

Comments welcome.

Social Welfare and Collective Goods Coercion in Public Economics

by

Stanley L. Winer^a, George Tridimas^b and Walter Hettich^c

This draft, October 21, 2006

An earlier version was presented at the University of Eastern Piedmont May 2006, and at Bar Ilan University and the IIPF Congress, Paphos, August 2006. We are indebted for comments by Alberto Cassone, Thomas Dalton, Mario Ferrer, Arye Hillman, Sebastien Kessing, Carla Marchese, Pierre Pestieau and Dan Usher. Errors and omissions remain the responsibility of the authors.

^a: School of Public Policy and Administration and Department of Economics, Carleton University, Ottawa, Canada K1S5B6. Tel: (613)520-2600 ext. 2630. Fax: (613)520-2551. E-mail: stan_winer@carleton.ca; ^b: School of Economics and Politics, University of Ulster, Shore Road, Newtownabbey, Co. Antrim BT37 0QB, UK Tel.: +44(0)28 90368273; Fax: +44 (0) 28 90 366847. E-mail: G.Tridimas@ulster.ac.uk; ^c: Department of Economics, California State University at Fullerton, E-mail: whettich@fullerton.edu. Winer's research is supported by the Canada Research Chair program. He is also indebted to the International Center for Economic Research, Turin and the Center for Economic Studies, Munich.

Extended Abstract

Although philosophers and constitutional experts have examined the nature of coercion at length, economists have not offered a comprehensive analysis of its role in normative theory. The social planner model, the main theoretical engine of normative public economics, does deal with redistribution, which may indeed be coercive, but it does not assign any normative weight to the coercive aspects of redistributive policies per se, acknowledging only the social welfare generated by redistribution whatever its coercive impact.

There are two well-known normative approaches falling in the literature on collective choice that deal more directly with coercion. Wicksell (1896) proposed that all policies be implemented using a unanimity rule, an idea that Lindahl (1919) attempted to improve upon by devising a system of voluntary consent for the financing of a public good. Buchanan and Tullock (1962) showed how to find an optimal policy rule in the face of coercion and decision costs. All of these authors placed their focus on process rather than on the welfare implications of particular policies, which limits application of their analysis to evaluation of alternative policy choices. Nor did they investigate how to precisely define and measure coercion.

The power of the social planning literature derives in large part from its ability to consider the implications of specific policy choices. This allows one to compare different outcomes and to evaluate them in the context of social welfare. A purely process-oriented approach cannot accomplish this, limiting the analyst to giving advice on institutions and their design rather than on specific policies. Yet, the literature focusing on process has made an important contribution by drawing attention to the central role of coercion in public life.

This paper develops an ecumenical approach to normative policy choice when coercion is explicitly acknowledged. We examine maximization of social welfare when coercion stemming from the collective provision of public goods, as distinct from redistribution, plays a central role.

The precise definition of coercion, the difference between such coercion and redistribution, and the method of its incorporation into a broader approach to social welfare optimization are the key issues in the paper. We investigate the implications of the expanded normative framework for the nature of optimal policy, showing how the Samuelson condition and tax policy rules must be substantially modified. We map the trade-off between social welfare and coercion under specific conditions, and consider the implications of this trade-off for normative policy analysis.

Key words: Coercion, optimal taxation, public goods, collective choice, unanimity, solidarity

JEL Categories: H1, H2, H3, H4

The essential feature that defines a democratic government is voluntary agreement by members of the public to subject themselves to its coercion.

William Baumol (2003, 613)

From the point of view of general solidarity...parties and social classes should...share an expense from which they receive no great or direct benefit. Give and take is a firm foundation of lasting friendship... It is quite a different matter, however, to be forced so to contribute. Coercion is always an evil in itself and its exercise, in my opinion, can be justified only in cases of clear necessity.

Knut Wicksell (1896/1958, 90)

1. Introduction

Coercion is a fundamental and unavoidable aspect of public life. Although philosophers and constitutional experts have examined its nature at length, economists have not offered a comprehensive analysis of its role in traditional normative theory. The planner model, the main theoretical engine of normative public economics, does deal with redistribution, which may indeed be coercive, but it does not assign any normative weight to the coercive aspects of redistributive policies per se, acknowledging only the social welfare generated by redistribution, whatever its coercive impact. Nor does it deal with other aspects of coercion that limit the freedom of individuals to make important marginal adjustments. Politics or institutions that may represent a response to various aspects of coercion remain outside the analysis, except for occasional ad hoc adjustments made by astute policy advisors when results are applied to actual situations.

Coercion in collective decision making occurs in many situations. Consider, for example, a group of people who have come together in a room for a common purpose and who must set the temperature on a thermostat and then pay for the resulting use of energy. Since Lindahl pricing is not feasible, some will be too hot and some too cold, and even those who are neither may be unhappy with the balance that results between what they pay and what they get. Individuals can escape the resulting situation if they move to another room or out of the building that represents the collectivity in this example. But if they stay, they must cope with the coercion implied by their assent to the collective decision.

As Baumol (2003) points out, people participate in a collectivity despite its coercive aspects because they are better off with some coercion than with the alternative. In other words, they will accept a certain amount of coercion in situations like those described as a necessary cost of having a community. But this is a keenly felt cost, as Wicksell (1896), Buchanan (1967), Breton (1974), Usher (1981) and others have stressed. If it becomes too great for too many, unrest, emigration and eventually failure of the state as a productive enterprise may occur. In more general terms, too much coercion endangers the operation of the collective choice process and the production of public services.

There are two well-known normative approaches falling outside the social planning literature that deal formally with coercion inherent in public policies. Wicksell (1896) proposed that all policies be implemented using a qualified unanimity rule, and Lindahl (1919) devised a method of voluntary exchange that would have the same result.¹

¹ The relationship between Lindahl equilibrium and collective choice, on the one hand, and coercion, on the other, has been discussed extensively in the literature, although not in a context reflecting recent normative analysis in

Unanimity and Lindahl's voluntary exchange both result in efficiency and the absence of coercion. As has been pointed out repeatedly however, such processes lead to high decision costs and are not descriptive of the real world. A different approach, developed by Buchanan and Tullock (1962), aims at finding an optimal policy rule in the face of decision costs and coercion. Like the two Scandinavian economists, they focus primarily on process and do not link their analysis to the welfare implications of particular policies. In none of these contributions is it clear how coercion is defined or how it is to be measured.

The power of the social planning literature derives in large part from its ability to consider the implications of specific policy choices. This allows one to evaluate and compare different outcomes in terms of the social welfare they imply. A purely process-oriented approach cannot accomplish this, limiting the analyst to giving advice on institutions and their design rather than on specific policies. Yet, the literature focusing on process has made an important contribution by drawing attention to the central role of coercion in public life, by linking it to different collective choice processes and by pointing out the need to formally acknowledge it as a significant factor in policy analysis.

This paper develops an ecumenical approach to normative policy choice by formally introducing coercion into social planning, while staying within the context of this tradition.² We examine maximization of social welfare when the planner is formally bound by constraints on the degree of coercion, as distinct from redistribution, that may be generated as part of a social plan to supply public goods financed by compulsory taxation. The necessity of coercion in a democratic society when public goods are collectively supplied, and the difference between what we shall call “collective goods coercion” and redistribution, and its incorporation into a broader approach to social welfare optimization as constraints on the planner are key issues addressed in section 2 of the paper.

The precise definition of collective goods coercion is developed in section 3. It is important to note here that coercion is involved in the origins of the state (see Usher 1993, Baumol 2003, and Perra and Scharf 2003) and can be imposed in many ways including, for example, through conscription (Levy 1997) or by regulating access to and the scope of private markets (Wiseman 1989). The term 'collective goods coercion' acknowledges that in this paper, we confine the analysis to coercion that arises when public goods are collectively provided.

In section 4 we investigate the implications of the expanded social planning framework for the nature of optimal policy, showing how the Samuelson rule and the optimal nature of tax structure must be substantially modified. (In an Appendix, we use the same approach to consider how the Ramsey rule (1927) for optimal commodity taxation must be modified to account for coercion.) In section 5 we map the trade-off between social welfare and coercion under specific conditions, and consider the implications of this trade-off for normative policy analysis. Section 6 concludes.

We realize, of course, that imposing constraints on a planner that are derived from a concern with the

public economics. Authors, such as Escarrez (1967) and Slutsky (1979) have focused on fiscal systems that are required when Lindahl equilibria are ruled out and their implications for coercion. Less formally, and in much earlier work, Henry Simons (1939) was concerned primarily with limiting the government's ability to use taxation as a way of intervening selectively and coercively in markets with the establishment of standards for horizontal equity. More recently, Buchanan and Congleton (1998) adopted a similar approach to this in their extensive proposal for new fiscal arrangements.

² Breton (1974), Dalton (1977) and Breton (1996) also analyze coercion implicit in public finance using a definition of coercion that depends exclusively on the level of the public good, as discussed below.

quality of collective choice, as we do here, extends the analysis beyond criteria generally accepted in the social planning or optimal tax literature. We believe that our approach is justified because it allows us to address significant questions about the role of coercion and its relationship to social welfare, and because we think that a concern with coercion must be an important and legitimate focus in the analysis of the public sector.

2. Coercion Versus Redistribution In Policy Analysis

Coercion in a democratic society is inevitable as a consequence of individuals interpreting and reacting to the nature and outcome of collective decision making. Jointness in supply, problems with preference revelation, and economies of scale in consumption create a situation in which public goods and services cannot be efficiently provided in private markets. Since unanimity doesn't work as a decision mechanism because of possible strategic behavior, some form of majority rule is required as long as we are interested in a democratic society, and majority rule in whatever form leads to situations in which there will be an imperfect matching of what people pay in taxes and what they actually receive in public services.

The individual's evaluation of the imperfect matching between what he or she gets and what is paid represents the essence of, though not a precise definition of what we referred to as "collective goods coercion" in the introduction. The problem arises as a by-product of any practical collective choice mechanism, even though such a mechanism is needed to reap the benefits from participating in a collectivity. (We note here that for simplicity of expression, we shall simply use the term coercion in the rest of the paper, rather than to the full term "collective goods coercion".) In other words, following Buchanan (2001, 588) we can look at participation in a political community as an imperfect exchange, where of necessity benefits and costs are not fully matched on an individual basis.³

Wicksell (1896) reminds us that this imperfect matching differs from, and goes beyond, voluntary redistribution. Nor is it the same as redistribution instituted by a planner as part of a program aimed at increasing social welfare. To see in general terms that coercion in public life is distinct from income redistribution, it is instructive to consider the United States Bill of Rights and similar documents or unwritten constitutional rules of other countries. The rights afforded by these documents are intended to apply equally to the poor and rich; they were not created with reference to income levels, but with reference to individual lives. There may, of course, be an interaction of the extent of redistribution such as that called for by a traditional social plan and of coercion, an issue to be dealt with later in the paper, but this only reinforces the insight that coercion arises from a variety of sources.⁴

³ "Politics is misconceived if it is interpreted as a separable allocation of costs and distribution of benefits. In an ultimate sense, collective action embodies an exchange" The costs are borne in order to secure the promised benefit flows." (Buchanan, 2001, 588). Buchanan goes on to acknowledge the imperfection of the exchange.

⁴ In an interesting paper that is complementary to the current one, Perroni and Scharf (2003) develop a positive theory of the self-enforcing fiscal system. The problem they begin with is that there is no external power to enforce the power to tax so that ultimately, in their view, all fiscal systems must be self-enforcing equilibria where the continual consent of the public (to the degree of taxing power, or power of the state to coerce) is sought. They search for efficient, self-enforcing equilibria that are robust to renegotiation among groups of citizens in future periods. As a consequence, they claim (result 4) that when citizens have identical preferences, efficiency and renegotiation proofness requires horizontal equity in taxation. But, as they explicitly state, this is "fully unrelated to any

One may ask more directly why maximization of social welfare, defined as the weighted or unweighted sum of utilities, does not deal with coercion even though the difference between benefits and costs for each individual is reflected in individual utility and therefore in the social objective. The reason is that the approach posits no limits on the loss or gain in utility for particular individuals that may occur as part of a social plan if this is required for social welfare maximization. (This will be shown formally below when the coercion-constrained social planner's optimization problem is explicitly stated and compared to a traditional planning problem.) As a consequence, the practical importance of coercion in public life is ignored, leading to ad hoc adjustments to social plans when policy advisors are forced to take 'political realities' into account.

It might be argued that application of the Pareto criterion - that only reallocations that leave every one better off are permissible - rather than the use of a social welfare function, can attenuate concern with coercion. Strict application of the Pareto criterion limits the degree of individual coercion for alternative moves from the status quo. It does not, however, alleviate any mismatch between benefits received and taxes paid that is embedded in the status quo itself. Moreover, it should be recalled that in much applied work, social welfare analysis goes beyond the strict Pareto criterion, which is too weak to allow for most social action, using the Hicks-Kaldor criterion of potential compensation instead. In that case, reallocations are desirable and permitted even if some people are worse off as long as gainers could *in principle* compensate losers and still be better off. For this reason, an explicit concern with coercion is justified and needed in most practical instances.

2.1 Our approach to modeling the role of coercion

We stated in the Introduction that to develop a normative approach that allows us to compare and evaluate specific policies, we proceed by imposing specific coercion constraints on maximization of social welfare as usually defined. We shall then investigate the nature of public policies that emerge from maximization of social welfare subject to such constraints and compare them to policies that are consistent with the traditional social planning approach. In this way we use an expanded approach to social planning that still falls into the neo-classical tradition in public economics.

Is it reasonable to use coercion constraints to acknowledge the importance of collective goods coercion rather than to reformulate the social welfare function that is to be maximized? Consider an analogy to modeling the social role of money. Macroeconomists have tried to come to grips with the role of money in society either by putting money into the utility function following Patinkin (1965), an obvious approximation to the social role of money, or by adding constraints to the specification of the economy (e.g. the cash-in-advance constraints of Robert Clower 1967). Our approach is analogous to the second method. We add coercion constraints to a planning problem in order to incorporate an important aspect of collective choice in a simple manner.

Using a constraint to deal with coercion also has a second justification. The most important way in which limits on coercion have been introduced into political arrangements is by constitutional provisions that restrict the power of government to abridge certain rights. Such provisions do not in principle allow for a trade-off between the rights that are given and other policy objectives. They may, of course, be subject to interpretation by the courts, but always with the understanding that the rights take precedence over other public aims. Setting of boundaries or constraints on public action

distributional goal" (p. 9). Rather, in their approach it is a matter of insuring the stability or viability of society as a whole.

thus represents a well-known and tested approach to dealing with coercion in public life.

In the analysis that follows, we distinguish between definitions of coercion on an individual basis, and those that are analogous to the use of Hicks-Kaldor potential compensation criterion, since collective goods coercion will be of heightened concern when the planner is allowed to make trade-offs among individuals without explicit compensation.

One should acknowledge that coercion may serve as a method of reducing the excess burden of taxation, thus having a productive as well as a harmful social role. In this paper, we assume that the state will coerce individuals up to the maximum possible, so that coercion constraints are generally introduced as equalities. We do not incorporate the possibility of a direct productive value for coercion in other ways. Excess burdens are defined in the usual manner, independently of the degree of coercion.⁵ The emphasis is on defining coercion arising from the collective provision and financing of collective goods and on investigating how an acknowledgment of limits to such coercion alters the structure of the fiscal system in a broader social welfare framework.⁶

3. Defining Coercion

We have so far informally introduced coercion for any citizen as the difference between what he or she gets in public services and what they pay in taxes. This will prove useful in approximating coercion. More formally, we shall define coercion of an individual as the difference between this person's utility under what he or she regards as *appropriate* treatment by the public sector, and the utility they actually enjoy as a result of social planning.⁷

To make this definition concrete, it is necessary to define what appropriate treatment means. There are at least two approaches to this issue, each of which corresponds to a particular perspective on the relationship between the individual and the state. Both are illustrated in Figure 1. One possibility is to think of the individual as judging social outcomes from a perspective in which they alone decide what is best for them and for others. In this '*individual-as-dictator*' approach, the counterfactual utility, denoted V_j^D , is defined by maximizing this person's (indirect) utility subject to the government budget restraint $t(G)$ that shows all feasible combinations of income tax rates and actual public good levels. Coercion can then be defined as $[V_j^D - V_j^a]$, where the 'a' denotes the actual level of utility. The corresponding counterfactual level of the public good is G_j^D .

[Figure 1 here]

An alternative definition sees the individual as a social being who does not pine for dictatorial outcomes. In this '*individual-in-society*' approach, the counterfactual is defined by the maximum individual utility attainable if the individual could adjust the level of the public good at the tax-prices

⁵ For example, we do not explicitly allow the planner to force independent evaluations of ability on taxpayers, or to coercively uncover potential for, or actual economic activity, thereby relaxing incentive compatibility constraints by the application of coercive means.

⁶ Kaplow (forthcoming) argues that what may appear to be a departure from a welfarist framework, in many cases is instead an attempt to come to terms with broader aspects of it. And this is so in our case

⁷ Breton (1974) defines coercion as depending on the deviation of marginal evaluations of public services from tax-prices. While (the total amount of) coercion as defined below varies with this difference, as we show below, it is not coercion itself.

they actually face. To illustrate this case, we let the individual's actual average tax-share be $\tau_j = (T_j/G)$ where T_j is his total tax payment, and assume, as in Buchanan (1967) and Breton (1974), that this tax-share is the one that the individual thinks he would pay if he could quantity-adjust. Then the relevant tax rate as a fraction of income for defining the counterfactual is $t_j = (\tau_j P/Y_j)G$, where P is the supply price of the public good - so $\tau_j P$ is the tax-price - and Y_j is income. The counterfactual level of utility then can be shown in Figure 1 as V_j^* where utility is maximized subject to the tax-share line with slope $(\tau_j P/Y_j)$. Coercion is $[V_j^* - V_j^a]$ with G_j^* the corresponding counterfactual level of the public good.

More formally, in the individual-in-society approach coercion for an individual is

$$[V_j^*(G^*, W_j, \tau_j P) - V_j^a] \quad \text{where } G_j^* = \arg \max_{\{G\}} V_j(G, W_j, \tau_j P), \quad (1)$$

where, in addition to the definitions above, $V_j^a =$ actual utility $V(G^a, W_j, \tau_j P)$; and $W_j =$ the wage or ability of individual j . We shall primarily use this approach to defining a counterfactual in our analysis. It is the one that is implicit in the work of Wicksell and Lindahl, Buchanan and Breton. It treats the individual as accepting of some coercion by society, along with a socially determined tax-price. Nonetheless, we shall note at various points how the individual-as-dictator approach would alter results. It turns out that both definitions carry with them similar implications for optimal fiscal structure and the coercion-welfare trade-off.

Before we can specify the coercion constraints that we shall impose on the social planner, there are two additional dimensions of the definition of coercion to consider: First, coercion can be defined on an individual basis as implied by the preceding discussion, or it can also be defined for a group. While applying constraints to each individual is consistent with the tradition initiated by Wicksell and Lindahl, we also want to explore the use of a definition based on a group approach that allows for a great degree of coercion. Although there isn't a complete parallel, defining coercion over a group of individuals is similar to the use of the Hicks-Kaldor potential compensation criterion (instead of a strict Pareto criterion) which allows for stronger policy judgments.⁸

Second, coercion can be defined using utility or, as proves useful in working out some examples, we can follow Buchanan (1967) and Breton (1974, 1996) and approximate changes in utility using levels of the public good.

The resulting four possibilities, each of which depends on the choice of a counterfactual, are summarized as coercion constraints in Table 1, where only the individual-in-society approach to the counterfactual is employed for illustrative purposes. Here we use K_j for the 'degree' of coercion applied to individual j (and, later, κ for the associated Lagrange multiplier) because in Greek, the word for coercion is '*katanagasmos*'. We note that although inequality constraints are used in the specification of the constraints in the table, in our formal analysis we assume that equalities apply and

⁸ It is interesting to note that Becker (1983) has proposed a *positive* theory of political outcomes in a democratic state that combines Hicks-Kaldor potential compensation with an understanding of *actual or existing* inequalities in political influence. If gainers from a policy action gain more than the loss to losers, they will, according to his argument, spend more and be more influential in the political process, unless there is some inequality in the distribution of political influence that favours the losers. Here we explore a normative theory that links the Hicks-Kaldor criterion with normatively desirable constraints on the extent to which gainers or losers *should be* coerced by the public sector.

that solutions are interior with respect to all constraints.

[Table 1 here]

It should be noted that for each case represented in the Table, the counterfactual level of utility and of the public good, individualized tax-prices, and coercion must all be simultaneously determined because tax-prices in part determine coercion and, in turn, the degree of coercion will be taken into account in deciding upon the tax system and its implied tax-prices.

3.1 Coercion defined by levels of the public good

Table 1 indicates that coercion can be approximated using the level of the public good, either on an individual or on an aggregate basis. The argument is illustrated in Figure 2. Because the marginal evaluation of the public good declines with the size of the public sector, the difference in utility in (1) above is monotonically related to the difference between the level of the public good in the counterfactual and that provided by the planner, ($G^* - G$). Thus we can use the difference in public good levels as an index of coercion. This will be so regardless of whether the individual prefers a level of the public good (given his or her tax-price $\tau_i P$) that is lower or higher than that determined by the planner.

We shall treat people who want less of the public good in the counterfactual symmetrically with those who would like more. The first type of citizen, illustrated in Figure 2, is losing utility because they would like a lower quantity of the good at the tax-price they face. The second type also fails to get their desired amount, but gains from the fact that they would be willing to pay more for what they receive than they are required to pay. Although one may argue that the first type are those who are 'coerced' in the popular sense of the word, we adopt a broader perspective that relates to the overall functioning of the system. All differences between what people would like and what they get are perceived as damaging shall be recognized in choosing a fiscal system. If deviations in either direction become too large, the collective choice that is being represented here by the incorporation of coercion constraints loses its legitimacy. To allow for both cases, we may use the absolute value of the difference between hypothetical levels of G and planned levels, as shown in case 3 of Table 1. Case 4 in Table 1 is the analogue to case 2 where an aggregate definition of coercion is used.

[Figure 2 here]

4. Coercion-Constrained Optimal Linear Income Taxation with a Public Good

We now show how acceptance of coercion constraints alters the welfare analysis of a fiscal system in which a pure public good is financed with a linear income tax.⁹ In the Appendix we show how the Ramsey (1927) analysis of the structure of commodity taxation is altered by the introduction of a concern with coercion.

For convenience in what follows, we drop the superscript 'a' when referring to actual quantities as distinct from those in a counterfactual. Unless otherwise noted, we adopt the individual-in-society

⁹ With $a = 0$ the tax is proportional to income, while with $a > (<) 0$ the tax is progressive (regressive). For convenience, here and below we follow Sandmo's (1998) notation as far as possible.

definition of the counterfactual used to define coercion.

Assume there are N individuals indexed by j , each maximizing utility defined over a private good X_j , leisure L_j and a public good G :

$$U_j = U_j(X_j, L_j, G), \quad j = 1, \dots, N. \quad (1)$$

Utility is maximized subject to the budget constraint

$$X_j = (1-t)W_j(1-L_j) + a \quad (2)$$

where W_j denotes the wage rate or ability and H_j denotes the supply of labour, with $L_j + H_j = 1$. The fact that the lump sum component a does not vary across individuals is a simple way of introducing the excess burden of taxation, and also of ruling out a Lindahl voluntary exchange equilibrium in which taxes are raised without any welfare loss.

Maximization of utility subject to the individual's budget constraint yields the usual condition $U_{jL}/U_{jX} = (1-t)W_j$, final demand for the private good $X_j = X_j[(1-t)W_j, a, G]$, labor supply $H_j = H_j[(1-t)W_j, a, G]$, and indirect utility function $V_j = V_j[(1-t)W_j, a, G]$.¹⁰

4.1 Establishing the counterfactual

To establish the counterfactual, we consider the individual when he is free to choose the level of the public good G_j for a given tax share, assumed to be constant with respect to the level of the public good. His optimization problem is to maximize

$$U_j = U_j(X_j, L_j, G_j^*), \quad (1')$$

subject to the budget

$$X_j + \tau_j P G_j = W_j(1-L_j), \quad (2')$$

where $\tau_j P$ is the tax-price per unit, with P the unit cost of the public good, and τ_j the tax share of person j , defined as the ratio of the tax paid by j to the total tax revenue, $\tau_j = T_j / \sum_j T_j$.

We note for later use that with the linear tax system $T_j = tY_j - a$, this tax share can be written as

$$\tau_j = \frac{t Y_j - a}{t \sum_j Y_j - Na}. \quad (3)$$

¹⁰ Denoting the marginal utility of income by λ_j , the partial derivatives of utility with respect to the fiscal variables are: $V_{jL} = -\lambda_j W_j$; $V_{ja} = -\lambda_j$ and $V_{jG} = U_{jG}$, and the marginal willingness to pay for the public good is $m_j = U_{jG}/U_{jX} = V_{jG}/\lambda_j$.

It should be pointed out that there are many ways to perform the translation of t into τ besides using (3). We assume, here and below, that the average tax price implied by the tax system is also the one that applies to marginal changes in public services when viewed from the perspective of each individual.

Maximization of (1') subject to (2') yields the usual first order conditions $U_{jX} = \lambda_j^*$, $U_{jL} = \lambda_j^* W_j$, and $U_{jG} = \lambda_j^* \tau_j P$, the indirect or *counterfactual* level of utility $V_j^* = V_j^*[W_j, \tau_j P]$, and also the effect of a change in the tax share on utility $V_{j\tau}^* = -\lambda_j^* P G_j^*$. Again the (*) reflects the fact that the individual is considered to be choosing his or her most preferred level of G at the *given* tax-price.

4.2 Social welfare maximization under aggregate coercion

In choosing fiscal policy instruments, the coercion-constrained planner is assumed to maximize the sum of individual utility functions,

$$S = \sum_j V_j \quad (4)$$

subject to the usual budget constraint of the government

$$t \sum_j W_j H_j - Na = PG. \quad (5)$$

In addition, the planner faces one or more coercion constraints. To begin, we consider case 2 in Table 1 where coercion is defined using utility levels and aggregated across individuals. As we have pointed out earlier, this case is analogous to the use of the Hicks-Kaldor criterion in cost-benefit analysis.

Let κ denote the Lagrangean multiplier of the coercion constraint. Then incorporating the coercion constraint, the problem is to maximize

$$\mathcal{L} = \sum_j V_j + \mu [t \sum_j W_j H_j - Na - PG] + \kappa [\sum_j (V_j^* - V_j) - K]. \quad (6)$$

Before proceeding to explore the solution, it is useful to point out here that this Lagrangean illustrates clearly that acknowledging coercion does not amount to simply placing added weight on the utility of some individuals in a social plan, or, in other words, the Lagrangean shows that redistribution and coercion are not equivalent concepts. In (6) it can be seen that a concern with coercion importantly requires that weight be given to the *counterfactual* level of utility for each individual.¹¹

Differentiating (6) with respect to policy instruments t , a and G and using also the definition of $V_{j\tau}^*$ we have first order conditions:

$$(1-\kappa) \sum_j \lambda_j W_j H_j + \kappa \sum_j \lambda_j^* P G_j^* (\partial \tau_j / \partial t) = \mu [\sum_j W_j H_j + t \sum_j W_j (\partial H_j / \partial t)] \quad (7.1)$$

$$(1-\kappa) \sum_j \lambda_j - \kappa \sum_j \lambda_j^* P G_j^* (\partial \tau_j / \partial a) = \mu [N - t \sum_j W_j (\partial H_j / \partial a)] \quad (7.2)$$

$$(1-\kappa) \sum_j \lambda_j m_j = \mu [P - t \sum_j W_j (\partial H_j / \partial G)] \quad (7.3)$$

where m_j is the marginal rate of substitution between public and private goods.

¹¹ One may note also that if there is only one person, or if everyone is identical in all respects, at an optimum there will be no difference between V^* and V , and any coercion constraint will be irrelevant. Coercion has no meaning in a single agent social planning model.

These equations feature two new important elements that are absent from traditional optimal taxation and which are always present in the analysis of collective goods coercion, as long as one uses the individual-in-society approach to the counterfactual: (a) the translation of tax structure into the tax price - here shown as $(\partial\tau_j/\partial t)$ and $(\partial\tau_j/\partial a)$; and (b) the translation of the tax price into the demand for G - here shown as $\lambda^*_j P G^*_j$. These two elements always appear because coercion involves the difference between what people pay at given tax-prices and what they get in public services. These elements do *not* appear if one uses the individual-as-dictator approach, since then the individual picks a fiscal system directly from the set of solutions consistent with the government budget constraint.

Using the covariance formula $\sigma^2_{\lambda m} = (\sum_j \lambda_j m_j / N) - (\sum_j \lambda_j / N)(\sum_j m_j / N)$ and setting $\lambda = \sum_j \lambda_j / N$, the mean value of the marginal utility of income, (7.3) can be rewritten as

$$(1-\kappa)[N\sigma^2_{\lambda m} + \lambda\sum_j m_j] = \mu [P - t\sum_j W_j (\partial H_j / \partial G)].$$

Denote $m = \sum_j m_j / N$, the mean value of the marginal rate of substitution between public and private goods, and let $\delta \equiv \sigma^2_{\lambda m} / \lambda m$ be the normalized covariance between the marginal rate of substitution and the marginal utility of income, reflecting the distributional characteristics of the public good. Then using these definitions in the above equation and manipulating yields

$$(\sum_j m_j)(1+\delta)(1-\kappa) = \frac{\mu}{\lambda} \left(P - t\sum_j W_j \frac{\partial H_j}{\partial G} \right). \quad (8)$$

Equation (8) represents the condition for optimal provision of the public good when the planner cannot breach a coercion constraint. It is a generalization of the Samuelson condition as amended by Atkinson and Stern (1974) to allow for a concern with coercion. It turns out that this condition is unaffected by the choice of the counterfactual.¹²

The left-hand-side represents the social marginal benefit from public provision in the presence of the coercion constraint. It is the product of three terms: (a) the sum of the marginal rates of substitution between the private and the public good; (b) term $1+\delta$, which adjusts the sum of MRSs for the distributional characteristics of its provision – captured by the expression $\sigma^2_{\lambda m} / \lambda m$; and (c) the term $(1-\kappa)$, which reflects the effect of the coercion constraint.

The right-hand-side represents the social marginal cost of public provision. It is the product of two components, familiar from the standard social welfare maximization case: (a) The marginal valuation of government revenue, μ/λ (see Sandmo 1998 or Atkinson and Stern); and (b) The net marginal rate of transformation of the public good, $P - t\sum_j W_j (\partial H_j / \partial G)$, which equals the unit cost of production of the public good adjusted for the effects of public provision on income and therefore on income taxation .

The standard rule of optimal public good provision can be derived as a special case when it is assumed that the planner does not observe a coercion constraint, so that $\kappa=0$. In this case

¹² To see this last result, one solves for indirect utility under the assumption that the individual chooses t , a and G directly, and proceeds as before. Note that in this case, it is crucial that in the counterfactual, the partial derivatives of indirect utility with respect to fiscal parameters are zero since the individual chooses fiscal structure directly to maximize his or her welfare.

$$(\sum_j m_j)(1 + \delta) = \frac{\mu}{\lambda} \left(P - t \sum_j W_j \frac{\partial H_j}{\partial G} \right). \quad (8')$$

Comparing (8) with (8') we observe that since κ is positive, $\sum_j m_j$ must be larger to maintain the equality in (8), and thus G and t must be lower in the presence of a coercion constraint than in the standard planning solution, and this is so whichever definition of the counterfactual is adopted. Assuming δ is positive, the reason is that the declining marginal rate of substitution between public and private goods for all individuals implies that those who are less coerced when G is increased gain less than do those who want less G when G is decreased.

The comparison also shows that the coercion adjusted marginal cost of funds (MCF) that is appropriate for policy analysis when one is formally cognizant of coercion is the term $\mu / [\lambda(1 - \kappa)]$. It must be higher than when coercion is ignored as long as κ is positive, as we should expect. But note that the MCF as a concept is still relevant here.

To derive the optimal income tax rate in the presence of coercion, we proceed as follows. Multiplying (7.1) by $(1/N)$ and (7.2) by $(\sum_j W_j H_j / N^2)$, subtracting the latter from the former¹³, and using the Slutsky decomposition, $\partial H_j / \partial t = s_j - W_j H_j (\partial H_j / \partial a)$, yields

$$(1 - \kappa) \sigma_{\lambda Y}^2 = t \frac{\mu}{N} \left[\sum_j W_j s_j - \sum_j W_j H_j \frac{\partial H_j}{\partial a} + \left(\sum_j W_j \frac{\partial H_j}{\partial a} \right) \left(\frac{\sum_j W_j H_j}{N} \right) \right] - \frac{\kappa}{N} \sum_j \lambda^*_{j} PG^*_{j} \left(\frac{\partial \tau_j}{\partial t} + \frac{\partial \tau_j}{\partial a} \bar{Y} \right) \quad (9)$$

where $Y_j = W_j$, and the negative covariance $\sigma_{Y\lambda}^2$ shows the relationship between the marginal utility of income and income from work and reflects the distributional effects of income taxation.¹⁴ The term $\bar{WS} = \sum_j W_j s_j / N$ is the mean substitution effect of taxation on labor supply, which is also negative, and the covariance term

$$\sigma_{Ya}^2 = \frac{1}{N} \left[\sum_j W_j H_j \frac{\partial H_j}{\partial a} - \left(\sum_j W_j \frac{\partial H_j}{\partial a} \right) \left(\frac{\sum_j W_j H_j}{N} \right) \right]$$

shows the relationship between income and the income effect of taxation; it is non-negative when those with large income are characterized by small income effects on labor supply.

The quantity $q_j = \partial \tau_j / \partial t + (\partial \tau_j / \partial a) \bar{Y}$ in (9) is the change in the tax share of j given by (3) when the tax rate and the lump-sum transfer both change. If we let $\psi_j \equiv V^*_{j\tau} = \lambda^*_{j} PG^*_{j}$ be the marginal utility of the tax share, we can then write $\sum_j \lambda^*_{j} PG^*_{j} [\partial \tau_j / \partial t + (\partial \tau_j / \partial a) \bar{Y}] = \sum_j \psi_j q_j$. Using the covariance formula, the right side of this last expression is $\sum_j \psi_j q_j = N \sigma_{\psi q}^2 + N \bar{\psi} \bar{q}$, where $\bar{\psi}$ and \bar{q} denote the mean values of ψ_j and q_j respectively. Here $\sigma_{\psi q}^2$ shows the relationship between the marginal utility of

¹³ Yielding as an intermediate step:

$$(1 - \kappa) \left(\frac{\sum_j \lambda_j W_j H_j}{N} - \frac{\sum_j \lambda_j}{N} \frac{\sum_j W_j H_j}{N} \right) = t \frac{\mu}{N} \left[\left(\sum_j W_j \frac{\partial H_j}{\partial t} \right) + \left(\sum_j W_j \frac{\partial H_j}{\partial a} \right) \left(\frac{\sum_j W_j H_j}{N} \right) \right] - \frac{\kappa}{N} \sum_j \lambda^*_{j} PG^*_{j} \left(\frac{\partial \tau_j}{\partial t} + \frac{\partial \tau_j}{\partial a} \bar{Y} \right).$$

¹⁴ $\sigma_{Y\lambda}^2$ is negative since the higher the level of income the lower is the marginal utility.

the tax share and the marginal tax share¹⁵. Also from differentiating the tax share, we have after some work¹⁶

$$\sum_j \frac{\partial \tau_j}{\partial t} = \sum_j \frac{\partial \tau_j}{\partial a} \bar{Y} = 0, \quad (10)$$

and thus $\bar{q} = 0$.

Substituting (10) into (9) and using the definition of $\sigma_{\psi q}^2$ then leads to the coercion-constrained optimal income tax rate:

$$t = \frac{(1 - \kappa) \sigma_{\lambda Y}^2 - \kappa \sigma_{\psi q}^2}{\mu (\overline{WS} - \sigma_{Y_a}^2)}. \quad (11)$$

We see that the optimal tax rate is decreasing in the marginal utility of the coercion constraint κ . That is, the less social solidarity or tolerance for coercion, the lower the optimal income tax rate. Given that both $\sigma_{\lambda Y}^2$ and the denominator are negative, the optimal tax rate will be lower if $\sigma_{\psi q}^2$ is positive; that is, if the marginal utility of the tax share rises with the tax share.

The standard linear optimal tax rate can be obtained as a special case of (11) when $\kappa = 0$:

$$t = \frac{\sigma_{\lambda Y}^2}{\mu (\overline{WS} - \sigma_{Y_a}^2)}. \quad (11')$$

By comparing (11) and (11') we see that the more general formulation of the optimal income tax rate features two new terms in comparison to the standard formula: (i) the marginal utility of the coercion constraint, κ ; and (ii) the covariance of marginal utility of the tax share and the marginal tax share, $\sigma_{\psi q}^2$.¹⁷ Were the individual-as-dictator approach to the counterfactual to be adopted, the optimal tax rate would then no longer depend on the covariance between the marginal utility of the tax share and the tax share $\sigma_{\psi q}^2$, since the tax share is not relevant to the definition of that counterfactual.

¹⁵ The value of $\sigma_{\psi q}^2$ depends on the size of the parameters of the utility function and is therefore an empirical matter. If tax payers who experience a large increase in their tax shares will also experience a significant fall in utility, $\sigma_{\psi q}^2$ will be negative.

¹⁶ Differentiating τ_j in (3) with respect to the income tax rate t and the lump-sum transfer a , and recognizing that a change in t and a affects the level of income,

$$\frac{\partial \tau_j}{\partial t} = \frac{t^2 (\bar{Y} Y_{jt} - Y_j \bar{Y}_t) + a (\bar{Y} - Y_j) + at (\bar{Y}_t - Y_{jt})}{N(t\bar{Y} - a)^2} \quad \text{and} \quad \frac{\partial \tau_j}{\partial a} = \frac{t(Y_{jt} \bar{Y} - Y_j \bar{Y}_{jt}) + Y_j - \bar{Y} + a(Y_{jt} - \bar{Y}_a)}{N(t\bar{Y} - a)^2} \quad \text{where } \bar{Y} = \sum_j Y_j / N \text{ is mean income.}$$

Equation (10) follows from these expressions.

¹⁷ For completeness, we can contrast this result with the opposite extreme, when the planner must not allow *any* coercion, i.e. when $\sum_j (V_j^* - V_j) = 0$. In this case each citizen/consumes the size of the public good which maximizes his or her utility given their tax share. Attaining such a Lindahl-like solution requires that all individuals consume the *same* level of the public good. We call this a Lindahl-like solution since taxation is still imposed by the planner rather than representing the result of voluntary exchange.

4.3 Individual coercion constraints

When the planner is constrained by how much it can coerce each individual taxpayer separately (as in case 1 of Table 1), the Lagrangean for the planning problem now becomes

$$\mathcal{L} = \sum_j V_j + \mu [t \sum_j W_j H_j - N a - P_G G] + \sum_j \kappa_j (V_j - V_j - K_j). \quad (12)$$

The resulting first order conditions,

$$\sum_j (1 - \kappa_j) \lambda_j W_j H_j + \sum_j \kappa_j \lambda_j^* P_G G^* (\partial \tau_j / \partial t) = \mu [\sum_j W_j H_j + t \sum_j W_j (\partial H_j / \partial t)] \quad (12.1)$$

$$\sum_j (1 - \kappa_j) \lambda_j - \sum_j \kappa_j \lambda_j^* P_G G^* (\partial \tau_j / \partial a) = \mu [N - t \sum_j W_j (\partial H_j / \partial a)] \quad (12.2)$$

$$\sum_j (1 - \kappa_j) \lambda_j m_j = \mu [P_G - t \sum_j W_j (\partial H_j / \partial G)] \quad (12.3)$$

are considerably more complicated than what we have encountered before because they feature three additional distributions: of the κ_j , λ_j , and m_j .

Like $s_{\lambda m} \equiv \sigma_{\lambda m}^2 / \lambda m$ ($\equiv \delta$), which denotes the normalized covariance between the marginal rate of substitution and the marginal utility of income, we may define $s_{\kappa \lambda} \equiv \sigma_{\kappa \lambda}^2 / \lambda \kappa$ as the normalized covariance between coercion and the marginal utility of income; $s_{\kappa m} \equiv \sigma_{\kappa m}^2 / \kappa m$ as the normalized covariance between coercion and the marginal rate of substitution; and $s_{\kappa \lambda m} \equiv \sigma_{\kappa \lambda m}^2 / \kappa \lambda m$ as the normalized covariance between coercion, the marginal utility of income and the marginal rate of substitution. Letting the sum of these normalized covariance terms be

$$\phi \equiv s_{\lambda m} + s_{\kappa \lambda} + s_{\kappa m} + s_{\kappa \lambda m}, \quad (13)$$

substituting into (12.3) and manipulating,¹⁸ then gives

$$\left(\sum_j m_j \right) \left[(1 + \delta)(1 - \kappa) - \kappa \phi \right] = \frac{\mu}{\lambda} \left(P_G - t \sum_j W_j \frac{\partial H_j}{\partial G} \right). \quad (14)$$

The left-hand-side of (14) again shows the marginal benefit from the public good. But now it is the product of the sum of marginal rates of substitution times the adjustment for the combined effect of the distributional characteristics of the public good *and* the effects of coercion. In the present case of individual coercion constraints, the adjustment for coercion contains two new elements, (i) the average effect of coercion captured by the term $(1 + \delta)(1 - \kappa)$, a term that also appears in the previous case of aggregate coercion, and (ii) the distributional characteristics of coercion captured by the term $\kappa \phi$, which corrects the aggregate term for the distributional characteristics of coercion. The right-hand-side of equation (14) is already familiar: it is the product of the marginal valuation of government revenue times the net marginal rate of transformation of the public good.

The solution in (14) differs from the solution in (11) where coercion is defined on an aggregate basis by

¹⁸ The left-hand-side of (12.3) is written as $\sum_j (1 - \kappa_j) \lambda_j m_j = \sum_j \lambda_j m_j - \sum_j \kappa_j \lambda_j m_j$. Recall that κ , λ , and m are the means of κ_j , λ_j , and m_j respectively, $\text{cov}(\kappa_j \lambda_j) = (1/N) \sum_j (\lambda_j - \lambda)(\kappa_j - \kappa)$ and $\text{cov}(\kappa_j \lambda_j m_j) = (1/N) \sum_j (\kappa_j - \kappa)(\lambda_j - \lambda)(m_j - m)$. Manipulating, these covariances and using (12.3) yields as intermediate steps: $\sum_j \lambda_j m_j = N \sigma_{\lambda m}^2 + \lambda \sum_j m_j$ and $\sum_j \kappa_j \lambda_j m_j = [\kappa \lambda \sum_j m_j + N(\kappa \sigma_{\lambda m}^2 + \lambda \sigma_{\kappa m}^2 + \sigma_{\kappa \lambda}^2 + \sigma_{\kappa \lambda m}^2)]$.

the term $\kappa\phi$, reflecting the fact that not just aggregate coercion matters but also its distribution. Now the benefit from public provision falls: (i) If the rich (low λ) have more social solidarity (low κ), so that $\sigma^2_{\kappa\lambda} > 0$ and $s_{\kappa\lambda} > 0$; (ii) If those who value public goods less (low m) have more social solidarity (low κ), then $\sigma^2_{\kappa m} > 0$ and $s_{\kappa m} > 0$; and (iii) If the rich (low λ) also value public goods less (low m) and have more social solidarity (low κ), then $\sigma^2_{\kappa\lambda m} > 0$ and $s_{\kappa\lambda m} > 0$, because the previous two effects are compounded.

In these cases, G (and t) in (14) will be less than when coercion is defined in an aggregate sense. That is, the Hicks-Kaldor-like solution for a coercion-constrained optimum (11) involves more spending than when coercion is defined on an individual basis. However, unlike the aggregate coercion case with positive δ , it could go the other way in principle, and it will be interesting to try and figure out in practice which case is likely to apply.

In order to derive the optimal income tax rate under individual coercion constraints, we work in a manner similar to the one before. Multiplying (12.1) by $(1/N)$ and (12.2) by $(\sum_j W_j H_j / N^2)$ and subtracting the latter from the former we have

$$\left(\frac{\sum_j \lambda_j W_j H_j}{N} - \frac{\sum_j \lambda_j}{N} \frac{\sum_j W_j H_j}{N} \right) - \left(\frac{\sum_j \kappa_j \lambda_j W_j H_j}{N} - \frac{\sum_j \kappa_j \lambda_j}{N} \frac{\sum_j W_j H_j}{N} \right) = t \frac{\mu}{N} \left[\left(\sum_j W_j \frac{\partial H_j}{\partial t} \right) + \left(\sum_j W_j \frac{\partial H_j}{\partial a} \right) \left(\frac{\sum_j W_j H_j}{N} \right) \right] + \frac{1}{N} \sum_j \kappa_j (-\lambda^*_j PG^*_j) \left(\frac{\partial \tau_j}{\partial t} + \frac{\partial \tau_j}{\partial a} \bar{Y} \right). \quad (15)$$

The left-hand-side involves the frequency distributions of three variables, the individual coercion constraint κ_j , the marginal utility of income λ_j , and income Y_j . Similarly, the right-hand-side features the individual coercion constraint κ_j , the marginal utility of the tax share $\psi_j = -\lambda^*_j PG^*_j$, and the marginal tax share, $q_j = [(\partial \tau_j / \partial t) + (\partial \tau_j / \partial a) \bar{Y}]$, as well as the effect of income tax on labour supply.

Let $\sigma^2_{\kappa\lambda Y}$ = covariance of κ_j , λ_j and Y_j ; $\sigma^2_{\kappa\psi q}$ = covariance of κ_j , ψ_j and q_j ; $\sigma^2_{\kappa Y}$ = covariance of κ_j and Y_j ; $\sigma^2_{\psi q}$ = covariance of ψ_j and q_j ; $\sigma^2_{\kappa q}$ = covariance of κ_j and q_j ; and κ , λ , ψ and q are the mean values of κ_j , λ_j , ψ_j and q_j respectively. Then using these definitions and the Slutsky equation gives the formula for the coercion-constrained optimal income tax rate:

$$t = \frac{(1 - \kappa) \sigma^2_{\lambda Y} - \kappa \sigma^2_{\psi q} - \lambda \sigma^2_{\kappa Y} - \psi \sigma^2_{\kappa q} - \sigma^2_{\kappa\lambda Y} - \sigma^2_{\kappa\psi q}}{\mu (\bar{WS} - \sigma^2_{Y a})}. \quad (16)$$

The optimal income tax rate depends as usual on the income distribution effect of taxation, captured by $\sigma^2_{\lambda Y}$, and the efficiency effect of taxation on labour (shown again by the denominator). In common with the case of aggregate coercion, it also depends on the relationship between the marginal utility of the tax share and the marginal tax share $\sigma^2_{\psi q}$.

In addition, the optimal tax rate depends on the distribution effects of coercion as the remaining four covariance terms make clear. The optimal tax rate will be lower than when coercion is defined on an aggregate basis: (i) If the rich (high Y) value social solidarity more (low κ) so that $\sigma^2_{\kappa Y} < 0$; (ii) If taxpayers who experience a large change in their tax share (high q) have low social solidarity (high κ), so that $\sigma^2_{\kappa q} > 0$ (since ψ is negative); (iii) If the rich (high Y and low λ) have more social solidarity (low κ), in which case $\sigma^2_{\kappa\lambda Y} < 0$; and (iv) If $\sigma^2_{\kappa\psi q} < 0$, which is the case when (a) those who experience a

small change in the tax share suffer a large utility loss also have less solidarity, or (b) those who experience a large change in the tax share suffer a large utility loss also have more solidarity.

5. The Trade-off between Social Welfare and Coercion in Income Taxation

In the linear income tax example, there is an optimal level of welfare corresponding to each assumed level of coercion. In this section, we explore the trade-off between welfare and coercion that is implied by this relationship. To allow a closed form solution for the trade-off, we assume utility is Cobb-Douglas and we adopt an aggregate definition of coercion based on the level of the public good (case 4 in Table 1). Later we shall also aggregate individuals into two groups.

The utility function of taxpayer j is written as

$$U_j = \alpha_j \log X_j + \beta_j \log L_j + \gamma_j \log G. \quad (19)$$

Consumers are assumed to have different preferences for the private good, leisure and the public good so that $\alpha_j \neq \alpha_i$, $\beta_j \neq \beta_i$ and $\gamma_j \neq \gamma_i$, for $j \neq i$. For algebraic tractability we also assume that a proportional income tax system is applied at a single rate t . The budget constraint of the j th person taxpayer is then

$$X_j = (1-t)W_j(1-L_j). \quad (20)$$

Maximizing the utility subject to the budget constraint, solving for X_j and L_j and substituting back we derive for later use the following expressions for the partial derivatives of income and the indirect utility function with respect to t and G :

$$\begin{aligned} Y_j &= \alpha_j W_j / (\alpha_j + \beta_j) & Y_{jt} &= 0 & Y_{jG} &= 0 \\ V_{jt} &= -\lambda_j = -(\alpha_j + \beta_j) / W_j & V_{jt} &= -\alpha_j / (1-t) & V_{jG} &= \gamma_j; \quad j=1, \dots, N \end{aligned}$$

In deriving the welfare-coercion trade-off, it proves useful to first derive the traditional social planning solution. This is the fiscal system which maximize the social welfare function $\sum_j V_j$ subject to the aggregate government budget constraint $PG = t \sum_j W_j H_j$.

To further simplify, we divide the population into two homogeneous groups with taste parameters for the public good γ_1 and γ_2 , and wage rates W_1 and W_2 . The numbers of individuals in the two groups are N_1 and N_2 , with $N=N_1+N_2$. Denoting the mean value of α_j and γ_j by $\bar{\alpha}$ and $\bar{\gamma}$, the socially optimal level (o) of the public good and tax rate are

$$G^o = \frac{N_1 \gamma_1 + N_2 \gamma_2}{N_1 (\alpha_1 + \gamma_1) + N_2 (\alpha_2 + \gamma_2)} \left(\frac{N_1 \alpha_1 W_1}{\alpha_1 + \beta_1} + \frac{N_2 \alpha_2 W_2}{\alpha_2 + \beta_2} \right) \frac{1}{P} \quad (21.1)$$

and

$$t^o = \frac{N_1 \gamma_1 + N_2 \gamma_2}{N_1 (\alpha_1 + \gamma_1) + N_2 (\alpha_2 + \gamma_2)} = \frac{\bar{\gamma}}{\bar{\alpha} + \bar{\gamma}}. \quad (21.2)$$

As before, to define coercion we must define for the counterfactual what individuals would like to have

at their given tax-prices. To do so, we maximize utility (19) subject to the budget constraint

$$W_j(1-L_j) = X_j + \tau_j P G_j, \quad (20')$$

where the tax share is $\tau_j = tY_j/\sum_j tY_j = W_j/\sum_j W_j$.

Solving yields

$$X_j^* = \alpha_j W_j; \quad 1-L_j^* = 1-\beta_j \quad \text{and} \quad G_j^* = \frac{\gamma_j W_j}{\tau_j P} = \frac{\gamma_j \sum_j W_j}{P}. \quad (22)$$

Note that with Cobb-Douglas utility, the level of the public good desired at the given tax-price does not depend on the preference coefficients for private consumption and leisure, but only on the tastes for the public good.

5.1 The coercion implied by social planning, and the welfare-coercion trade-off

We now proceed by deriving the coercion-constrained social welfare optimum. This solution can then be used to calculate the degree of coercion implied by traditional social planning, which is of interest in its own right - it is *not* zero - and which is used in deriving the social-welfare-coercion trade-off.

With the population divided into two groups, it follows from (22) that the corresponding demands for the public good are

$$G_1^* = \frac{\gamma_1(N_1W_1 + N_2W_2)}{P} \quad \text{and} \quad G_2^* = \frac{\gamma_2(N_1W_1 + N_2W_2)}{P}. \quad (22')$$

Of course only a single level of the public good G is to be provided.

Assume that taxpayers in group 1 consume a quantity of the public good larger than the quantity that they would have chosen freely, that is, $G > G_1^*$, and that taxpayers in 2 consume less than they would like, that is, $G < G_2^*$. Coercion must be limited, so $(G - G_1^*)P \leq K_1$ and $(G_2^* - G)P \leq K_2$.

Summing the coercion constraints over the two groups, the aggregate coercion constraint is

$$N_1(G - G_1^*)P + N_2(G_2^* - G)P = N_1K_1 + N_2K_2. \quad (23)$$

Setting $K = N_1K_1 + N_2K_2$, substituting from (22), and manipulating, implies that the coercion-optimal size of G is¹⁹:

¹⁹ In the present setting, the planner possesses only two instruments, the income tax rate, t , and the size of public provision, G . As the two depend on each other through the budget constraint, there is a single free instrument, whose value is found by solving the aggregate coercion constraint. In the more general case of more than two instruments, their values are found by maximizing the social welfare function subject to the budget constraint and the coercion constraint.

$$G^C = \frac{(N_1W_1 + N_2W_2)(N_2\gamma_2 - N_1\gamma_1) - K}{(N_2 - N_1)P}. \quad (24.1)$$

For G^C to be positive, it must be that $N_2 - N_1 > 0$ and $N_2\gamma_2 - N_1\gamma_1 > K/(N_1W_1 + N_2W_2)$ or $N_2 - N_1 < 0$ and $N_2\gamma_2 - N_1\gamma_1 < K/(N_1W_1 + N_2W_2)$. In what follows we assume that the former two inequalities hold. In addition, we also assume that $\gamma_2 > \gamma_1$, so that $N_2\gamma_2 > N_1\gamma_1$. Since we are exploring case 4 in Table 1, coercion only exists if the two groups are not balanced with respect to size and preference for the public good.

Substituting (24) into the budget constraint of the government (written to reflect the existence of the two groups)

$$PG = t \left(\frac{N_1\alpha_1W_1}{\alpha_1 + \beta_1} + \frac{N_2\alpha_2W_2}{\alpha_2 + \beta_2} \right),$$

leads to the optimally coercive income tax rate

$$t^C = \frac{(\alpha_1 + \beta_1)(\alpha_2 + \beta_2)}{(\alpha_2 + \beta_2)N_1\alpha_1W_1 + (\alpha_1 + \beta_1)N_2\alpha_2W_2} \frac{(N_2\gamma_2 - N_1\gamma_1)[(N_1W_1 + N_2W_2) - K]}{N_2 - N_1}. \quad (24.2)$$

It is immediately seen that the less coercion society tolerates, the smaller the income tax rate and the smaller the size of public provision. In addition, we see that the more intense are the preferences for the public good of group 2, the higher is the optimal tax rate, and the lower the intensity of preferences for the public good by group 1, the lower the tax rate.²⁰

From (21.2) and (24.2) it is possible to find the level of coercion K_{OT} implied by the standard social planning solution, by setting $t^C = t^O$ and solving for K . This yields the level of coercion for which the coercion-constrained, welfare maximizing planner would levy the same tax rate as the unconstrained planner. Thus,

$$K_{OT} = (N_2\gamma_2 - N_1\gamma_1)N\bar{W} - (N_2 - N_1) \frac{\bar{\gamma}}{\alpha + \gamma} N\bar{Y}, \quad (25)$$

where a bar denotes a mean value of α_j , γ_j , W_j and Y_j .²¹ Note that our assumptions that $N_2 > N_1$ and $\gamma_2 > \gamma_1$, ensure that $K_{OT} > 0$.

Putting (24.1) and (24.2) into the individual utility function (19) and summing over all individuals leads to what we shall call '*coercion-constrained social welfare*', the basis for exploring the welfare-coercion trade-off :

²⁰ For completeness, we state the level of public provision when $K=0$. Then the solution requires that all taxpayers consume the same level of G , in a Lindahl-like equilibrium (L). Inserting $G^*_{-j} = G_L$ in (22) and substituting the resulting expression for τ_j into $\sum_j \tau_j = 1$ shows that in our example, $G^L = (\gamma\bar{W} + \sigma^2_{\gamma W})(N_1 + N_2)/P$.

²¹ $\bar{\alpha} = \frac{N_1\alpha_1 + N_2\alpha_2}{N_1 + N_2}$, $\bar{\gamma} = \frac{N_1\gamma_1 + N_2\gamma_2}{N_1 + N_2}$, $\bar{W} = \frac{N_1W_1 + N_2W_2}{N_1 + N_2}$ and $\bar{Y} = \left(\frac{N_1\alpha_1W_1}{\alpha_1 + \beta_1} + \frac{N_2\alpha_2W_2}{\alpha_2 + \beta_2} \right) \frac{1}{N_1 + N_2}$.

$$S(K) = A + (N_1\alpha_1 + N_2\alpha_2) \log(1-t^c(K)) + N_1\alpha_1 \log W_1 + N_2\alpha_2 \log W_2 + (\gamma_1 N_1 + N_2\gamma_2) \log G \quad (26)$$

where $A = N_1[\alpha_1 \log \alpha_1 + \beta_1 \log \beta_1 - (\alpha_1 + \beta_1) \log(\alpha_1 + \beta_1)] + N_2[\alpha_2 \log \alpha_2 + \beta_2 \log \beta_2 - (\alpha_2 + \beta_2) \log(\alpha_2 + \beta_2)]$.

Differentiating (26) with respect to K shows that the trade-off between S and K is concave with an inflection point at $K = K_{OT}$:²²

$$\frac{dS}{dK} = \frac{B}{\Pi} (K_{OT} - K) > 0 \quad \text{for} \quad K_{OT} > K \quad (27.1)$$

and

$$\frac{d^2S}{dK^2} = -\frac{B}{\Pi} < 0, \quad (27.2)$$

where $B = (\alpha_1 + \beta_1)^2 (\alpha_2 + \beta_2)^2 [N_1(\alpha_1 + \gamma_1) + N_2(\alpha_2 + \gamma_2)] > 0$, and $\Pi = t^c(1-t^c)(N_2 - N_1)^2 [(\alpha_2 + \beta_2)N_1\alpha_1W_1 + N_2\alpha_2W_2(\alpha_1 + \beta_1)]^2 > 0$.

The trade-off is illustrated in Figure 3, where coercion-constrained welfare is shown on the vertical axis and the given degree of aggregate coercion is on the horizontal. The Appendix shows that the trade-off is the *same* if the individual-as-dictator counterfactual is used to define the coercion. Of course different patterns of numbers of individuals and of tastes will result in a different trade-off curve. But we stress that it is in principle possible to derive the trade-off using the methodology outlined above.

[Figure 3 here]

The figure records that coercion-constrained social welfare is increasing in K for $K < K_{OT}$. Intuitively, starting from low levels of coercion, the higher the coercion allowed, the easier it is for the planner to implement traditional social planning. This upward sloping part of the locus defines what we may call the 'consenting society'. Points in this region inside the curve represent combinations of coercion and welfare that are not Pareto-efficient. Coercion-constrained welfare reaches a maximum where $K = K_{OT}$, where the solutions for constrained and unconstrained planning are equal. Finally, coercion-constrained welfare is decreasing for values $K > K_{OT}$. For levels of coercion higher than K_{OT} the income tax rate falls below its Optimal Tax size, as seen from (24.2),

²² **Proof of equations (27):** Differentiation of $S(K)$ yields $\frac{dS}{dK} = \frac{-(N_1\alpha_1 + N_2\alpha_2)}{1-t} \frac{dt^c}{dK} + \frac{N_1\gamma_1\alpha_1 + N_2\gamma_2}{G} \frac{dG^c}{dK}$

Substituting from (24.1) and (24.2) into t^c and G^c and noting that upon differentiation

$$\frac{dt^c}{dK} = \frac{-(\alpha_1 + \beta_1)(\alpha_2 + \beta_2)}{[(\alpha_2 + \beta_2)N_1\alpha_1W_1 + (\alpha_1 + \beta_1)N_2\alpha_2W_2](N_1 - N_2)} \quad \text{and} \quad \frac{dG^c}{dK} = \frac{-1}{(N_1 - N_2)P},$$

$$\frac{dS}{dK} = \frac{(N_1\alpha_1 + N_2\alpha_2)(\alpha_1 + \beta_1)(\alpha_2 + \beta_2)}{(1-t^c)[(\alpha_2 + \beta_2)N_1\alpha_1W_1 + (\alpha_1 + \beta_1)N_2\alpha_2W_2](N_2 - N_1)} - \frac{(N_1\gamma_1 + N_2\gamma_2)}{(N_2\gamma_2 - N_1\gamma_1)(N_1W_1 + N_2W_2) - K}.$$

Manipulating and using the definition of K_{OT} in (25), the numerator is written as

$$(\alpha_1 + \beta_1)(\alpha_2 + \beta_2)[N_1(\alpha_1 + \gamma_1) + N_2(\alpha_2 + \gamma_2)](K_{OT} - K).$$

Using the definition of t^c in (24.1) and manipulating, the denominator is written as

$$\frac{t^c(1-t^c)(N_2 - N_1)^2 [(\alpha_2 + \beta_2)N_1\alpha_1W_1 + N_2\alpha_2W_2(\alpha_1 + \beta_1)]^2}{(\alpha_1 + \beta_1)(\alpha_2 + \beta_2)}.$$

Substituting back into dS/dK yields equations (27). QED.

dragging down the level of social welfare. The downward sloping part of the trade off is thus appropriately labelled the 'masochistic society'.

If coercion is of vital concern, it is desirable for society to locate on the upward segment of the curve, and not at the peak of the trade-off. Exactly where depends on the degree of coercion, which is not determined here. The analysis here shows that this unsolved problem is of central importance in public economics when the coercion in the life of the community is acknowledged.²³

But even without knowing K , it is still possible to conduct formal analysis of the structure of public policy. One can delineate the trade-off, and ask what policies look like if they are consistent with attainment of the coercion-welfare frontier, as we have done in the previous section by analyzing a fiscal system with a public good financed by a linear income tax. In the Appendix we reformulate the Ramsey Rule (1927) for the structure of commodity taxation, where it is shown that even when cross-elasticities are ignored, tax rates now depend not merely on own-price elasticities, but also on the distribution of tastes for the taxed goods, since the extent of coercion depends on the difference between what citizens receive and what they pay in taxes.

6. Conclusion

Although coercion is a central fact in the operation of the public sector, normative public economics based on the planning model has not made it an explicit element of the analysis. The coercive nature of collective choice has received much attention in the literature that links economics to political processes. Yet, work in this tradition has focused primarily on the design of public choice mechanisms and does not provide a normative framework for comparing and ranking collective choice procedures in relation to the degree of coercion.

In this paper, we formally introduce coercion into normative analysis by adding constraints that limit allowable coercion caused by tax and expenditure programs. We focus on what we call "collective goods coercion", a problem that arises when citizens experience a mismatch between what they receive in public goods and services and what they pay in taxes. While we make use of a methodology that is similar to that employed in optimal taxation, maximizing social welfare subject to coercion constraints, the focus of the analysis is quite different. Optimal taxation asks, given a consensus on the welfare function, what is the best public decision? Here, in view of the absence of consensus about how the collectivity should conduct its affairs, we ask about the trade-off between social welfare and coercion that must occur as a consequence of collective choice.

Our intent is to explore the grammar and logic of the analysis when coercion, as distinct from redistribution, is formally and explicitly acknowledged. One should note that the traditional planning model does incorporate the analysis of income redistribution, which may have a coercive aspect. However, it places no formal limits on the extent to which any one individual can be coerced when the social plan is implemented. Furthermore, distributional policies are only one possible source of coercion and differ in nature from public goods coercion which stems from the character of public goods and from the lack of a pricing mechanism for such goods as well as from collective choice.

²³ The public choice literature contains possible suggestions about how to approach the choice of K . Although not worked out in a quantifiable manner, the analysis of Buchanan and Tullock (1962, chp. 6) of the optimal decision process may serve as a guide. An alternative approach to determining K may rely on ideas from the contractarian literature.

To make the concept of coercion operational, a counterfactual specifying what individuals regard as appropriate treatment by the public sector is required. This way of proceeding parallels the approach used in defining violations of rights given to citizens in constitutional documents. As is well known, a standard of justice is a prerequisite for determining the extent and nature of any violations of basic rights. We have formulated with some precision several alternative standards of reference: we may formulate the ideal in terms of individual or aggregate utility or, following Breton (1974) and Buchanan (1967), using a convenient approximation that involves a reference level of government expenditures. The aggregate definitions are analogous to the use of the Hicks-Kaldor criterion and impose a less severe constraint on the decision maker than those having an individual basis.

Coercion constraints have significant and complex effects on a social plan. When a linear progressive system of taxation is assumed, tax rates and the level of government expenditures in an extended framework will differ from traditional results. When coercion constraints are defined with reference to individuals, they will become higher or lower than in the traditional analysis, depending on the marginal evaluations of coercion and on the covariance of these with tastes for the public good, among other parameters. When the definition of the counterfactual refers to aggregations of individuals, on the other hand, government expenditures and tax rates will be lower than in an optimal tax plan that fails to account for coercion, assuming only that the value of the Lagrange multiplier on the constraint is positive. Little is known at this point about the relevant empirical magnitudes in both cases.

The analysis also has significant implications for the marginal cost of funds (MCF), another central concept in public finance. In the case based on the aggregate definition of coercion, for example, the MCF must be amended to incorporate the value placed on relaxation of the coercion constraint. If this valuation is positive as expected and an aggregate definition of the coercion constraint is adopted, the adjusted MCF will be higher than the one determined in a framework that fails to account for public goods coercion.

A novel aspect of the analysis relates to the trade-off between social welfare as traditionally defined and coercion. Using a Cobb-Douglas formulation, we derive a trade-off function, as well as the degree of coercion implied by an unconstrained or social plan. The analysis allows us to examine how to achieve the highest level of traditionally defined welfare for a given degree of coercion or, in other words, how to be coercion-efficient. One should note that the term welfare is narrowly defined in this context. If such welfare falls when the coercion constraint is made stricter, we cannot infer that welfare, defined in a broader context, must also fall.

The trade-off function between narrowly defined welfare and public goods coercion is shown to be concave at least in a simplified case, with the maximum representing the point of maximum welfare as traditionally defined. A society (or a planner acting on its behalf) will prefer to be on the upward-sloping part of the relationship, a locus of points corresponding to what we have termed “the consenting society”.

Extensions are possible in several directions. The analysis of welfare-coercion trade-offs can be used to gain insight into how coercion can be reduced through institutional means. The analysis by Usher (1977) suggests that how the boundary between private and public sectors is drawn matters in this regard, and that the coercion-welfare frontier can be shifted favorably by removing certain types of economic activity from the public sphere, thereby increasing the chance of reaching stable political equilibria. The trade-off could be used to formalize this argument and analyze the optimal scope of the public sector. Federalism or decentralization of the public sector is another way of shifting the

trade-off, as suggested in the Introduction. The Tiebout(1956) model certainly is consistent with this sort of reform, and political scientists such as Pennock (1959) have explicitly investigated how federalism and majority rule interact, but the resulting normative economic theory of fiscal federalism that was prompted by Tiebout's paper has not been explicitly concerned with the coercion that precipitates interregional migration. The analysis here suggests how it may be possible to formally integrate coercion into discussion of the optimal assignment of functions in a federation.

Finally, we note that the welfare-coercion frontier also provides the basis for extending the analysis to the study of collective choice, by comparing political equilibria under alternative institutional arrangements in terms of their implied trade-offs between welfare and coercion.

Appendix

1. Demonstration that the social welfare - coercion trade-off derived in section 5 is unaffected by the choice of the counterfactual

Here we adopt the individual-as-dictator counterfactual rather than the individual-in-society perspective. We must first derive what a individual would like if he or she were allowed to decide for everyone the levels of t and G which maximize their own utility function.

We maximize the utility function $U_j = \alpha_j \log X_j + \beta_j \log L_j + \gamma_j \log G$ with respect to t and G subject to the government budget constraint $t \sum_j W_j [\alpha_j W_j / (\alpha_j + \beta_j)] = PG$. Solving yields $t_j^D = \frac{\gamma_j}{\alpha_j + \gamma_j}$ and

$$G_j^D = \frac{\gamma_j}{\alpha_j + \gamma_j} \frac{1}{P} \sum_j \frac{\alpha_j W_j}{1 - \gamma_j} .$$

The corresponding demands for the public good are

$$G_1^D = \frac{\gamma_1}{\alpha_1 + \gamma_1} \frac{1}{P} \left(\frac{N_1 \alpha_1 W_1}{1 - \gamma_1} + \frac{N_2 \alpha_2 W_2}{1 - \gamma_2} \right) \text{ and } G_2^D = \frac{\gamma_2}{\alpha_2 + \gamma_2} \frac{1}{P} \left(\frac{N_1 \alpha_1 W_1}{1 - \gamma_1} + \frac{N_2 \alpha_2 W_2}{1 - \gamma_2} \right).$$

Summing the coercion constraints over the two groups, the aggregate coercion constraint when it just bites is again $N_1(G - G^*_1)P + N_2(G^*_2 - G)P = N_1 K_1 + N_2 K_2$. Setting $K = N_1 K_1 + N_2 K_2$, substituting from above, and manipulating, implies that the coercion-optimal size of G is

$$G^D = \left[\left(\frac{N_2 \gamma_2}{\alpha_2 + \gamma_2} - \frac{N_1 \gamma_1}{\alpha_1 + \gamma_1} \right) \Sigma Y - K \right] \frac{1}{(N_2 - N_1)P} , \text{ where } \Sigma Y = \frac{N_1 \alpha_1 W_1}{1 - \gamma_1} + \frac{N_2 \alpha_2 W_2}{1 - \gamma_2} \text{ denotes total income}$$

Using the government budget restraint then leads to the optimally coercive income tax rate

$$t^D = \frac{1}{N_2 - N_1} \left(\frac{N_2 \gamma_2 (\alpha_1 + \gamma_1) - N_1 \gamma_1 (\alpha_2 + \gamma_2)}{(\alpha_1 + \gamma_1)(\alpha_2 + \gamma_2)} - \frac{K}{\Sigma Y} \right).$$

As in the text, setting $t^D = t^O$, the optimal solution for t under unconstrained social planning, and solving for K , then yields the level of coercion for which the coercion-constrained, welfare maximizing planner would levy the same tax rate as the unconstrained planner:

$$K_D = \frac{N_2 N_1 (\alpha_1 + \gamma_1 + \alpha_2 + \gamma_2) (\alpha_1 \gamma_2 - \alpha_2 \gamma_1)}{(\alpha_1 + \gamma_1)(\alpha_2 + \gamma_2) [N_1 (\alpha_1 + \gamma_1) + N_2 (\alpha_2 + \gamma_2)]} \Sigma Y .$$

To derive the trade-off following the text, we substitute G^D and t^D into the welfare function consisting of the sum of individual utilities. Differentiating this expression with respect to K shows that the trade-off between S and K is concave with an inflection point at $K = K_D$:

$$\frac{dS}{dK} = \frac{B'}{\Pi'} (K_D - K) > 0 \quad \text{for } K_D > K \quad \text{and} \quad \frac{d^2 S}{dK^2} = -\frac{B'}{\Pi'} < 0 ,$$

where $B' = (\alpha_1 + \gamma_1)^2 (\alpha_2 + \gamma_2)^2 [N_1 (\alpha_1 + \gamma_1) + N_2 (\alpha_2 + \gamma_2)] > 0$ and $\Pi' = t^D (1 - t^D) (N_2 - N_1)^2 (\alpha_1 + \gamma_1)^2 (\alpha_2 + \gamma_2)^2 (\Sigma Y)^2 > 0$.

This again leads to equations (27) in the text.

2. Optimally coercive commodity taxation

A longstanding problem concerns the relationship between elasticities of demand and the structure of commodity taxation required to minimize the aggregate excess burden of taxation. We re-examine this problem, first solved by Ramsey (1927), when coercion matters.

To keep the analysis simple, we assume there are two homogeneous groups $j=1, 2$ consuming two commodities $i = 1, 2$, whose prices are P_1 and P_2 . Individual consumption is denoted by X_{ji} . Each taxpayer pays ad-valorem commodity taxes t_1 and t_2 on each of two consumption goods. The tax revenue collected is returned lump sum to each individual in equal amounts, denoted by $R/2$. Taxpayer 1 is assumed to pay too much tax in comparison to what he receives and vice versa for taxpayer 2: this is a highly simplified version of case 4 in Table 1, where the *desired* size of tax payments is equal to $R/2$.

When coercion constraints bind, taxes must be set so that the overpayment made by taxpayer 1 does not exceed a given sum K_1 and the underpayment made by taxpayer 2 must not fall below a certain level K_2 : thus,

$$t_1 P_1 X_{11} + t_2 P_2 X_{12} - (R/2) = K_1 \quad \text{and} \quad t_1 P_1 X_{21} + t_2 P_2 X_{22} - (R/2) = -K_2 .$$

The planner chooses t_1 and t_2 to minimize the excess burden of commodity taxation after securing revenue R , subject to the coercion constraints. Denoting the (absolute value of the) price elasticity of demand for i by e_i , assuming that it is identical across groups, and ignoring cross-price effects, the total excess burden of the tax is

$$B = (1/2) e_1 t_1^2 P_1 (X_{11} + X_{21}) + (1/2) e_2 t_2^2 P_2 (X_{21} + X_{22}),$$

and the budget constraint of the government is

$$t_1 P_1 (X_{11} + X_{21}) + t_2 P_2 (X_{12} + X_{22}) = R .$$

Using κ_1 and κ_2 to denote the relevant multipliers of the coercion constraints, the solution to the problem of minimizing B subject to the government budget restraint and the coercion constraints is in the usual way seen to be:

$$\frac{t_1}{t_2} = \frac{\mu + \kappa_1 w_{11} + \kappa_2 w_{21}}{\mu + \kappa_1 w_{12} + \kappa_2 w_{22}} \frac{e_2}{e_1}, \quad (\text{A1})$$

where $w_{ji} = X_{ji} / \sum_i X_{ji}$ is group j 's share of the consumption of commodity j .

The standard inverse elasticity formula is obtained as a special case of (16) when $\kappa_1 = \kappa_2 = 0$. Then $t_1 / t_2 = e_2 / e_1$ which implies that the more inelastic good must be taxed more heavily.

However, it is now clear from (A1) that this standard result is no longer valid when coercion is taken into account. The optimally coercive commodity tax rate now depends not only on the inverse of the demand elasticity, but also on the level of coercion tolerated by the taxpayers, and this dependence may in turn

reverse the standard conclusion and require that the more elastic good must be taxed more heavily. For example, assume $e_2 > e_1$ so that the Ramsey formula indicates lower taxation of the elastic good 2, but that the ratio

$$p = (\mu + \kappa_1 w_{11} + \kappa_2 w_{21}) / (\mu + \kappa_1 w_{12} + \kappa_2 w_{22}) \quad (A2)$$

is lower than the elasticity ratio. Then for coercion-constrained efficiency it must be that $t_1 < t_2$ and the more elastic good 2 must now be taxed relatively more highly.

Although the actual size of the p -ratio is an empirical issue, one may expect that the higher the budget shares w_{12} and w_{22} (the more both people consume good 2), the more likely that the p -ratio is lower than the ratio of elasticities, and therefore that commodity 2 should be taxed more heavily, contrary to the Ramsey formula. Thus even when cross-elasticities are ignored, tax rates in the amended analysis now depend not merely on own-price elasticities, but also on the distribution of tastes for the taxed goods, since the extent of coercion depends on the difference between what citizens receive and what they pay in taxes.

The structure of commodity taxation and the nature of social solidarity

In order to study more carefully how the nature of the coercion constraints affects the structure of commodity taxation, we can differentiate the first order conditions for excess burden minimization with respect to t_1 , t_2 , κ_1 and κ_2 . Solving the resulting system of four equations, we find that²⁴

$$\frac{dt_1}{dK_1} = \frac{1}{P_1} \frac{X_{22}}{X_{11}X_{22} - X_{21}X_{21}}, \quad \frac{dt_1}{dK_2} = \frac{1}{P_1} \frac{X_{12}}{X_{11}X_{22} - X_{21}X_{21}}$$

and

$$\frac{dt_2}{dK_1} = \frac{-1}{P_2} \frac{X_{21}}{X_{11}X_{22} - X_{21}X_{21}}, \quad \frac{dt_2}{dK_2} = \frac{-1}{P_2} \frac{X_{11}}{X_{11}X_{22} - X_{21}X_{21}}.$$

The sign of these derivatives depends on the sign of the expression $(X_{11}X_{22} - X_{21}X_{12})$. Let us assume that group 1 is the intensive user of commodity 1. Thus, $X_{11} / X_{21} > X_{12} / X_{22}$, which implies that $X_{11}X_{22} - X_{21}X_{12} > 0$. We then have the following two results:

$$(i) \frac{dt_1}{dK_1} > 0, \frac{dt_1}{dK_2} > 0 \quad \text{and} \quad (ii) \frac{dt_2}{dK_1} < 0, \frac{dt_2}{dK_2} < 0 \quad (A3)$$

We consider only the first set of inequalities in (A3) since the second set are analogous. The derivatives show that when K_1 rises, which means that group 1 is coerced more, the optimal coercive tax rate on the commodity which it uses more intensively (commodity 1) rises too. And when K_2 rises, which means that group 2 is coerced less, the tax rate on the commodity which the other group uses more intensively rises.

²⁴ The first order conditions are: $e_1 t_1 P_1 X_1 = \mu P_1 X_1 + \kappa_1 P_1 X_{11} + \kappa_2 P_2 X_{21}$; $e_2 t_2 P_2 X_2 = \mu P_2 X_2 + \kappa_1 P_1 X_{12} + \kappa_2 P_2 X_{22}$; $t_1 P_1 X_{11} + t_2 P_2 X_{12} - (R/2) = K_1$; $t_1 P_1 X_{21} + t_2 P_2 X_{22} - (R/2) = K_2$. Here except for t_1 , t_2 , κ_1 and κ_2 all parameters are assumed fixed. Differentiating these equations with respect to t_1 , t_2 , κ_1 and κ_2 yields: $e_1 P_1 X_1 dt_1 + 0 \times dt_2 - P_1 X_{11} d\kappa_1 - P_2 X_{21} d\kappa_2 = 0$; $0 \times dt_1 + e_2 P_2 X_2 dt_2 - P_2 X_{12} d\kappa_1 - P_2 X_{22} d\kappa_2 = 0$; $P_1 X_{11} dt_1 + P_2 X_{12} dt_2 + 0 \times d\kappa_1 + 0 \times d\kappa_2 = dK_1$; and, $P_1 X_{21} dt_1 + P_2 X_{22} dt_2 + 0 \times d\kappa_1 + 0 \times d\kappa_2 = -dK_2$.

The intuition for this last result is as follows: recall that group 2 gets more than they pay, and that coercion is defined symmetrically for both groups. Hence, when K_2 rises, group 2 gets on balance even more than before, which requires that additional taxes be efficiently raised from group 1.

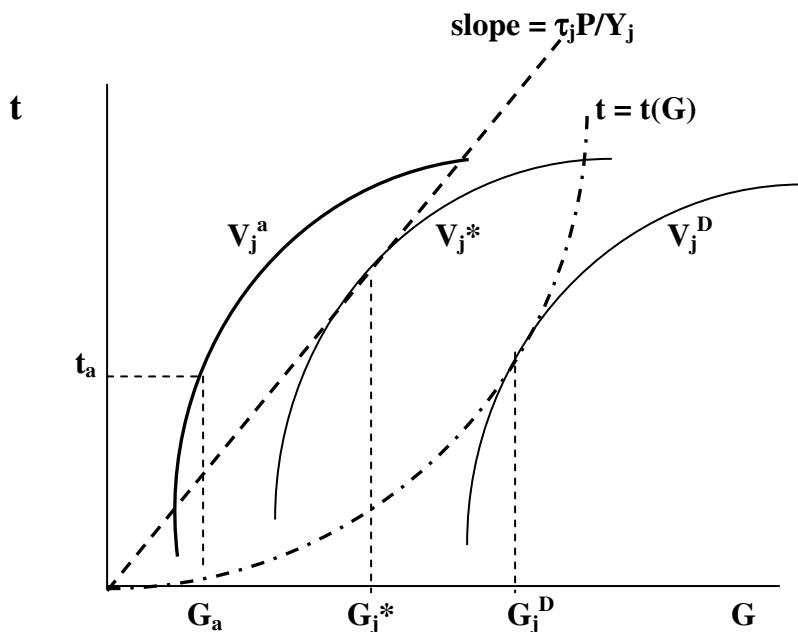
References

- Atkinson, Anthony B. and Nicholas H. Stern (1974). "Pigou, Taxation and Public Goods". *Review of Economic Studies* 41, 119 - 128.
- Atkinson, Anthony, B and Joseph Stiglitz (1980). *Lectures on Public Economics*. McGraw-Hill.
- Baumol (2003), "Welfare Economics and the Theory of the State" , in C. Rowley and F. Schneider (eds.). *The Encyclopaedia of Public Choice*, Kluwer, 610 - 613.
- Breton , Albert (1974). *The Economic Theory of Representative Government*. Chicago: Aldine.
- Breton, Albert. (1996). *Competitive Governments: An Economic Theory of Politics and Public Finance*. Cambridge University Press, Cambridge.
- Buchanan, James, M. and Gordon Tullock (1962). *The calculus of consent. Logical foundations of constitutional democracy*. Ann Arbor, University of Michigan Press.
- Buchanan, James (1968). *The Demand and Supply of Public Goods*. Rand McNally
- Buchanan, James and Roger Congleton (1998). *Politics by Principle, Not Interest*. Cambridge University Press.
- Clower, Robert (1967). "A Reconsideration of the Microfoundations of Monetary Theory", *Western Economic Journal*
- Dalton, Thomas (1977). "Citizen Ignorance and Political Activity". *Public Choice*, 85 - 99.
- Escarrez, D. R. (1967) ' Wicksell and Lindahl: Theories of Public Expenditure and Tax Justice Reconsidered' *National Tax Journal* 20, 137-148
- Hettich, Walter and Stanley L. Winer (1985). " Blueprints and Pathways: The Shifting Foundations of Tax Reform". *National Tax Journal*.
- Hettich, Walter and Stanley L. Winer (1999). *Democratic Choice and Taxation: A Theoretical and Empirical Investigation*. Cambridge University Press.
- Kaplow, Louis (2005). "Pareto Principles and Competing Principles". *The New Palgrave Dictionary of Economics and the Law*, 2nd Edition. Forthcoming. (Harvard Law School Discussion paper 519, August 2005).
- Levy, Margaret (1997). *Consent, Dissent, and Patriotism*. Cambridge University Press.

- Lindahl, Eric (1919). "Just taxation, a Positive Solution". in R. Musgrave and A. Peacock (1958). *Classics in the Theory of Public Finance*. Macmillan 1958, 168-176.
- Mirrlees, James A. (1971). "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175-208.
- Patinkin, Don (1965). *Money, Interest and Prices: An Integration of Monetary and Value Theory*. 2nd Edition. Harper and Row.
- Pennock, Roland (1959). "Federal and Unitary Government - Disharmony and Frustration". *Behavioral Science* 4.
- Perroni, Carlo and Kimberley Scharf (2003). "Viable Tax Constitutions". Unpublished, University of Warwick, September.
- Ramsey, Frank P. (1927). "A contribution to the theory of taxation." *Economic Journal* 37, 47-61.
- Sandmo, Agnar (1984). "Some Insights from the New Theory of Public Finance," *Empirica* 2, 111-124.
- Sandmo, Agnar (1983). "Progressive taxation, redistribution and labor supply," *Scandinavian Journal of Economics* 85, 311-323
- Sandmo, Agnar (1998). "Redistribution and the marginal cost of public funds" *Journal of Public Economics* 70, 365-382.
- Simons, Henry C. (1938). *Personal Income Taxation: The Definition of Income as a Problem of Fiscal Policy*. University of Chicago Press.
- Slutsky, Steven M. (1979). 'Equilibrium Under α -Majority Voting'. *Econometrica* 47, 1113-1125.
- Usher, Dan. (1977). "The welfare economics of the socialization of commodities". *Journal of Public Economics* 8, 151-168.
- Usher, Dan (1993). *The Welfare Economics of Markets, Voting and Predation*. University of Michigan Press
- Usher, Dan (1981). *The Economic Prerequisites of Democracy*. Routledge.
- Weldon, Jack (1968?) On federalism showing that there is no reason for it in standard theory (you could just as well have administrative decentralization).
- Wicksell, Knut (1896). "A New Principle of Just Taxation". in R. Musgrave and A. Peacock (1958). *Classics in the Theory of Public Finance*. Macmillan, 72-118.
- Winer, Stanley L. and Hirofumi Shibata editors (2002). *Political Economy and Public Finance: The Role of Political Economy in the Theory and practice of Public Economics*. Edward Elgar.

Wiseman, Jack (1989). *Cost and Choice and Political Economy*. Edward Elgar Publishing.

Figure 1
Definitions of Coercion and the Choice of Counterfactual



Legend:

t_a : the actual tax rate paid

$t = t(G)$: the feasible combinations of the income tax rate and the level of the public good given the tax system

$t = (\tau_j P / Y_j) G_j$: the implied income tax rate tax at which 'the individual in society' is assumed to be able to quantity adjust the level of the public good, given his tax-share τ_j , the price of the public good P and income Y_j

G_a : actual level of the public good

G^D : the level of the public good which maximizes the utility of j when j chooses G and t to maximize his utility 'as if he were a dictator'

V_j^D : the level of utility attained when j consumes G^D ; this is the counterfactual with the individual as dictator

$V_j^D - V_j^a$: coercion, individual as dictator version

G_j^* : the level of a public good that the individual would like the community to provide at his given tax-price

V_j^* : maximum desired utility at the individual's given tax-price if that person could quantity adjust the level of the public good ($=G_j^*$); this is the Wicksell-Lindahl type counterfactual

$V^* - V_a$: coercion, individual in society version.

Notes to Figure 1:

1. The indifference curves in the (t, G) space are constructed from the individual utility function $U = U(X, G)$, where X is the sole private good, and the individual budget constraint $X = Y(1 - t)$. Substituting the latter into the utility function and totally differentiating we obtain the equation for the slope of the indifference curves in (t, G) space $\frac{dt}{dG} = \frac{\frac{\partial U}{\partial G}}{\frac{\partial U}{\partial X}} Y = (MRS_{GX}) \times Y > 0$

Table 1
Definitions of Coercion

	<i>Type of coercion constraint</i> (The Greek word for coercion is 'katanagasmos')		
<i>Coercion defined With respect to:</i>	Counterfactual	<i>Individual (j=1,..N)</i>	<i>Aggregate</i>
Utility	<i>Utility public good level is what the individual wants at the tax-price determined by the social plan</i>	Case 1: $(V_j^* - V_j) \leq K_j$	Case 2: $\sum_j (V_j^* - V_j) \leq K$
Level of public good	<i>Desired level of the public good compared to its supply in the social plan</i>	Case 3: $ P_G(G_j^* - G) \leq K_j$	Case 4: $n_1 P_G \sum_j (G - G_j^*) + n_2 P_G \sum_j (G_j^* - G) \leq K$

Legend:

G_j^* = the level of a public good that the individual would like the community to provide at his given tax price

G = the level of the public service provided

K_j = the 'degree' of coercion for citizen j .

K (unsubscripted) = an aggregate level of coercion

n_i = the number of taxpayer/citizens of a given type i

P_G = marginal cost of the public good (assumed constant)

V_j^* = maximum desired utility at the individual's given tax price if that person could determine the level of the public good.

V_j = actual level of utility

Figure 2
Coercion Measured by the Level of a Public Good, When $G^* < G$

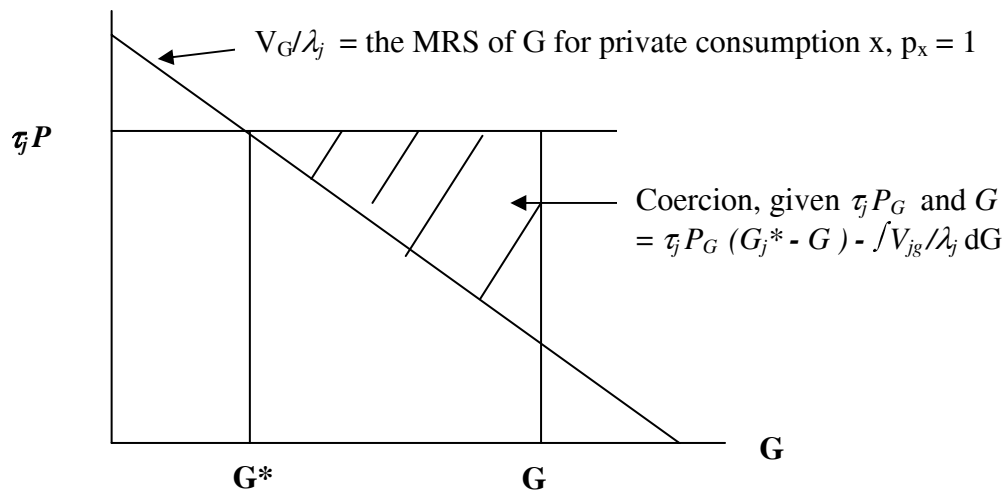
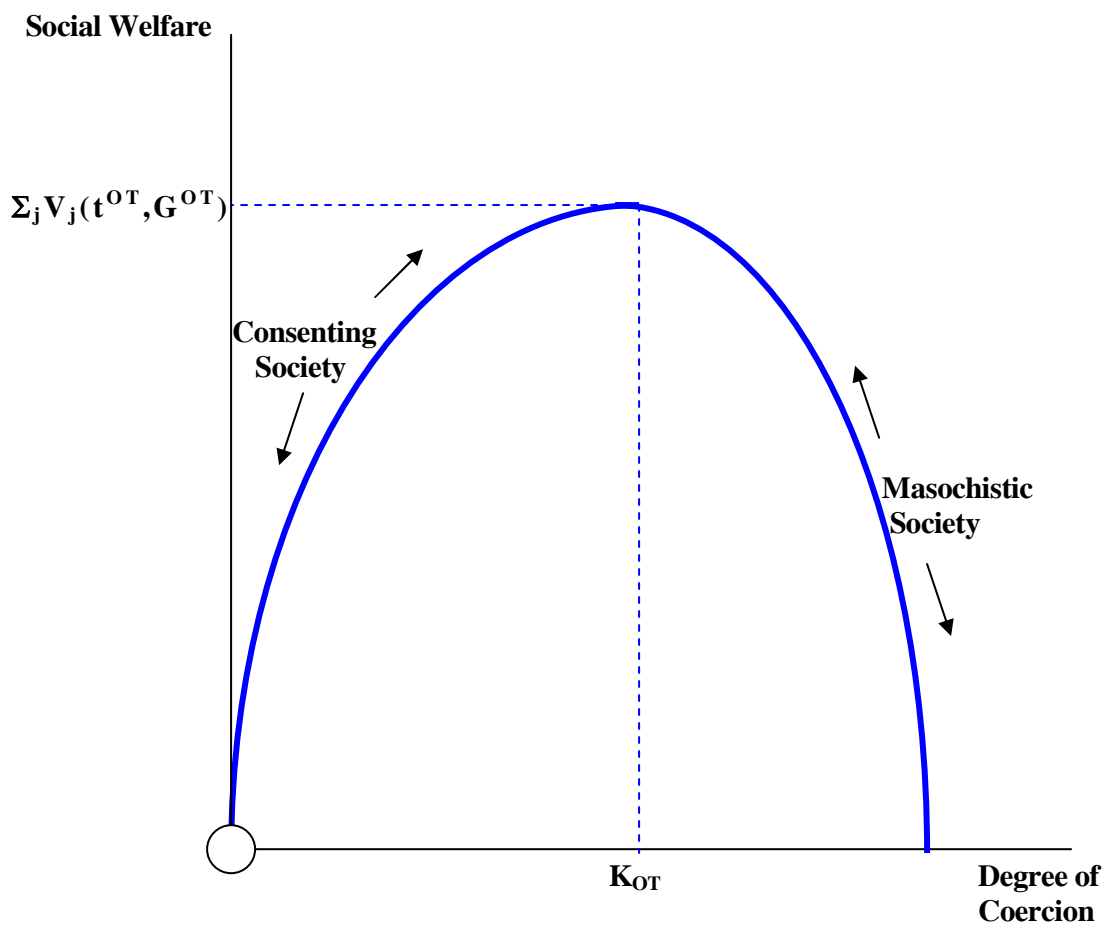


Figure 3
The Coercion - Welfare Frontier



$\Sigma_j V_j(t^{OT}, G^{OT})$: Social welfare (sum of individual utilities) with Optimal Tax solution.

At the origin $K=0$, the tax rate takes the value of zero and the planner's social welfare maximization problem is not defined.