Regulation versus Taxation*

Alberto Alesina[†] Francesco Passarelli[‡]

July 2011

Abstract

We study which policy tool and at what level is chosen by majority voting in order to reduce negative externalities, such as pollution. We consider three instruments: a rule, that sets an upper limit to the activity which produces the negative externality, a quota that forces a proportional reduction of the activity, and a proportional tax on it. For all instruments the majority chooses levels which are too restrictive when the activity is performed mainly by a small fraction of the population, and when costs for reducing activities or paying taxes are quite convex. However, in case of a rule too restrictive levels are more frequent than in case of a tax or a quota.

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.

Keywords: Externalities, Pollution, Voting, Rules, Taxation.

JEL Codes: H23, P16, D62, D72, L51, K32.

.

*We thank Pierpaolo Battigalli, Allan Drazen, Firouz Gahvari, Vincenzo Galasso, Erzo Luttmer, Eugenio Peluso, Andrei Shleifer, Jim Snyder, Pierre Yared and seminar partcipants at Bocconi, Harvard and the 2010 NBER Summer Institute for useful comments and suggestions.

[†]Harvard University and Igier Bocconi.

[‡]University of Teramo and Bocconi University.

1 Introduction

There are three ways of reducing the level of an activity generating negative externalities which are routinely used: a rule that sets an upper bound to this activity, a compulsory proportional reduction for everybody of the activity or a proportional tax on it.[1] 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, in addition, taxes generate revenues which could be used, for instance, to provide useful public goods.

This paper adopts 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 yield 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. When this occurs, the same majority has also an incentive to set a restrictive level of the rule.

We can relate the nature of the departure from optimally 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, a sort of tyranny of the majority. Thus we should observe restrictive rules for activities enjoyed by some specific individuals/firms/sectors or for very large "polluters", while rules might be too lax for activities that a majority enjoys.[2]

This is consistent with the evidence that in many cases a host of rules are introduced

[1] A fourth way is tradeable permits. We do not study them in the present paper because doing so would require a different model in which political distortions eventually emerge in the definition of property rights; i.e. who has the right to pollute and who has to pay whom.

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

while taxation is not always chosen as the preferred policy instrument. For example, in agriculture the limits in the use of pesticides are quite frequent whereas taxes on them are rare. In the case of air pollution, there is a sharp contrast between taxes and emission standards. The latter are preferred when polluters are concentrated in specific industries or plants, such as emissions of pollutants by power generation industries or by steel and cement makers.[3] Low level of taxation are chosen instead when the polluters are the majority; an example would be low taxes on gasoline or heating fuel.[4]

We define our negative externality "pollution" but there is a sense in which this term should be considered a shortcut for more general issues. What we intend to capture are situations in which a majority has a chance of imposing its "standards" to the minority with rules of behavior limiting activities that the majority does not like. Examples range from purely economic situations, like technical standards for firms, to social situations, like noise in public contexts, housing construction rules, pets in condominiums, smoking rules, etc. On the other hand 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 dissipate rapidly and only affect the currently alive.

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 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

[3] On November 22, 2010 the The Wall Street Journal reported that since Mr. Obama took office, the US Environment Protection Agency (EPA) had proposed or finalized 29 major regulations and 172 major policy rules. Requiring energy sources to install *the best available control technology* to limit greenhouse gas emissions, would impose the electric industry costly capital expenditures to meet the increasingly strict burden.

[4] For a gasoline tax, the International Center for Technology Assessment computed that indirect costs to society total around \$12 per gallon (\$3.17 per liter). Once added to the production cost, the gallon price raises to \$15.

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.[5] On the government spending side we examine both the case in which revenues are used to produce a public good and the case in which the revenues are redistributed lump sum (in Appendix 8.2). As regards the sign of political distortions the results are qualitatively the same. We frame our model in the tradition of the political economy literature on redistributive fiscal policy;[6] but we focus upon rules and externalities rather than redistributive taxation.

The dilemma between regulation and taxation is old in the literature, but it has been mainly posed in a normative perspective. The idea that the two instrument perform differently when uncertainty regards either costs or benefits dates back to Weitzman (1974). The small literature which introduces political economy considerations in this area is confined to environmental issues.[7] 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 élite, and the rest of the population. They have to choose the amount of a costly environmental standard. Since the standard reduces aggregate income, the élite 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. Fluctuating 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

[5] We provide some normative details on linear and non-lnear taxes in Appendix 8.3.

[6] See the seminal works by Roberts (1977), Romer (1977) and Meltzer and Richard (1981). For a survey of this literature, see Persson and Tabellini (2000).

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

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.[8]

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 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, and Appendix 8.3 explores socially optimal taxation and private information.

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 cost. A behavior different 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, a type represents the profit maximizing level of polluting emission for a firm, the utility maximizing level of noise made by an individual, the utility maximizing speed adopted on a highway by a driver, the production of waste that maximizes an individual's utility. The adjustment cost function, c, is the same for all individuals:

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

[8]In addition, our argument is different from the *tollbooth* literature which holds that politicians use regulation to create rents and to extract them with campaign contributions (see Shleifer and Vishny, 1998, and references therein). This literature explains excess regulation in poor and corrupted countries. Our argument in this paper applies to a wider notion of regulation and it explains why its use may be eccessive also in advanced democratic societies.

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 social loss function is the total (per capita) externality produced in the society:

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

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. 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 thus the aggregate level of 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 this kind of 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. Examples abound, from pollution caps, performance standards, recycling rules, safety standards, speed limits, minimum contents of hazardous ingredients or components, smoking restrictions, typologies of behavior that are banned, i.e. a rule of zero for certain activities, etc.

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 select their behavior. The rule, $\hat{\rho}$, is fully enforced. All types higher than the rule have to adjust and pay the cost; 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. Formally, the concavity of $U_i(\rho)$ requires that 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. In fact, F needs to be "smooth" overall. By "smoothness" we mean that $f(t)$ must never be too high locally, otherwise, for some ρ, the marginal gains from reducing the rule might not be decreasing. Formally, $-\varepsilon'(\rho) \cdot f(\rho)$ in the right end side of (5), which is positive, might be too high due to large local density of new individuals affected. Under this condition, which we assume from now on, any individual i has a uniquely preferred rule and 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.[9] 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 $\rho_i^* \in [0, t_i)$. Moreover, lower types prefer lower rules:

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

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

$$\hat{\rho}_m = \rho_m^*$$

The socially optimal rule $\hat{\rho}^*$ in general differs from the voting outcome, $\hat{\rho}_m$. In fact $\hat{\rho}^*$ maximizes a "social" policy preference schedule, call it $W(\rho)$, that 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.[10] 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 $c'(|\rho - t_m|)$ 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

[9]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.

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

case is more likely when the median voter is a low type, compared to the other affected people.

Say that, if $\hat{\rho}_m < \hat{\rho}^*$, the rule is *too restrictive*; if $\hat{\rho}_m > \hat{\rho}^*$, the rule is *too permissive.* In the case both $\hat{\rho}_m$ and $\hat{\rho}^*$ are interior, 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.

Another incentive to set a too restrictive rule arises from the distribution of types. If the median voter is a rather low type and there is a concentration of types close but below his position, lowering the rule is highly productive for him since he can affect the behavior of many 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 median type is substantially lower than the average type.[11]

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.[12]

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

[11] 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 a tax level efficiency is the distance between the median income and its average.

[12] EPA's estimated benefits from further standards on light duty vehicles amount to \$0.5-1.2 billion without including a social cost of carbon and benefits from greenhouse gas emissions.

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. This may reflect situations in which the majority's goals about the behavior of minorities lead to laws and norms which are difficult to justify from a utilitarian perspective. Through the democratic process the majority has the chance, almost for free, to force the minority to endorse its own goals. In fact Baron (2003) claims that moralistic goals about how the others should behave are prominent in political choice. Roth (2007) argues that in many cases repugnance by some societies in certain time periods is the real constraint to some transactions (e.g. lending money for interest, eating horse meat, selling of organs for transplantation, ...). Our model provides a political-economy rationale for these restrictions.

4 Voting on a quota

We now analyze a policy which requires a reduction of the activity by a proportion $\tau \in [0,1]$ that we call "quota". Examples of quotas include the "20-20-20" plan of emission reduction adopted by the European Union in 2007 or the California's tailpipe standards which require a 30% reduction in emissions from new cars by 2016. Quotas are probably motivated by the idea of sharing the burden of externality reduction more equally. However, we will show later that there are no efficiency reasons for using this instrument, which is dominated by a proportional tax.

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:

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

It is easy to see that $U_i(\tau)$ is concave, and that each voter i's most preferred quota is negatively related to i's type. Under majority rule, the pivot is the median type, and the voting outcome is his most preferred quota. Therefore in the voting equilibrium,

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

where $a\varepsilon'(\tau) = -\int_0^1 t\varepsilon'((1-\tau)t)f(t)dt$ is the (positive per capita) marginal externality produced, after the quota has been enforced.[13]

The social welfare function is:

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

$W(\tau)$ is concave and the social optimum satisfies:

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

Again, social optimum is reached where the marginal benefit from a tax, $a\varepsilon'(\tau)$, equals the average marginal cost, $ac'(\tau) = \int_0^1 tc'(\tau t)f(t)dt$. By comparing (9) with (10) it is clear that any differences between the median voter and the social planner 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.

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.*

Inefficiency of the voting outcome is due to two factors. First, the median's relative position: a low median is inclined to prefer a high quota because his costs remain relatively low at the margin. Second, cost convexity: when costs are quite convex adjustments imposed to high types are socially too costly.[14]

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".[15] Interestingly, however, the vice

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

[14]The reader may notice that this model with a quota perfectly parallels Meltzer and Richard's (1981) model of taxation: incomes are "replaced" by types and proportional taxation is substituted by proportional adjustments in behaviors. Here marginal costs are not necessarily linear. In case they were, inefficiency would be due only to the first factor, in the sense that a median type below the average type would be necessary and sufficient for a too restrictive quota emerging. In case the median coincided with the average, the majority rule would deliver the socially optimal quota.

[15]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., $t_m c'(\tau_m^* t_m)/ac'(\tau_m^*) > 1$). 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., $c'(|\rho_m^* - t_m|) \cdot (1 - F(\rho_m^*))/ac'(\rho_m^*) < 1$), the rule is too restrictive.

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 the next Section we will use this result to compare the efficiency of taxes and rules.

In synthesis with a rule the risk of a too restrictive outcome is always higher than with a quota. The reason is that a rule offers a low median a larger scope for opportunistic voting. Thus whenever he selects a too restrictive quota he will also select a too harsh rule. 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. Pigouvian taxes aimed at influencing people's behavior are rather frequent in reality. Examples are the so-called ecotaxes, intended to promote ecologically sustainable activities.[16]

We assume that the government uses tax revenues to provide a non excludable public good, g. This will allow us to investigate the "double dividend" effect namely the fact

[16]Taxes on motor fuels and vehicles represent almost 90% of the revenue from environmentally related taxes in Europe. During the past decade, European countries have increased the use of this instrument. Environmental taxes target a broad array of bases (e.g. fertilizer, pesticides, plastic bags, landfill waste, batteries, etc.).

In the US motor fuel taxes remain substantially below the European levels. In 1993 Clinton's proposal for a broad-based energy tax on the Btu content of fuels failed to pass the Congress. Revenues from federal environmentally related taxes represent 3.5% of total tax revenues, compared to an average of 7% for the OECD countries.

that a tax reduces negative externalities and finances a useful public good. In Appendix 8.2 we consider the case in which proceeds are redistributed lump sum to population.[17] 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.

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:

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

Without a tax, non-atomic individuals do not contribute to the public good, unilaterally. The 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.

Individuals lower their behavior below their types and pay taxes on after-tax behavior. The relationship between behavior and tax derives from the individual cost optimization for given tax. The solution is given by the FOC:[18]

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

If the tax increases, individuals reduce their behavior and for any tax, higher types will prefer higher behaviors. Thus, the FOC pins down, in implicit form, the incentive

---

[17] The results are qualitatively similar. Some differences in the conditions leading to inefficiency are dicussed in Appendix.

Lump sum transfers of the proceeds is a standard assumption in taxation models. This assumption is however rather unrealistic for at least two reasons. First, externality taxes are frequently used to increase the budget size. Second, transfers are usually implemented "through" fiscal reliefs from other existing taxes. Thus, as long as individual contributions to those taxes are not constant rebate effects cannot be considered lump sum.

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

compatible behavior:[19]

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

Because of taxes individuals lower their behaviors below their types until their marginal cost from paying taxes equals the marginal cost of reducing behavior. Let this effect be the "first dividend". The "second dividend" of taxation is the financing of the public good. 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:

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

where $\omega\left(., t_i\right) = c(|b(t_i, \mu) - t_i|) + d(\mu b(t_i, \mu))$ and $b(t, \mu)$ is, for all i, the incentive constraint in (12). The first term in the right-hand side of (13) represents the externality from the after-tax behaviors of 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 (see also the proof of Lemma 2 below). 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 $\omega'_\mu\left(., t_i\right) = c'(.) \cdot b'_\mu(.) + d'(.) \cdot$

[19] Function b is not indexed. In fact, 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.

$\left(b(.) + \mu b'_\mu(.)\right)$. Since c and d are convex, then $\omega'_\mu(., t_i)$ is increasing in t_i.

The following Lemma is a tool to solve the voting stage.

Lemma 2 *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: $\hat{\mu}_m = \mu_m^*$.

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

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

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:

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

where $a\omega'_\mu$ is the average marginal cost and 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 social and private marginal benefits are equal and independent of type. 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 their marginal costs. This yields a result that parallels Propositions 1 and 2 above.

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.*[20]

A low median has an incentive to prefer high taxes because he gets large externality reductions and big amounts of public good compared to what he pays. This can generate a large inefficiency when costs are quite convex. Cost convexity and a relatively low position of the median tend to determine a too restrictive tax level, likewise the other two instruments. At this stage two questions may be raised: Is it possible that a tax is

[20] *The proof is similar to the proof of Proposition 2. Thus we omit it.*

too restrictive when a rule is too permissive? Can a tax be too restrictive if a quota is too permissive?

The answer to the first question is “No”. A low median has always a stronger incentive to set a too restrictive rule than a too restrictive tax. The reason is that a tax shares costs more evenly, thus a low median has to pay relatively more with a tax than with a rule.

The answer to the second question depends on the impact of a tax, namely how it affects the behavior of high and low types. If the tax induces higher types to less than proportional adjustments, costs are less concentrated on high types, then also the answer to the second question is “No”. A low median has stronger incentives to set a too restrictive quota than a too restrictive tax because he pays relatively more with a tax than with a quota. Vice versa, if the higher types reduce their after-tax behavior more than proportionally a low median pays relatively less with a tax than with a quota, thus he has stronger incentives to set a too restrictive tax. Proposition 5 establishes these results and states that more/less than proportional adjustments depend on the curvature of marginal costs.

Proposition 5 *i) a) A tax cannot be too restrictive whenever a rule is too permissive. b) A tax may be too permissive when a rule is too restrictive.*
ii) a) If c' is more convex than d'_b, after-tax behavior is less than proportional to type. In this case, a tax cannot be too restrictive whenever a quota is too permissive. b) If c' is less convex than d'_b, after-tax behavior is more than proportional to type. A quota cannot be too restrictive whenever a tax is too permissive.

Part *i)* of Proposition 5 should not be interpreted as: “a tax is less restrictive than a rule”, but rather as: “in case of a rule, too restrictive outcomes are more frequent than in case of a tax”. In fact, in order to say something on the “amount” of the political distortion we would need to be more specific about the shapes of the externality and cost functions. Nonetheless we can make an informal point. The political distortion is small when there is little discrepancy between the median and the average marginal cost. We expect this discrepancy to be smaller with a tax than with a rule, because a tax allocates costs more evenly.[21] Conceivably, political distortion is smaller with a tax than with a rule.

[21] A well known result in probability theory states that the median and the mean of a distribution can never differ from each other by more than one standard deviation. Since the standard deviation of

This model predicts that too restrictive rules occur more frequently in reality than too high taxes. This explains why, for example, taxation on smoking or motor fuels is below the socially optimal level or why road congestion pricing in urban areas is usually rather low.[22]

6 The choice of the policy instrument

6.1 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 generally a rule is very restrictive. Another result is that quotas are always dominated by taxes. For this reason, that will be discussed later, we restrict our attention to a binary choice between rules and taxes.

Benefits and costs of rules (and 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. We use this concept to represent within the same policy space the choice among 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 10 units of taxes plus reducing their behavior by 0.1. Note that this is not restrictive since the benefits of taxes and polluting

the marginal cost ditribution is lower with a tax than with a rule, we claim that median and average marginal costs are closer with a tax than with a rule.

[22] As pointed out in the Introduction, socially optimal gasoline taxes in the US should be much higher. In 2006 the Center for Disease Control and Prevention calculated the social cost of smoking cigarettes at \$10.47 per pack (including the cost of treating smoking-related illnesses and the production loss due to illnesses).

Cordon area congestion pricing is a fee paid to enter a restricted area. Stockholm and Milan have set the congestion charge at about \$3: a quite low level. Singapore adopts more differenciated tariffs whose the average level is rather low. The \$15 London congestion charge is probably the highest in the world. This kind of taxes are becoming popular, their implementation is however quite difficult.

behavior are then evaluated by means of generically different utility and cost functions. Call $RE(.)$ the revenue equivalent of an instrument. For a rule and a tax we have, respectively:

$$RE(\rho) = \int_{\rho}^{1} tdF(t) - \rho(1 - F(\rho))$$
$$RE(\mu) = \int_{0}^{1} [t - (1 - \mu) \cdot b(\mu, t)]\, dF(t)$$

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, the revenue equivalent is due to total behavioral reductions ($\int_{0}^{1} [t - b(\mu, t)]\, dF(t)$) plus total tax revenues on residual behaviors ($\int_{0}^{1} \mu \cdot b(\mu, t) dF(t)$). Summing and rearranging yields the left-hand side.[23]

Below we approach the instrument choice with a cost-benefit analysis, in which benefits and costs are a function of RE. After, we look for the Condorcet winner, then we discuss efficiency.

6.2 Private benefits from RE

When externalities are quite concave a rule always performs better than the other instruments. In this case, the main objective is restricting the behavior of top polluters, then a rule is the best way to do it. Thus the benefits from a rule are larger for any amount of RE.

If this is not the case (i.e. if concavity of externalities is low), a tax may produce larger benefits provided that the double dividend is strong. Call $B(RE(\rho))$ and $B(RE(\mu))$ the benefit function from revenue equivalents due to a rule and to a tax. Examples of possible shapes are in Figure 1: initially the rule performs better because the top polluters are affected; for large RE amounts a tax is better because of the double dividend.[24]

6.3 Private costs from RE

Private costs differ not only across instruments but also across types.

[23] In the case of a quota we have that $RE(\tau) = \tau \int_{0}^{1} tdF(t)$.

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

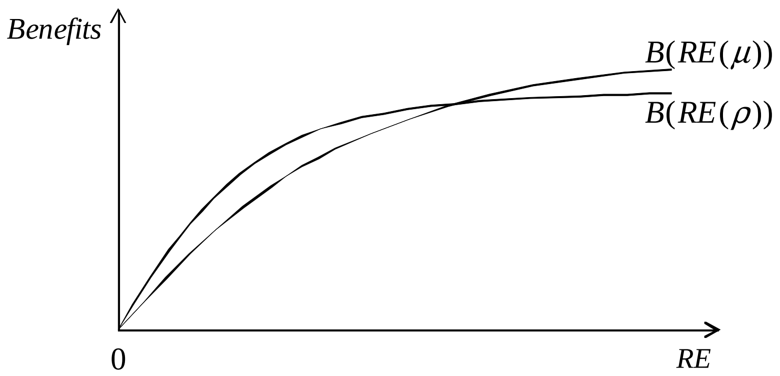

Figure 1: Benefits from rules and taxes

1. 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 private costs; 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. Denote with $C(RE(\rho), t_i)$ the cost function of type i. Figure 2 shows a possible shape for this function. For a low type, the zero-cost area is extended and the function is rather flat since the burden of producing additional RE is shared among many people. In fact, marginal costs are inversely proportional to the affected population.[25] Vice versa, for a high type the zero-cost area is narrow and the curve is quite steep.

2. Let $C(RE(\mu), t_i)$ be the cost of RE generated by a tax. A tentative shape is shown in Figure 2. Assume that the sacrifice of paying a tax and the cost of reducing behavior are the same function (i.e. $c(.) = d(.)$; an assumption that we will keep hereafter). As pointed out earlier, in this case revenue equivalents come from proportional adjustments,

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

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

Since,

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

then,

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

The marginal cost of RE is inversely proportional to the affected population.

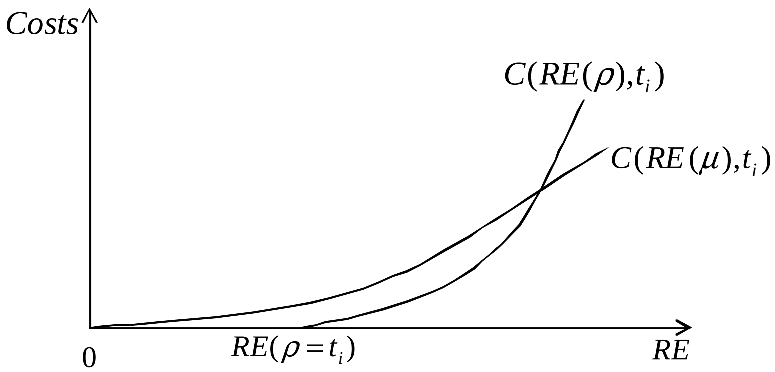

Figure 2: Costs of rules and taxes

hence marginal costs are proportional to types.[26]

Observe that a quota is always more costly than a tax. In fact, with both instruments individuals make proportional RE contributions. However with a tax individuals can optimize the cost of any RE by choosing between behavioral adjustments and tax payments. 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.

Despite their inefficiency, quotas are frequent in reality. One reason is that proportional reductions are perceived as a fair method of sharing the sacrifices of curbing externalities. Another reason is that quotas do not involve monetary payments or redistributions. This may become an important issue when people are concerned about rent-seeking by politicians, or when tax collection devices are inefficient or not available. For example, many supranational institutions or international agreements have a quite limited ability to tax (e.g. the Kyoto protocol). In these cases parties resort to quotas as a substitute of taxes.[27]

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

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

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

[27] In the European Union there is a sharp constrast between a limited budgetary ability (only 1.4% of member's GNI) and an extended regulatory activity. The already mentioned 20-20-20 plan is mainly based on the use of command-and-control measures (like proportional quotas), rather than taxes. The most likely reason is that countries do not want to relinquish their monopoly in taxation.

6.4 Existence of a Condorcet winner: sequential voting

Let us start with the case in which voting takes place sequentially: first, the majority selects the instrument; then it chooses its level. In this case, a Condorcet winner always exists.[28] Individuals know that, whatever the instrument, the level that will pass at the second stage is the one preferred by the median, either ρ_m^* or μ_m^*. The revenue equivalent outcome is either $RE(\rho_m^*)$ or $RE(\mu_m^*)$. 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 which we explore in the next Section.

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. Utility from the public good (or from tax rebates) is not sufficient to make benefits from a tax more attractive because a rule is a much cheaper way to generate RE. With a rule they can induce virtuous behavior by others paying zero or a very low cost. There might also be intermediate types who are available to pay a significant private cost because the benefits from a rule are substantially larger.

Thus a rule is more likely than a tax when there are many low types (with potential political distortions) or when the rule is a more efficient way to cope with the externality problem (i.e. externalities are quite concave and the double dividend is low).

6.5 Existence of a Condorcet winner: simultaneous voting

Suppose now that a majority has to form on the instrument choice and on its level simultaneously. Voting takes place 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

[28] 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.

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).

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 instrument is a rule and the median level is ρ_m^*. Thus the Condorcet winner is $RE(\rho_m^*)$.[29]

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 a tax. Nonetheless, we can give a *sufficient* condition for it. The idea is simple: take the median type t_m. Assume that he prefers a rule. We require that for all lower types the rule is the cheapest way to generate RE. This happens if the marginal costs of lower types decrease more with a rule than with a tax. As pointed out earlier, with a rule marginal costs are inversely proportional to $1 - F(\rho)$ and with a tax they are inversely proportional to $\bar{t}/t_i$. Thus, a rule is the cheapest instrument if $1 - F(\rho) < \bar{t}/t_i$ for all i below the median. This point is formally made in the following lemma.

Lemma 3 *i) If the median prefers a rule to a tax, and if*

$$1 - F(\rho_m^*) < \frac{\bar{t}}{t_m} \tag{16}$$

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

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

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

Inequality (16) may be explained with an example. Suppose that $(1 - F(\rho_m^*)) = 0.5$ and $\bar{t}/t_m = 1$. Recall that for the median the marginal cost of a rule is proportional to $1/(1 - F(\rho_m^*)) = 2$, and the marginal cost of a tax is proportional to $t_m/\bar{t} = 1$. In this case a marginal decrease in t_m will cause a reduction in marginal cost that is double in the case of a rule than in the case of a tax. The reason is that with a rule the group of

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

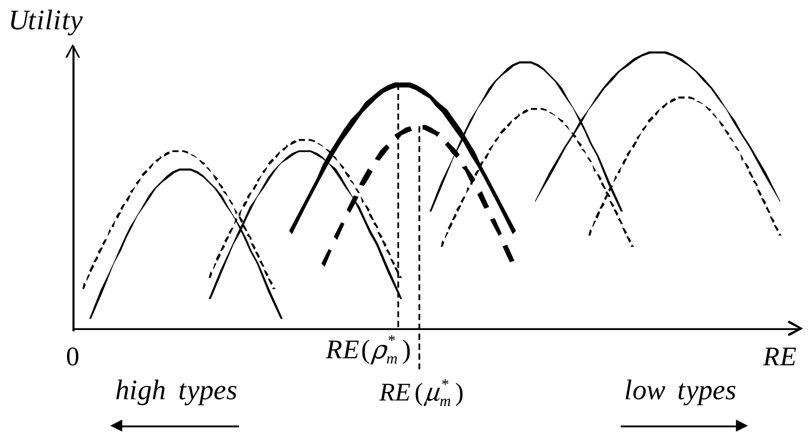

Figure 3: Monotonic preferences in both dimensions

people with whom a lower type can share the cost enlarges grows by a larger proportion. Therefore, if the median prefers a rule *notwithstanding* he shares costs with 50% of the population only, then a lower type has a stronger incentive to prefer a rule. [30]

This voting equilibrium is represented in Figure 3. The x-axis is the second dimension of the policy space, the set of RE.[31] 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 inverse-monotonic in RE and the median bliss points belong to the median type in both cases. Under Lemma 3, 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 $RE(\rho_m^*)$.

Proposition 6 *If Lemma 3 applies, then the majority prefers a rule and it selects the median's most preferred level.*

[30] Observe that (16) introduces a "local" requisite that concerns only a marginal decrease in t_m. Part ii) of Lemma 3 extends the requisite to all types below t_m.

The reader may also notice that, by assuming that the benefits from rules and taxes are the same and by requiring that costs are monotonic in both dymensions, we are de facto applying a separation argument that is similar to Grandmont's (1978) sufficient condition for the existence of a Condorcet winner in multidimesional policy problems.

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

This kind of equilibrium is more likely when the double dividend is low and externalities are quite concave, so that the benefits of a rule are high. Moreover, since a rule can be very cheap for a low median, a majority of low types prefers a rule in order to charge the minority a larger burden for the externality reduction.

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 3 applies:

Corollary 1 *If the median is below the average and Lemma 3-ii) applies, then the majority always prefers a rule.*

6.6 Normative aspects

Without any restrictions on the shape of the tax schedule a tax would always be superior to other policy instruments. However because of realism and to guarantee existence of a Condorcet winner we are restricting ourselves here to the case of proportional tax. Therefore a proportional tax may not be always superior to a rule. That is suppose that the optimal unrestricted tax would be very progressive, i.e. taxing high types at very high rates. Then a rule may be a closer approximation to such a tax than a proportional tax.[32] This is the case when externalities are quite concave, the double dividend is limited and cost convexity is low. In all other cases a proportional tax performs better. It generates larger benefits from the double dividend and it is "socially" cheaper than a rule.

How are these normative conclusions related to the positive results above? Suppose that the social planner prefers a proportional tax because externalities are rather flat and costs are quite convex. Suppose also that the distribution is rightward slanted so that there is sufficient support for a rule. By Proposition 1 we know that slanted distributions and convex costs yield quite restrictive rules. Thus the general idea is that if the social planner prefers a tax and the majority prefers a rule, a too restrictive rule is very likely.

Vice versa, when externalities are quite concave and marginal costs are rather flat the socially optimal instrument is a rule. Suppose that the majority chooses a tax because the median is in a relatively high position. In this case, the tax level is too low. If the social planner prefers a rule and the majority prefers a tax, a too low tax is quite likely.

[32] Observe that an appropriate tax schedule can reach the social optimum even if types are private. This happens if disutility of paying taxes is linear and it happens in any case when a public good is provided. Appendix 8.3 offers some further details on these points.

In a sense one may talk of a "double political distortion": if the majority chooses the wrong instrument then also its level is wrong. More precisely, when polluting activities are mainly due minorities, there will be too many rules instead of taxes (first distortion) and their level will be too restrictive (second distortion). If pollution derives from activities that are enjoyed by the majority, then there will be a preference for taxes and the level will be too low.

This possibly explains why motor fuel taxation is too low, as discussed earlier. This is also consistent with the recent debate on obesity policies, which is a major concern in the US compared to Europe.[33] What we observe is that the EU prefers regulation (e.g. more standards on fat contents, clearer food labelling, improving the nutritional content of school and office meals), while the US is possibly oriented to taxation. Recently, in the US Congress there have been proposals for an obesity tax, which is expected to be low compared to social costs.

7 Conclusion

We have examined the political economy of how to curb activities which generate negative externalities. We showed three things. First, for given policy tool (rules, quota or taxes)generally the median voter does not choose the efficient level of the policy instrument. Too stringent restrictions (or too high taxes) are chosen when the activity with negative externality is concentrated on a relatively small fraction of the population and when the cost of limiting the polluting activity grows more than proportionally. In the opposite case instead regulation or taxation would be too lenient. Second, when regulation is chosen it is always more restrictive compared to taxation. With a tax the private cost of the median voter is closer to the social cost, thus the political distortion is smaller. Third, we showed that when majorities can also choose the policy instrument they 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. We also showed that the two political distortions (wrong instruments and wrong level) are

[33] According to OECD statistics, more than 30% of Americans are clinically obese. In Europe the percentage is slightly above 10%.

The health care expenditures related to obesity range from $100- to $150 billion per year in the US, with a growing pattern.

correlated. When a majority prefers a rule, whereas the optimal instrument is a tax, a too restrictive rule is likely to come about. Vice versa, if the majority prefers a tax and the best instrument is a rule, then a too low tax level will be chosen.

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 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.

Third, thus far we have imposed that rules and quotas are self enforcing. 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.

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.[34] The problem is that a super-majority assigns blocking power not only to

[34] 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.

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.[35] 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. 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.

[35] 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.

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,

$$\frac{\partial \rho_i^*}{\partial t_i} = -\frac{-c''(|\rho - t_i|)}{\varepsilon''(\rho)\,(1 - F(\rho)) - \varepsilon'(\rho) f(\rho) - c''(|\rho - t_i|)} \tag{17}$$

The denominator in the right-hand side of (17) 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 $\hat{\rho}_m < \hat{\rho}^*$, we have

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

and

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

Therefore,

$$c'(|\hat{\rho}_m - t_m|) > ac'(\hat{\rho}_m)$$

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

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

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

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

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

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

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, $\hat{\tau}_m > \hat{\tau}^*$. From (9) and from (10), we have:

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

and

$$a\varepsilon'(\hat{\tau}_m) = t_m c'(|-\hat{\tau}_m t_m|)$$

Therefore,

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

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

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

We want to show that, for any τ,

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

Consider that $ac'(\hat{\rho}_m)$ is an average in which the only non-zero elements are the $(1 - F(\hat{\rho}_m))$ marginal costs of the affected people above $F(\hat{\rho}_m)$; where $1 - F(\hat{\rho}_m) > 0.5$.

Moreover 50% of these elements are larger than $c'(|\hat{\rho}_m - t_m|)$. 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 $F(\hat{\rho}_m)$ 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 $ac'(\tau) > ac'(\hat{\rho}_m)$, but $ac'(\tau)$ cannot be larger that the median's marginal cost, $t_m c'(\tau t_m)$. ■

Lemma 2 *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 (12). 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 (12) is satisfied for higher marginal costs. Thus, total cost $\omega\left(., b(\mu, t_i), \mu, t_i\right)$ is convex in μ, and, for any μ, marginal costs are higher for higher types. Observe that benefits, i.e. the first term 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. ■

Proposition 5 *i) a) A tax cannot be too restrictive whenever a rule is too permissive. b) A tax may be too permissive when a rule is too restrictive.*
ii) a) If c' is more convex than d'_b, after-tax behavior is less than proportional to type. In this case a tax cannot be too restrictive whenever a quota is too permissive. b) If c' is less convex than d'_b, after-tax behavior is more than proportional to type. A quota cannot be too restrictive whenever a tax is too permissive.

Proof. Part i) can be proved using the same argument of Proposition 3.

As for part ii), observe that:
a) If $c''' > d'''$ then $\frac{\partial^2 b}{\partial t^2} > 0$; i.e. the relationship between type and after-tax behavior is convex . This can be proved by second implicit differentiation of (11). To save notation, let us rewrite (11) as $c'(t-b) - d'(b) = 0 = H(t, b)$. By Dini's theorem,

$$\frac{\partial^2 b}{\partial t^2} = -\frac{H_b^2 H_{tt} - 2H_t H_b H_{tb} + H_t^2 H_{bb}}{H_b^3}$$

Substituting and simplifying,

$$\frac{\partial^2 b}{\partial t^2} = \frac{c'' \cdot (c''' - d''')}{(c'' + d'')^3} \tag{18}$$

which is positive if $c''' > d'''$. This means that adjustments are less than proportional to types. With respect to the tax, a quota concentrates marginal adjustment costs on higher types. Applying the same rationale as in Proposition 3, we can state that insofar as with a quota the median marginal cost is higher than the average (i.e. the quota is too permissive) then the median marginal cost cannot be lower than the average if a tax instead of a quota is adopted. Thus a tax cannot be too restrictive if a quota is too permissive.

b) By (18) if $c''' < d'''$ then $\frac{\partial^2 b}{\partial t^2} < 0$. With a tax, adjustments are more than proportional to types and costs are concentrated on higher types. If the median marginal cost is higher than the average with a tax (i.e. a tax is too permissive) it has to be so also with a quota (i.e. a quota cannot be too restrictive).

Observe that $c''' = d_b'''$ then $\frac{\partial^2 b}{\partial t^2} = 0$ after-tax behavior is linearly related to types. In this case, a tax has a proportional impact on behaviors. The allocation of costs across population comes out of proportional adjustments, a mechanism that is similar to the quota. Thus a quota and a tax are either too restrictive or too permissive under the same circumstances. ■

Lemma 3 *i) If the median prefers a rule to a tax, and if*

$$1 - F(\rho_m^*) < \frac{\bar{t}}{t_m} \tag{19}$$

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

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

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

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

$$\frac{\partial B(RE(\hat{\rho}_m))}{\partial RE} = \frac{1}{1 - F(\hat{\rho}_m)} c'(t_m - \hat{\rho}_m)$$

Call $RE(\hat{\rho}_m)$ the revenue equivalent when the optimal rule is applied. Let μ^0 the tax level (without refunds) that yields the same revenue equivalent, and call $RE(\mu^0)$ that amount. By assumption benefits from rule and tax are the same. Therefore, $\frac{\partial B(RE(\hat{\rho}_m))}{\partial RE} =$

$\frac{\partial B(RE(\mu^0))}{\partial RE}$. Thus we have,

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

where $\omega'(.,\mu^0)=[b_i-(1-\mu^0)\cdot b_i']\cdot c'(.)$. A marginal decrease in t_m 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

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

i.e. if condition (19) 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_m prefers a rule and a marginally lower type prefers a tax. Part ii) can be proved applying the proof of part i) recursively. ■

Proposition 6 *If Lemma 3 applies, then the majority prefers a rule and it selects the median's most preferred level.*

Proof. The median prefers a rule when his position is low, externalities are quite concave and the double dividend is small. Conditions in Lemma 3 are sufficient to ensure that the median bliss points in both policy dimensions belong to the median type. Thus the Condorcet winner is $RE(\rho_m^*)$. ■

Corollary 1 *If the median is below the average and Lemma 3-ii) applies, then the majority always prefers a rule.*

Proof. Observe that if the median is below the average, then $\frac{\bar{t}}{t_m}>1$. Thus condition (19) is always satisfied. We only need that part ii) of Lemma *3* applies. ■

8.2 Taxes with lump sum transfers

We assume here that 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. A significant number of recent environmental taxes in Europe have followed a revenue-neutral approach which is a form of transfer (the so-called green tax shifting). As argued earlier, the fact that those transfers are lump sum is questionable. While standard and elegant from a theoretical viewpoint, lump sum transfers are quite rare in reality.

Each individual receives out of a balanced public budget a transfer which amounts to the average tax burden, $\mu \cdot \bar{b}$. Call $d(\mu\left(b_i - \bar{b}\right))$ the net cost that i bears from paying the tax when his behavior is b_i, with $d'(.) > 0$ and $d''(.) \geq 0$. Observe that, in contrast with the basic literature on optimal taxation, we do not assume that utility is quasi-linear in money. The reason is that the convexity of d plays a role in the efficiency of the voting outcome.

Individuals choose their behavior and pay taxes on after-tax behavior. The cost incurred by agent i is the following:

$$\psi\left(., b_i, \mu, t_i\right) = c(|b_i - t_i|) + d(\mu(b_i - \bar{b}))$$

Optimal after-tax behavior minimizes the cost function $\psi\left(., b_i, \mu, t_i\right)$:

$$b_i \in \arg\min_{\tilde{b}_i} \psi\left(., \tilde{b}_i, \mu, t_i\right)$$

Optimal behavior solves the following FOC:

$$-c'(|b_i - t_i|) = d'_b(.) \tag{20}$$

Given our assumptions the SOC is satisfied and the solution of (20) is unique. This solution pins down the equilibrium after-tax behavior, $b_i = b(\mu, t_i)$. Plugging the optimal behavior for all individuals into i's utility, yields i's indirect preferences for the tax rate:

$$U_i(\mu) = \int_0^1 \varepsilon(b(\mu, t))dF(t) - \psi\left(., b(\mu, t_i), \mu, t_i\right) \tag{21}$$

Maximizing (21) yields the bliss point of agent i. Applying the same rationale of Lemma 2 it is possible to see that bliss points are monotonically related to types and the voting outcome is the tax rate preferred by the median. The voting outcome, if interior, solves the following FOC:

$$\varepsilon'_\mu = \psi'_\mu\left(., b(\mu, t_m), \mu, t_m\right)$$

where $\varepsilon'_\mu = \int_0^1 \varepsilon'_b b'_\mu f(t)dt$ is the private marginal benefits from externality reduction and the right-hand side is the median's private marginal cost of taxation.

The policy benchmark maximizes the following social preference function, subject to (20) for all individuals:

$$W(\mu) = \int_0^1 \varepsilon(b(\mu,t))dF(t) + \int_0^1 \psi\left(.,b(\mu,t),\mu,t\right)dF(t)$$

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

$$\varepsilon'_\mu = a\omega'_\mu$$

The social planner would choose a tax such that per-capita marginal benefits are equal to per-capita (or 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.

In the case of lump sum transfers, the nature of the political distortion does not change with respect to public good provision. The benefit from lump sum transfers substitutes public good provision. Formally, public goods increase benefits, whereas fiscal reliefs reduce the cost of paying taxes. Arguably, some differences may arise when this substitution has not the same impact for all individuals. Suppose that the marginal sacrifice of paying taxes grows fast for low amounts (i.e. $d''' < 0$). In this case, fiscal reliefs have strong impact on the utility of low types, and may induce them to choose too high tax rates. This kind of distortion does not occur in the case of public good, since we have assumed that utility from the public good is the same for all individuals.

Summing up, with respect to the public good provision, an additional political distortion occurs in the case utility of transfers is different amongst individuals. Of course, if utility functions are quasi-linear, such a distortion disappears.

8.3 Optimality conditions and private types

While proportional taxes are widely used in reality, they might not be the socially optimal instrument. Specifically, a proportional tax implements the first best allocation if the social loss function is $\varepsilon(\int b(t)f(t)dt)$ and, $\omega\left(.,b_i,\mu,t_i\right) = c(|b_i - t_i|) + d(\mu b_i)$. In this case the social optimum requires that, for any t,

$$\varepsilon'(\int b(t)f(t)dt) = \omega'_{\mu}(t - b(t))$$

With this special functional form of ε, which is usually adopted in taxation literature, the marginal cost of pollution reduction is uniform across agents. Hence a constant price of pollution (i.e. a linear tax) can implement this first best allocation.

If instead, as in our model, the social loss is $\int_0^1 \varepsilon(b)dG(b)$, the optimality condition is

$$\varepsilon'(b(t)) = \omega'_{\mu}(t - b(t))$$

In this case the social planner could implement the optimum with a non-linear tax schedule where marginal taxes are variable in type so that high cost types are allowed to pollute more. The reader may verify that the same conclusion holds in the case tax proceeds are refunded lump sum.

An appropriate tax schedule yields the social optimum even if types are private information. This happens if utility is quasi-linear in money and it happens in any case when a public good is provided. Suppose the social planner does not know individual types, but it only knows the distribution $F(t)$. Call $b^{FB} = \int b(t)f(t)dt$. Suppose that marginal costs are independent of types (i.e. the social loss function is $\varepsilon(\int b(t)f(t)dt)$). After simple algebraic manipulations of the optimality condition given in the previous footnote we end up with

$$\varepsilon'(b^{FB}) = \omega'_{\mu}\left(E(t) - b^{FB}\right)$$

which pins down b^{FB}.

If instead the social loss is $\int_0^1 \varepsilon(b)dG(b)$, one can again prove that once we allow for uniform access to the public good (or linear transfers) the truth-telling constraint has no bite. Thus the social planner can establish a nice non-linear tax schedule that achieves the first best allocation. In other words, as long as tax rebates or public goods have the same utility for everybody, the inability of the government to perceive individuals' private costs does not impose any distortion on the economy. This is related to the more general insight that when trying to solve the Mirrlees problem with quasi-linear preferences, truth-telling constraints have no bite.

We are grateful to Pierre Yared for suggesting the idea above and the way our model is related to the Mirrlees problem.

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] Baron, Johathan (2003): "Value analysis of political behavior: self-interested, moralistic, altruistic, moral", *University of Pennsylvania Law Review*, 151, pp. 1135-1167.

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

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

[8] 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.

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

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

[11] 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.

[12] 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.

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

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

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

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

[17] 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.

[18] 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.

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

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

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

[22] Roth, Alvin E. (2007): "Repugnance as a Constraint on Markets", *Journal of Economic Perspectives*, Vol. 21, pp. 37–58.

[23] Shleifer, Andrei and Robert W. Vishny (1998): *The Grabbing Hand: Government Pathologies and their Cures*, Cambridge, MA, Harvard University Press.