ then $\mu_i^* \leq \mu_j^*$.*

Proof. For any tax rate, individuals set their behavior in order to satisfy condition (11). It is easy to see that, if t_i increases, this condition is satisfied for higher marginal costs. Moreover, for any t_i, if μ increases, then condition (11) is satisfied for higher marginal costs. Thus, total costs are convex in μ, and, for any μ, marginal costs are higher for higher types. Observe that benefits, i.e. the first two terms in the right-hand side of (13), are independent of the type. Moreover, the concavity of (13) in μ implies that the marginal cost curve crosses the marginal benefit curve from below. Thus, as type increases, the crossing point, i.e. the optimal level of μ, decreases. ■

Lemma 4 *a) If the median prefers a rule to a tax, if marginal benefits of rule and tax are the same and if*

$$1 - F(\rho_s^*) < \frac{\bar{t}}{t_s} \tag{20}$$

then a marginally lower type prefers a rule to a tax.
b) If for any $t_i < t_s$

$$1 - F(\rho_i^*) < \frac{\bar{t}}{t_i}$$

then all types below the median prefer a rule to a tax.

Proof. Consider part a) of the Lemma. Individual s's optimal choice of a rule implies that

$$\frac{\partial BE(\rho^*)}{\partial RE} = \frac{1}{1 - F(\rho^*)} c'(t_s - \rho^*)$$

Call $RE(\rho^*)$ the revenue equivalent when the optimal rule is applied. Let μ^0 the tax level that yields the same revenue equivalent, and call $RE(\mu^0)$ that amount. By assumption, $\frac{\partial BE(\rho^*)}{\partial RE} = \frac{\partial BE(\mu^0)}{\partial RE}$, thus,

$$\frac{1}{1 - F(\rho^*)} c'(t_s - \rho^*) = \frac{t_s}{\bar{t}} \omega'(., \mu^0)$$

where $\omega'(., \mu^0) = \left[b_i - (1 - \mu^0) \cdot b_i'\right] \cdot c'(.)$, as in Section 6. A marginal decrease in t_s results in the marginal changes (namely, marginal decreases) of $c'(t_i - \rho)$ and $\omega'(.)$. We can reasonably assume that their amounts are the same (i.e. the curvature of c is rather constant). Thus if

$$\frac{1}{1 - F(\rho^*)} > \frac{t_s}{\bar{t}}$$

i.e., if condition (20) is satisfied, then a lower type "saves" in marginal costs more if he chooses a rule than if he chooses a tax. Thus it cannot be that a t_s prefers a rule and a marginally lower type prefers a tax. Part b) can be proved applying the proof of part a) recursively. ■

Proposition 5 *If the median is relatively low, the double dividend is low and high behaviors produce large negative externalities, and conditions in Lemma 4 apply, then the majority prefers a rule and selects the median's most preferred one.*

Proof. The median prefers a rule when his position is low, externalities are quite concave and the double dividend is small. Conditions in Lemma 4 are sufficient to ensure that the median bliss points in both voting dimensions belong to the median type. Thus the Condorcet winner is $RE(\rho_s^*)$. ■

Corollary 1 *If the median is below the average and Lemma 4-b) applies, then the majority always prefers a rule.*

Proof. Observe that if the median is below the average, then $\frac{\bar{t}}{t_s} > 1$. Thus condition (20) is always satisfied. We only need that part b) of Lemma *4* applies. ■

8.2 Transfers rather than public goods

We assume here that the tax proceeds are redistributed with lump sum transfers to individuals and no public good is provided. We show that making fixed transfers instead of providing a public good does not change qualitatively the nature of our results.

Each individual receives out of a balanced public budget a transfer which amounts to the average tax burden, $\mu \cdot \bar{b}$. Individual i's indirect utility is:

$$U_i(\mu) = \int_0^1 \varepsilon(b)dG(b) - \omega\left(., \mu, t_i\right)$$

with $\omega\left(., \mu, t_i\right) = c(|b_i - t_i|) + d(\mu(b_i - \bar{b}))$. Individual optimization leads to:

$$-c'(|b_i - t_i|) = \mu \cdot d'(\mu(b_i - \bar{b})) \qquad (21)$$

Transfers reduce the marginal cost of paying taxes (i.e., $d'(\mu(b_i - \bar{b}) < d'(\mu b_i))$. This effect on costs replaces the “double dividend” of Section 5. Thus, on the one hand, private marginal benefits are lower since no public good is supplied but, on the other hand, marginal costs are smaller. While the former effect is the same for all individuals, the latter is stronger for higher types, due to cost convexity. Below we show that this plays an important role on the efficiency of the voting outcome.

Bliss points are monotonically related to types and the voting outcome is the tax rate preferred by the median. The benchmark would be a tax rate such that the per capita marginal benefit is equal to the per capita (average) marginal costs. Similarly to the other instruments, a possible inefficiency is due to a discrepancy between the median and the average marginal cost: high cost convexity and rightward slanted distributions favor too high tax rates. Observe however that transfers “smooth down” differences among individual marginal costs. This means that, with respect to the case with public good, the median's marginal cost is closer to the average. The median chooses a tax rate that is closer to the one that would be chosen by the social planner: the possible inefficiency is smaller. As for the risk of too restrictive or too permissive outcomes, a result in Section 5 applies: if d is more convex than c then the tax with transfers is more likely to be too restrictive with respect to a quota, and it is more similar to a rule. Consider now benefits and costs as a function of RE. Due to the absence of the double dividend benefits are lower with transfers than with a public good, but also costs are lower since individuals pay a smaller net amount of money for any level of RE. Thus, if the utility from the public good is low or if the sacrifice of paying taxes is large, a tax with transfers is more attractive than a tax with a public good.

Let us look at the instrument choice. A tax with transfers is superior to a quota since for any amount of RE individuals optimize costs by choosing behavioral reductions and tax payments on residual behaviors. Thus the choice is between the tax and the rule. Our analysis in Section 6 does not change substantially: if the instrument/level choice is simultaneous the condition for a Condorcet winner is rather simple; if the choice is sequential a winner always exists. Let's consider the simultaneous choice. The marginal effect of tax increase on RE is the same as for the case with public good:

$$\frac{\partial RE}{\partial \mu} = \bar{b} - (1-\mu) \cdot \bar{b}'$$

where $\bar{b}$ is the after-tax average behavior and $\bar{b}' = \frac{\partial \bar{b}(\mu)}{\partial \mu}$. The marginal cost is lower. As in Section 6 assume that the sacrifice of changing behavior equals the cost of paying taxes ($c(.) = d(.)$). We can write:

$$\frac{\partial \omega(., t_i)}{\partial \mu} = \left[b_i - \bar{b} - (1-\mu) \cdot b_i' - \mu \bar{b}'\right] \cdot c'(.)$$

where $b_i = b(\mu, t_i)$ solves (21). Then,

$$\frac{\partial CE(\mu, t_i)}{\partial RE(\mu)} = \frac{b_i - \bar{b} - (1-\mu) \cdot b_i' - \mu \bar{b}'}{\bar{b} - (1-\mu) \cdot \bar{b}'} \cdot c'(.)$$

If we assume that the first term in the right-hand side of this equation is proportional to $t_i/\bar{t}$, then marginal costs are proportional to types. Thus Lemma 4 applies. Accordingly, condition $1 - F(\rho_i^*) < \bar{t}/t_i$ (for any $t_i \leq t_s$) is sufficient to ensure that the rule is the majority's choice. Observe however that this condition is even too stringent when the marginal benefits of a rule are larger than the marginal benefits of a tax. This is the case when no public good is provided. This implies that a stable majority on the rule is easy to form. In synthesis, with transfers it is less likely that a median prefers a rule, but once this happens the chance that a stable majority forms on a rule is higher. As for the sequential case, the argument of Section 6.5 holds. Since the choice is binary, there will always be a majority that behaves as a Stackelberg leader: it chooses either a tax with transfers or a rule, based on the level that will be voted at the second stage.

Summing up, a first result is that with transfers the majority chooses a less inefficient tax level because median and average marginal costs are closer. A second result is that with transfers the majority is more inclined towards a tax, then it is more likely to prefer the "right" instrument. However, a very low median prefers the rule in any case. In this case a stable majority on the rule is easier to form and the level is quite high (a too restrictive rule). On the contrary, a high median prefers a tax with transfers but, if the optimal instrument is the rule, inefficiency is lower with respect to the case with public good.

References

[1] Aghion, Philippe, Alberto Alesina, and Francesco Trebbi, (2004): "Endogeneous Political Institutions", *The Quarterly Journal of Economics*, 119 (2), pp. 565–611.

[2] Aghion, Philippe and Patrick Bolton (2003): "Incomplete Social Contracts", *Journal of the European Economic Association*, No. 1, pp. 38-67.

[3] Banks, Jeffrey S., John Duggan and Michel Le Breton (2006): "Social Choice and Electoral Competition in the General Spatial Model", *Journal of Economic Theory*, 126, pp. 194-234.

[4] Barbera, Salvador and Matthew O. Jackson (2006): "On the Weights of Nations: Assigning Voting Weights in a Heterogeneous Union", *Journal of Political Economy*, Vol. 114, pp. 317-339.

[5] Black, Duncan (1948): "On the Rationale of Group Decision-making", *Journal of Political Economy,* Vol. 56, pp. 23-34.

[6] Black, Duncan (1948b): "The Decisions of a Committee Using a Special Majority", *Econometrica*, Vol. 16, No. 3, pp. 245-261.

[7] Boyer, Marcel and Jean-Jacques Laffont (1998): "Toward a political theory of the emergence of environmental incentive regulation", *Rand Journal of Economics*, Vol. 30, No. 1, pp. 137-157.

[8] Buchanan, James M. and Gordon Tullock (1975): "Polluters' Profits and Political Response: Direct Controls Versus Taxes", *American Economic Review,* 65, pp. 139-147.

[9] Congleton, Roger D. (1992): "Political Institutions and Pollution Control", *The Review of Economics and Statistics*, Vol. 74, No. 3, pp. 412-421.

[10] Cremer, Helmut, Philippe De Donder and Firouz Gahvari (2004): "Political sustainability and the design of environmental taxes", *International Tax and Public Finance*, 11, pp. 703–719.

[11] Cremer, Helmut, Philippe De Donder and Firouz Gahvari (2008): "Political competition within and between parties: an application to environmental policy", *Journal of Public Economics*, 92, pp. 532–547.

[12] Davis, Otto A., Morris H. DeGroot and Melvin J. Hinich (1972): "Social Preference Orderings and Majority Rule", *Econometrica*, 40, pp. 147-157.

[13] De Donder, Philippe, Michel Le Breton and Eugenio Peluso (2010): "Majority Voting in Multidimensional Policy Spaces: Kramer-Shepsle versus Stackelberg", *IDEI Working Papers,* 593.

[14] Dixit Avinash, Gene M. Grossman and Faruk Gul (2000): "The Dynamics of Political Compromise", *Journal of Political Economy*, Vol. 108, pp. 531-568.

[15] Grandmont, Jean-Michel (1978): "Intermediate preferences and the majority rule", *Econometrica*, 46, pp. 317–330.

[16] Hepburn, Cameron (2006): "Regulation by Prices, Quantities, or Both: a review of instrument choice", *Oxford Review of Economic Policy*, Vol. 22, No. 2, pp. 226-247.

[17] Meltzer, Allan H. and Richard, Scott F. (1981): "A Rational Theory of the Size of Government", *Journal of Political Economy*, vol. 89(5), pp. 914-927.

[18] Mueller, Dennis C. (2003): *Public Choice III*, Cambridge, Cambridge University Press.

[19] Persson, Torsten and Guido Tabellini (2000): *Political Economics: Explaining Economic Policy,* Cambridge MA, The MIT Press.

[20] Roberts, Kevin W. S. (1977): "Voting over income tax schedules", *Journal of Public Economics*, vol. 8(3), pp. 329-340.