



**Norges  
Handelshøyskole**

*Norwegian School of Economics  
and Business Administration*



**ESSAYS ON TAXATION, EFFICIENCY,  
AND THE ENVIRONMENT**

by

**Lars Håkonsen**

**Disertation submitted to the Department of Economics at The  
Norwegian School of Economics and Business Administration in  
partial fulfilment for the degree of dr.oecon.**

**April, 1998**

36.22  
ille  
ks.

# Acknowledgements

During the time spent working on this thesis, I have benefited greatly from the help, support, and encouragement received from a number of people.

My supervisor, Lars Mathiesen, has always been available and ready to comment upon problems encountered during the various stages that writing a thesis involve. Whether questions have been large or small, concerned mathematics, exposition, or language, he has invariably provided the necessary input for improvements. Furthermore, during the joint work with him on chapter 5, I learned a lot about the many challenges involved in writing a research paper. For all this, I would like to express my gratitude.

The two other members of my supervising committee have been Agnar Sandmo and Tom Rutherford. Agnar Sandmo's writings during the last three decades, and his teaching during my years as a student, have both been important stimuli for my interest in the topics treated in the thesis. Also, his detailed knowledge of the relevant literature and his comments on the various papers have greatly improved the thesis.

Tom Rutherford has through his extensive work on numerical modelling and numerical solvers provided the platform on which chapter 5 is based. Furthermore, his teaching in various modelling courses and the many examples and problems included in these have been highly useful for the work on the other chapters.

I have also benefited greatly from numerous discussions with Bjørn Sandvik. His keen interest in my work and his detailed comments have been highly important for the progress of the thesis.

There is a rich and active research environment on taxation and optimum tax theory at the Department of Economics, which apart from Agnar Sandmo includes Kåre Petter Hagen, Petter Osmundsen, Guttorm Schjelderup, and Fred Schroyen. To all of these I am grateful for numerous interesting and relevant discussions. I would also like to express my gratitude to all my friends and colleagues at the Department, who have made my years here a very stimulating and pleasant experience.

Financial support from the Research Council of Norway and the Norwegian School of Economics and Business Administration is gratefully acknowledged.

98h002757

336.22

Hille

cles. 2

Finally, my warmest gratitude goes to my wife and son, Nina and Magnus, for bringing great pleasure to my life outside the office, and for reminding me that there are other important things in this world than searching for the dead weight loss from distortionary taxation.

Bergen, April 1998

Lars Håkonsen

# Contents

<b>Introduction</b>	<b>1</b>
<b>Chapter 1. On the second order conditions for optimal taxation. Some experience from numerical models</b>	<b>8</b>
<b>Chapter 2. An investigation into alternative representations of the marginal cost of public funds</b>	<b>26</b>
<b>Chapter 3. On green tax reforms and double dividends</b>	<b>45</b>
<b>Chapter 4. Negative externalities, dead weight losses, and the cost of public funds</b>	<b>64</b>
<b>Chapter 5. CO<sub>2</sub>-stabilisation may be a «no-regrets» policy. A general equilibrium analysis of the Norwegian economy.</b>	<b>98</b>

# Introduction\*

## General background

The main themes of this thesis are tax policy analysis and optimal taxation. The development of the theory of optimal taxation during the last decades might be split into two major branches. During the first half of the 1970s, the most of the research focused on models of linear taxation, e.g. indirect commodity taxation and linear income taxes. The fundamental contribution in this tradition is Diamond and Mirrlees (1971). Among other contributions from the 1970s of special importance for this thesis are e.g. Stiglitz and Dasgupta (1971) and Atkinson and Stern (1974), who explore the optimum provision of public goods, and Sandmo (1975), who develops optimal linear tax rules in the presence of negative externalities.

Another branch of the theory of optimal taxation focuses on optimal non-linear income taxes, with fundamental contributions by Mirrlees (1971) and Stiglitz (1982). The issue of optimum provision of public goods in this class of models has been treated by e.g. Boadway and Keen (1993). Of course, there is little reason to restrict the attention to either linear or non-linear taxes. The combined use of non-linear income taxation and linear commodity taxation (the so-called mixed taxation case) has been studied by e.g. Atkinson and Stiglitz<sup>1</sup> (1976), Atkinson (1977), and Christiansen (1984). The consequences of negative externalities in this class of models has recently been analysed by Pirttilä and Tuomala (1997).

As this brief selection of references to some of the contributions to the theory of optimal taxation indicates, the issue of linear taxation had its most active period of research more than 20 years ago. Two more recent issues, however, have greatly revived this branch of literature. The first is the issue of «the marginal cost of public funds» (MCF), which, loosely speaking, attempts to clarify how tax distortions affect the cost of public sector resource use, and which consequences such tax distortions have on the optimum level of public goods provision. The other issue is the so-called «double dividend<sup>2</sup>» from the introduction of environmental taxes in an economy where other distortionary taxes are initially present. Both

---

\* I am grateful to Lars Mathiesen, Jarle Møen, and Fred Schroyen for providing helpful comments and suggestions.

<sup>1</sup> Atkinson and Stiglitz also allow for non-linear commodity tax schedules.

<sup>2</sup> There are several alternative definitions of «double dividends». A generalised version of the double dividend claim is that an increase in tax rates on polluting activities might result in both higher environmental quality and «improvements in other desirable things». Among the «other desirable things» which have been proposed in the literature is i) higher welfare exclusive of environmental quality, ii) higher employment, iii) higher economic growth, and iv) reduced tax distortions.

issues have resulted in a considerable number of published articles during the last few years, most of which have been based on models of linear taxation.

The marginal cost of public funds and the double dividend issues have also received much attention outside academic circles, as both topics are of great relevance to problems high up on the political agenda in most developed countries. It is a common feature for these countries that the size of the public sector has grown considerably during the post-war period. A growing public sector is naturally accompanied by higher tax rates, accentuating the problems of tax distortions and dead weight losses. Due to this, many economists and politicians feel that the marginal tax rates are about to reach their «upper limits». Another topic of political concern in the developed countries is the challenge imposed by environmental problems of various kinds. The idea that environmental taxation might lead to both improved environmental quality and benefits of other kinds, is therefore obviously an attractive one for politicians. In fact, the great political interest in the double dividend issue is perhaps best understood by combining the problems of high initial tax rates and environmental concerns. The opportunity to reap benefits by reducing existing, distortionary tax rates as the environmental taxes are raised to reduce pollution, certainly are more attractive the higher the initial tax rates.<sup>3</sup>

Although most of the literature on double dividends and the marginal cost of public funds has focused on linear taxation, there is of course no particular reason not considering these issues within a non-linear tax framework as well. When this thesis does not consider cases of mixed taxation, it is first of all due to the seemingly rich possibilities of bringing contributions to an active field of research which adopts the linear taxation framework. Another motivation for the use of linear taxation models is that the income tax schedules in most countries are relatively simple, i.e., linear over relatively large income intervals, such that linear taxation becomes a reasonable approximation. In terms of empirical relevance, linear taxation may even become more attractive, since several countries seem to consider flat income tax schedules (i.e., constant marginal tax rate) as potentially interesting future tax systems.

---

<sup>3</sup> From a political economy perspective, it would be tempting for politicians to use environmental concerns as an excuse to impose higher taxes and thereby generate more income which they would benefit from.

## **On the research strategy**

The thesis consists of five separate essays, organised in five chapters. The first four chapters are based on a model framework with a linear production technology in which labour is the only primary factor of production. The last chapter is based on a computable general equilibrium (CGE) model of the Norwegian economy, using relatively detailed National Accounts data, with a particular focus on energy inputs and energy intensive production activities.

In each chapter, the model structure is kept as simple as possible in order to highlight the main issue under investigation. An example may illustrate this point. In chapter 2, the fundamental issue is the distortion of the labour-leisure decision of a household facing either a linear tax on labour income or an indirect consumption tax. In this chapter, therefore, preferences are defined over only two commodities, leisure and a consumption aggregate. Chapters 3 and 4, on the other hand, focus on environmental problems, necessitating a split of the private consumption aggregate into «clean» and «dirty» consumption goods, where the consumption of the latter category results in reduced environmental quality.

In the first four chapters, the theoretical analyses are supplemented with the results from numerical models. The heavy reliance on numerical model examples is open to criticism for only providing special results. Since the models are relatively simple, however, the number of parameter values to be specified is relatively limited. Thus, the numerical examples may uncover important elements of the more general insight by providing sensitivity analyses of a few central parameters, e.g., the elasticities of substitution and the choice of functional form in the utility function. In terms of analysing the effects of policy changes on the consumed quantities of the model variables, numerical analysis is both relatively simple and highly effective. Numerical examples have a central place in e.g. the analyses and comparisons of alternative definitions double dividend definitions in chapter 3. The use of numerical computations is central also in the chapters on the marginal cost of public funds (chapters 2 and 4). In this branch of literature, numerical computations are relatively common, since explicit estimates of MCF for alternative model assumptions and parameter values are central to the discussion.

## **Contents in brief**

Chapter one, on the second order properties of optimal taxation, takes as its starting point the standard textbook model of optimal commodity taxation as described in e.g. Auerbach (1985) and Myles (1995), section 4.3. While first order conditions for optimal taxation have been explored extensively by several authors, little has been said about the second order conditions. Indeed, the only statement which typically is included is that second order conditions are problematic due to the specific nature of the curvature of the maximand and the constraint set. This lack of analyses motivates my approach, which is to solve and illustrate tax optima in small-scale numerical examples. Given the restrictive assumptions of a linear production technology and labour as the only factor of production, a unique tax optimum is found for several alternative specifications of the preference structure of a representative consumer. The figures makes it relatively simple to understand the requirements for an optimal tax model to be well-behaved, and are therefore useful from a pedagogical point of view.

Chapter 2 first provides an overview of several measures of the marginal cost of public funds (MCF) which appear in the literature, and then looks into alternative measures. The chapter attempts to clarify how the various measures relate to the marginal dead weight loss from distortionary taxes. Special attention is given to the invariance properties of the alternative MCF measures with respect to the choice of untaxed commodity and transformations of the utility index. The chapter also attempts to clarify an old and central discussion in this literature, namely whether or not labour income taxation in the special case of Cobb-Douglas preferences should be viewed as distortionary. This discussion is motivated by the fact that the optimality condition for public goods provision in this case appears to be the same as in the first best case, i.e., to equate the sum of the individuals marginal rates of substitution between the public good and the numeraire good to the corresponding marginal rate of transformation.

Chapter 3, on green taxes and double dividends, adopts the model framework of the seminal article in the double dividend literature; Bovenberg and de Mooij (1994). More recent contributions have extended the results of Bovenberg and de Mooij by introducing e.g. intermediate factors of production and labour market imperfections. My approach in this chapter is to clarify and reinterpret the results obtained in the original model. This is motivated by the fact that Bovenberg and de Mooij's results to some extent have been misinterpreted and misunderstood. Two alternative double dividend definitions are examined:

i) a revenue neutral green tax reform leads to higher environmental quality and higher employment, and ii) a revenue neutral green tax reform leads to higher environmental quality and higher welfare exclusive of environmental quality. In the literature on double dividends, the distinction between these two alternatives has not been made sufficiently clear, such that the conditions for version i) to come true have been taken to be relevant also for version ii). The chapter shows that this is a false conclusion, and that whether or not the two alternative double dividends materialise is highly sensitive to the initial tax rates and preference structure.

Chapter 4 is based on the same model framework as Chapter 3, and explores the effects of negative consumption externalities for alternative measures of the marginal cost of public funds. Without externalities in an otherwise perfectly competitive economy, any commodity tax will violate the conditions for Pareto optimality and thereby create a dead weight loss.<sup>4</sup> With a negative externality in the economy, however, this result is turned upside-down. More precisely, it is the no-tax case which violates the conditions for Pareto optimality, while a Pigouvian tax on the source of the negative externality becomes an instrument for bringing about the efficient allocation of resources. While these facts are well established, few attempts have so far been made to clarify how the presence of negative externalities influences MCF. A main conclusion in Chapter 4 is that the presence of negative externalities significantly reduces MCF at low levels of total tax revenues, but that MCF converges towards the value in the no-externality case as total tax revenue increases. In addition to comparing MCF with and without externalities, the chapter also supplements the insight from Chapter 2 regarding the effects of alternative assumptions about preference structures and parameter values on the alternative MCF measures.

The last chapter, Chapter 5, (co-authored with Lars Mathiesen) is closely related to the previous ones in the sense that it emphasises the effects of alternative tax policies on environmental quality and economic efficiency. However, the analysis is no longer based on a stylised theoretical model, but a relatively detailed computable general equilibrium (CGE) model of the Norwegian economy. This work was in part motivated by the fact that several earlier articles on the economic consequences of CO<sub>2</sub> reductions adopted a framework without distortionary tax rates in the benchmark equilibrium. In such a setting, a tax on CO<sub>2</sub> emissions

---

<sup>4</sup> One exception should be mentioned: If relative price changes do not give rise to substitution effects (Leontief preferences), commodity taxes are equivalent in terms of efficiency to lump sum taxes.

necessarily violates the conditions for Pareto optimality and thus by definition reduces the welfare of a representative household.<sup>5</sup> The model in Chapter 5 includes two important elements which potentially could reverse this result. First, since there are distortionary taxes present in the economy to begin with, taxation of CO<sub>2</sub> does not take the economy away from an initially efficient allocation. Rather, increased taxation of CO<sub>2</sub> becomes a tax reform, where one tax rate is increased while one or several other taxes can be reduced accordingly in order to maintain tax revenue neutrality. If the initial tax system is not second best, one cannot exclude the possibility of both CO<sub>2</sub> reductions and efficiency improvements. Second, the model includes several other emissions than CO<sub>2</sub>, e.g. CO, SO<sub>2</sub>, and NO<sub>x</sub>. All of these emissions give rise to negative externalities like e.g. damages to vegetation and health problems. Since these emissions are complements<sup>6</sup> to CO<sub>2</sub> in our model, one will overestimate the cost of CO<sub>2</sub> reduction unless the positive side effects associated with reduced emissions of CO, SO<sub>2</sub> and NO<sub>x</sub> are accounted for. Although the size of such side benefits are highly uncertain, our results indicate that the cost associated with some given goal for CO<sub>2</sub> reduction may be significantly reduced and possibly become negative.

---

<sup>5</sup> This line of reasoning presupposes that positive effects of reduced greenhouse emissions are ignored in the model.

<sup>6</sup> I.e., the emissions of CO, SO<sub>2</sub>, and NO<sub>x</sub> are reduced alongside with CO<sub>2</sub> when introducing a tax on CO<sub>2</sub> emissions.

## References:

- Atkinson, A.B. (1977): Optimal taxation and the direct versus indirect tax controversy, *Canadian Journal of Economics* 10, 590-606.
- Atkinson, A.B. and N.H. Stern (1974): Pigou, taxation and public goods, *Review of Economic Studies* 41, 119-128.
- Atkinson, A.B. and J.E. Stiglitz (1976): The design of tax structure: Direct versus indirect taxation, *Journal of Public Economics* 6, 55-75.
- Auerbach, A.J. (1985): The Theory of Excess Burden and Optimal Taxation, in Auerbach, A.J. and M. Feldstein (eds.): *Handbook of Public Economics*, North-Holland, Amsterdam.
- Boadway, R. and M.J. Keen (1993): Public goods, self-selection and optimal income taxation, *International Economic Review* 34, 463-478.
- Bovenberg, A.L. and R. de Mooij (1994): Environmental levies and distortionary taxation, *American Economic Review* 84, 1085-1089.
- Christiansen, V. (1984): Which commodity taxes should supplement the income tax?, *Journal of Public Economics* 24, 195-220.
- Diamond, P.A. and J.A. Mirrlees (1971): Optimal taxation and public production, *American Economic Review* 61, 8-27 and 261-278.
- Mirrlees, J.A. (1971): An exploration in the theory of optimum income taxation, *Review of Economic Studies* 38, 175-208.
- Myles, G.D. (1995): *Public Economics*, Cambridge University Press.
- Pirttilä, J. and M. Tuomala (1997): Income tax, commodity tax and environmental policy, *International Tax and Public Finance* 4, 379-393.
- Sandmo, A. (1975): Optimal taxation in the presence of externalities, *Swedish Journal of Economics* 77, 86-98.
- Stiglitz, J.E. (1982): Self-selection and Pareto efficient taxation, *Journal of Public Economics* 17, 213-240.
- Stiglitz, J.E. and P.S. Dasgupta (1971): Differential taxation, public goods, and economic efficiency, *Review of Economic Studies* 38, 151-74.

# Chapter 1. On the second-order properties of optimal taxation. Some experience from numerical models\*

## 1. Introduction

Optimal taxation is a well-established area of economic theory, where the main body of research took place in the 1970's. A typical feature of most articles on optimal taxation is a detailed analysis of first-order conditions, while second-order conditions are hardly treated at all. In general, the second order conditions for optimal taxation seems to be regarded as a somewhat problematic topic. This is due to the fact that the maximand – the indirect utility function – is quasi-convex in consumer prices. The following quotation from Dixit (1990), p. 84, summarises the fundamental problem of the second-order properties of optimal taxation:

[The indirect utility function is quasi-convex in P.] «In other words, the lower contour sets of the indirect utility function are convex. This has an unfortunate consequence. When the government chooses indirect taxes optimally, it is in effect choosing prices to maximize an indirect utility function. Our result says that the objective function has the wrong curvature for a maximization problem. Therefore sufficient conditions for optimal taxation are hard to verify.»

Myles (1995) also provides a warning concerning the problems inherent in the structure of optimal taxation in his section 4.3, p.113, «A cautionary note»:

For the Ramsey Rule, the objective function was the household's indirect utility function and hence was quasi-convex..... In addition, the set of taxes that generate at least the required revenue may not be a convex set.

For these reasons the standard sufficiency theorems of quasi-concave programming cannot be appealed to so that there is no guarantee that the first-order conditions actually describe a maximum..... [This problem] is often put to one side and it is simply assumed that the first-order conditions will correctly describe the optima.

In this paper, it is showed that a selection of numerical examples of the standard model of optimal taxation has a unique, well-behaved optimum despite the problem indicated by Dixit and Myles above, viz. the quasi-convexity of the objective function. If concave or quasi-concave programming were applicable for the optimal tax problem, a tangency point found by

---

\* Comments and suggestions from Lars Mathiesen, Petter Osmundsen, Tom Rutherford, Agnar Sandmo and Bjørn Sandvik are gratefully acknowledged.

solving the first-order conditions would be the global maximum. However, concave and quasi-concave programming require *too much* in the sense that although the optimal tax problem violates the conditions for concave or quasi-concave programming, it may still be perfectly well-behaved.

As always, the solution of a constrained maximisation problem is found at a point of tangency between the border of the constraint set and a contour of the maximand. An illustration of such a tangency point has not previously been seen in the optimal taxation literature, but seems to be a nice pedagogical tool for analysing the second order properties of the optimal tax problem.

## 2. Problem statement and solution procedure

The simplest variety of the optimal tax problem is the case where a representative consumer has an exogenous amount of income, and where the producer prices are fixed, see Auerbach (1985), section 5.1. In this paper, I adopt a setting where a representative consumer is equipped with an exogenous time endowment,  $e_0$ , which is allocated optimally between leisure consumption,  $x_0$ , and labour supply,  $(e_0 - x_0)$ . There are  $(n+1)$  private consumption goods (including leisure),  $\mathbf{x} \equiv (x_0, x_1, \dots, x_n)$ , and the consumer's preferences are described by the utility function  $U(\mathbf{x})$ , assumed to be twice continuously differentiable, non-decreasing, and strictly quasi-concave. In addition to private consumption goods there is a good,  $G$ , which is financed and consumed by the public sector. Labour is the only factor of production, and the production possibilities of the economy are described by a fixed coefficient aggregate production technology,

$$-(e_0 - x_0) + \sum_{i=1}^n \alpha_i x_i + \alpha_G G = 0. \quad (1)$$

For the private consumption goods, define the two price vectors  $\mathbf{P}$  and  $\mathbf{p}$ , where  $\mathbf{P} = (P_0, P_1, \dots, P_n)$  is the consumer price vector and  $\mathbf{p} = (p_0, p_1, \dots, p_n)$  is the producer price vector. For the good financed and consumed by the public sector, the producer price is denoted by  $p_G$ . It is assumed that the public sector does not impose taxes on this particular good.

Using leisure (labour) as the untaxed numeraire good, we have that  $P_0 = p_0 = 1$ . Assuming competitive behaviour in the production sectors, (1) implies that  $p_i = \alpha_i$  for  $i = (1, \dots, n)$  and  $p_G = \alpha_G$ , i.e., producer prices are equal to marginal costs measured in labour units.

The representative consumer maximises  $U(x)$  subject to the budget constraint

$$\sum_{i=0}^n P_i x_i(\mathbf{P}) = e_0, \text{ giving the system of Marshallian demand functions}^1 x_i(\mathbf{P}) \forall i \in \{0, \dots, n\},$$

and the corresponding indirect utility function  $U(x(\mathbf{P})) \equiv V(\mathbf{P})$ .  $V(\mathbf{P})$  is continuous, non-increasing, and quasi-convex in  $\mathbf{P}$ .

We now introduce a system of indirect taxes, i.e., a vector of unit taxes  $\mathbf{t} \equiv (t_1, t_2, \dots, t_n)$ . Since the producer prices are  $p_i = \alpha_i$  for  $i = (1, \dots, n)$ , we have that the consumer prices are given by  $P_i = \alpha_i + t_i$  for  $i = (1, \dots, n)$ . With fixed producer prices we may therefore for simplicity express the indirect utility function and the Marshallian demand functions as functions of the vector of tax rates;  $V(\mathbf{t})$  and  $\mathbf{x}(\mathbf{t})$ . Finally, the government's tax revenue is defined as  $R(\mathbf{t}) \equiv \sum_{i=1}^n t_i x_i(\mathbf{t})$ . Given that the Marshallian demand functions  $x_i(\mathbf{t})$  are continuously differentiable,  $R(\mathbf{t})$  is continuously differentiable. The optimal tax problem is the constrained maximisation problem

$$\max_{\mathbf{t}} V(\mathbf{t}) \text{ s.t. } R(\mathbf{t}) \geq \alpha_G G^0,$$

where  $\alpha_G G^0$  is the cost of producing an exogenous amount  $G^0$  of the good financed and consumed by the public sector. The Lagrange function is

$$L(\mathbf{t}, \mu) = V(\mathbf{t}) + \mu[R(\mathbf{t}) - \alpha_G G^0], \quad (2)$$

where  $\mu$  is a Lagrange multiplier and the orientation of the constraint is such that  $\mu$  is positive at the optimum. The first order conditions are

$$\frac{\partial L}{\partial t_k} = -\lambda x_k + \mu \left[ x_k + \sum_{i=1}^n t_i \frac{\partial x_i}{\partial t_k} \right] = 0, \quad \forall k \in \{1, \dots, n\}, \text{ and} \quad (3)$$

---

<sup>1</sup> The labour supply function is the time endowment minus the demand function for leisure, i.e.,  $e_0 - x_0(\mathbf{P})$ .

$$\frac{\partial L}{\partial \mu} = R(\mathbf{t}) - \alpha_G G^0 \geq 0, \quad \mu \geq 0, \quad \mu \frac{\partial L}{\partial \mu} = 0. \quad (4)$$

The Lagrange theorem requires that a constraint qualification must hold. In this particular case, we only have one constraint,  $R(\mathbf{t}) \geq \alpha_G G^0$ . The relevant constraint qualification is then that

$$\frac{\partial R(\mathbf{t})}{\partial t_k} \neq 0 \text{ for at least one } k \in \{1, \dots, n\}. \quad (5)$$

This simply rules out cases where  $R(\mathbf{t})$  has stationary points. We may now state the Lagrange theorem.

### The Lagrange theorem

Suppose  $\mathbf{t}^*$  maximises  $V(\mathbf{t})$  subject to  $R(\mathbf{t}) \geq \alpha_G G^0$ , and the constraint qualification (5) holds. Then there exists a value  $\mu^*$  such that

$$\begin{aligned} \frac{\partial L(\mathbf{t}^*, \mu^*)}{\partial t_k} &= 0, \quad \forall k \in \{1, \dots, n\}, \text{ and} \\ \frac{\partial L(\mathbf{t}^*, \mu^*)}{\partial \mu} &\geq 0, \quad \mu^* \geq 0, \quad \mu^* \frac{\partial L(\mathbf{t}^*, \mu^*)}{\partial \mu} = 0. \end{aligned}$$

In other words, (3) and (4) are necessary conditions for having a constrained maximum for  $V(\mathbf{t})$ , provided that the constraint qualification (5) holds. They are not sufficient, however, since a tax vector  $\mathbf{t}'$  satisfying (3), (4), and the constraint qualification (5) may also be a constrained *minimum* for  $V(\mathbf{t})$ . This is the crux of the matter, since one may not appeal to the theorems of concave or quasi-concave programming for ruling out this possibility in the optimal tax problem<sup>2</sup>.

---

<sup>2</sup> Suppose that both  $V(\mathbf{t})$  and  $R(\mathbf{t})$  are globally concave in  $\mathbf{t}$ . We then have a concave programming problem, with  $L(\mathbf{t}, \mu)$  being the sum of globally concave functions, and thus globally concave in  $\mathbf{t}$ . It follows that a stationary point  $(\mathbf{t}', \mu')$  for the Lagrange function maximises  $L(\mathbf{t}, \mu)$  with respect to  $\mathbf{t}$  given  $\mu'$ . From (4) we then have that  $\mathbf{t}'$  is the global maximum for  $V(\mathbf{t})$  subject to the constraint  $R(\mathbf{t}) \geq G^0$  since (4) ensures that  $L(\mathbf{t}, \mu) = V(\mathbf{t})$  at  $(\mathbf{t}', \mu')$ . Quasi-concave programming cannot be treated this briefly, but the conclusion would still be that a stationary point of the Lagrangean is the global maximum of  $V(\mathbf{t})$ . For further details, see Arrow and Enthoven (1961) or Takayama (1994), Appendix B.

While a theorem ensuring global optimality of a tax vector satisfying (3) and (4) is not available, we may of course still verify that a tax vector satisfying the first-order conditions is a strict local maximum. Suppose that  $(\mathbf{t}', \mu')$  is a solution to (3) and (4), and that the constraint qualification holds. Form the  $(n+1) \times (n+1)$  bordered Hessian of the Lagrange function,

$$H = \begin{bmatrix} 0 & \frac{\partial R}{\partial t_1} & \cdots & \frac{\partial R}{\partial t_n} \\ \frac{\partial R}{\partial t_1} & \frac{\partial^2 L}{\partial t_1^2} & \cdots & \frac{\partial^2 L}{\partial t_n \partial t_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial R}{\partial t_n} & \frac{\partial^2 L}{\partial t_1 \partial t_n} & \cdots & \frac{\partial^2 L}{\partial t_n^2} \end{bmatrix}. \quad (6)$$

### Second order condition

If the sign of the determinant of  $H(\mathbf{t}', \mu')$  is  $(-1)^n$  and the last  $(n-1)$  leading principal minors of  $H(\mathbf{t}', \mu')$  alternate in sign, then  $\mathbf{t}'$  is a strict local constrained maximum of  $V(\mathbf{t})$ <sup>3</sup>.

The second order condition rules out possible minima found by solving the first order conditions, and reduces the number of candidates for being the global constrained maximum. The first order conditions (3) and (4) are only necessary conditions provided that the constraint qualification holds, however. If there exists a tax vector  $\mathbf{t}'$  where  $R(\mathbf{t}') \geq \alpha_G G^0$  but where the constraint qualification is violated, then  $\mathbf{t}'$  is a candidate for being the solution to the maximisation problem. The full procedure for solving the problem therefore is the following (cf. Sydsæter (1990), section 4.13, p. 215).

### Solution procedure

- i) Find all tax vectors satisfying the constraint qualification (5) and the first order conditions (3) and (4).
- ii) Among the points found from i), use the second order condition to rule out local minima.
- iii) Find all points in the constraint set  $\{\mathbf{t} : R(\mathbf{t}) \geq \alpha_G G^0\}$  where the constraint qualification (5) is violated.
- iv) The tax vectors remaining from i) -iii) are all candidates for being the global

<sup>3</sup> See e.g. Simon and Blume (1994), Theorem 19.6.

maximum. The global maximum is the tax vector  $t^*$  among the candidates found from i) - iii) where  $V(\cdot)$  takes the highest value.

These are standard procedures, but it remains to be seen how our problem fits into this general framework. The problem is obviously to be found at stage ii), the second order test. Writing out (6) in full, we find that  $H$  encompasses the Marshallian demand functions, their first- and second-order partial derivatives,  $\lambda$  (the marginal utility of lump-sum income) and its first derivatives, and the multiplier  $\mu$ .

$$H = \begin{bmatrix} 0 & \left( x_1 + \sum_{i=1}^n t_i \frac{\partial x_i}{\partial t_1} \right) & \dots & \left( x_n + \sum_{i=1}^n t_i \frac{\partial x_i}{\partial t_n} \right) \\ \left( x_1 + \sum_{i=1}^n t_i \frac{\partial x_i}{\partial t_1} \right) & \begin{pmatrix} -\lambda \frac{\partial x_1}{\partial t_1} - x_1 \frac{\partial \lambda}{\partial t_1} \\ + \mu \left[ 2 \frac{\partial x_1}{\partial t_1} + \sum_{i=1}^n t_i \frac{\partial^2 x_i}{\partial t_1^2} \right] \end{pmatrix} & \dots & \begin{pmatrix} -\lambda \frac{\partial x_1}{\partial t_n} - x_1 \frac{\partial \lambda}{\partial t_n} \\ + \mu \left[ \frac{\partial x_1}{\partial t_n} + \frac{\partial x_n}{\partial t_1} + \sum_{i=1}^n t_i \frac{\partial^2 x_i}{\partial t_1 \partial t_n} \right] \end{pmatrix} \\ \vdots & \vdots & \ddots & \vdots \\ \left( x_n + \sum_{i=1}^n t_i \frac{\partial x_i}{\partial t_n} \right) & \begin{pmatrix} -\lambda \frac{\partial x_n}{\partial t_1} - x_n \frac{\partial \lambda}{\partial t_1} \\ + \mu \left[ \frac{\partial x_n}{\partial t_1} + \frac{\partial x_1}{\partial t_n} + \sum_{i=1}^n t_i \frac{\partial^2 x_i}{\partial t_n \partial t_1} \right] \end{pmatrix} & \dots & \begin{pmatrix} -\lambda \frac{\partial x_n}{\partial t_n} - x_n \frac{\partial \lambda}{\partial t_n} \\ + \mu \left[ 2 \frac{\partial x_n}{\partial t_n} + \sum_{i=1}^n t_i \frac{\partial^2 x_i}{\partial t_n^2} \right] \end{pmatrix} \end{bmatrix} \quad (6')$$

Ideally we would like to show that the second order condition is fulfilled for a general class of preferences, i.e., without using explicit functional forms. Let us denote the first principal

minor of  $H$  (a 3x3 matrix) by  $D_1$ . By using Young's theorem ( $\frac{\partial^2 L}{\partial t_1 \partial t_2} = \frac{\partial^2 L}{\partial t_2 \partial t_1}$ ), we obtain

$$D_1 = 2 \left( x_1 + \sum_{i=1}^n t_i \frac{\partial x_i}{\partial t_1} \right) \left( x_2 + \sum_{i=1}^n t_i \frac{\partial x_i}{\partial t_2} \right) \left( -\lambda \frac{\partial x_1}{\partial t_2} - x_1 \frac{\partial \lambda}{\partial t_2} + \mu \left[ \frac{\partial x_1}{\partial t_2} + \frac{\partial x_2}{\partial t_1} + \sum_{i=1}^n t_i \frac{\partial^2 x_i}{\partial t_1 \partial t_2} \right] \right) \\ - \left( x_1 + \sum_{i=1}^n t_i \frac{\partial x_i}{\partial t_1} \right)^2 \left( -\lambda \frac{\partial x_2}{\partial t_2} - x_2 \frac{\partial \lambda}{\partial t_2} + \mu \left[ 2 \frac{\partial x_2}{\partial t_2} + \sum_{i=1}^n t_i \frac{\partial^2 x_i}{\partial t_2^2} \right] \right) \\ - \left( x_2 + \sum_{i=1}^n t_i \frac{\partial x_i}{\partial t_2} \right)^2 \left( -\lambda \frac{\partial x_1}{\partial t_1} - x_1 \frac{\partial \lambda}{\partial t_1} + \mu \left[ 2 \frac{\partial x_1}{\partial t_1} + \sum_{i=1}^n t_i \frac{\partial^2 x_i}{\partial t_1^2} \right] \right) \quad (7)$$

From the second order condition, we require that  $D_1 > 0$ . From (7) it seems that the sign of  $D_1$  (and the higher principal minors) cannot be derived unambiguously from general properties of the demand system. Therefore, the only achievement so far is to confirm Dixit's observation

that «.....sufficient conditions for optimal taxation are hard to verify». This seems to be as far as we get without turning to numerical analysis.

In the remaining sections of the paper, we therefore employ numerical examples of the optimal tax problem stated above. The examples suggest that the quotations in section 1 (Dixit (1990) and Myles (1995)) are on the pessimistic side. In all the numerical examples of the optimal tax problem, a unique global maximum is found by solving the first order conditions. Apart from testing the second order properties, the examples also provide useful insight as to how the optimal tax solution changes for alternative specifications of the consumer's preferences.

### 3. Optimal taxation with a one-level CES utility function

In order to illustrate the computed tax optima, we use the model of the previous section, with the following commodities: leisure,  $x_0$ , labour,  $e_0 - x_0$ , two private consumption goods,  $x_1$  and  $x_2$ , and the good  $G$  financed and consumed by the public sector. Throughout this section, the representative consumer's preferences are described by a CES-function<sup>4</sup>,

$$U(x_0, x_1, x_2) = \left( x_0^{\frac{\sigma-1}{\sigma}} + x_1^{\frac{\sigma-1}{\sigma}} + x_2^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \quad (8)$$

where  $\sigma$  denotes the elasticity of substitution between any two goods in  $U(\cdot)$ . For simplicity, we set all the productivity coefficients in the aggregate production technology equal to one, obtaining

$$-(e_0 - x_0) + x_1 + x_2 + G = 0. \quad (1a)$$

Accordingly, all producer prices are 1.0, the tax vector is  $\mathbf{t} = (t_1, t_2)$ , and the vector of consumer prices is  $\mathbf{P} = (P_1, P_2, P_3) = (1.0, 1+t_1, 1+t_2)$ . The indirect utility function  $V(\cdot)$ , the Marshallian demand functions,  $x_i(\cdot)$ , and the tax revenue function  $R(\cdot)$  are thus functions of the tax rates  $\mathbf{t}$ . To complete the numerical specification of the model, we choose a time endowment  $e_0 = 100$ , which with a price of one represents the potential income or full endowment income of the economy. The level of public expenditures is set equal to 25% of full endowment income,  $G^0 = 25$ .

### 3.1. Cobb-Douglas preferences

Consider the solution procedure described in section 2 for the Cobb-Douglas case where  $\sigma = 1.0$ . In step (i), (4) and (5) generate one candidate for global maximum;  $(t_1, t_2) = (0.6, 0.6)$ . In step (ii) we check the local second order condition. The bordered Hessian (6) is in this example a 3x3-matrix, implying that there is only one principal minor to be computed, namely the determinant of the full 3x3 matrix. We find that  $\det(H) = 3227$ , which verifies that  $H$  is negative definite at  $(t_1, t_2) = (0.6, 0.6)$ , i.e., this point is a local constrained maximum for  $V(\cdot)$ , and a candidate for being the global maximum as well. In step (iii), we find that  $R(t_1, t_2)$  has no stationary points. Rather,  $R(t_1, t_2)$  grows asymptotically towards  $200/3$ , i.e.,  $R(t_1, t_2) \geq G^0$  is not a closed and bounded set<sup>5</sup>. Thus, the stationary point of the Lagrange function,  $(t_1^*, t_2^*) = (0.6, 0.6)$  is the global constrained maximum for  $V(\cdot)$ .

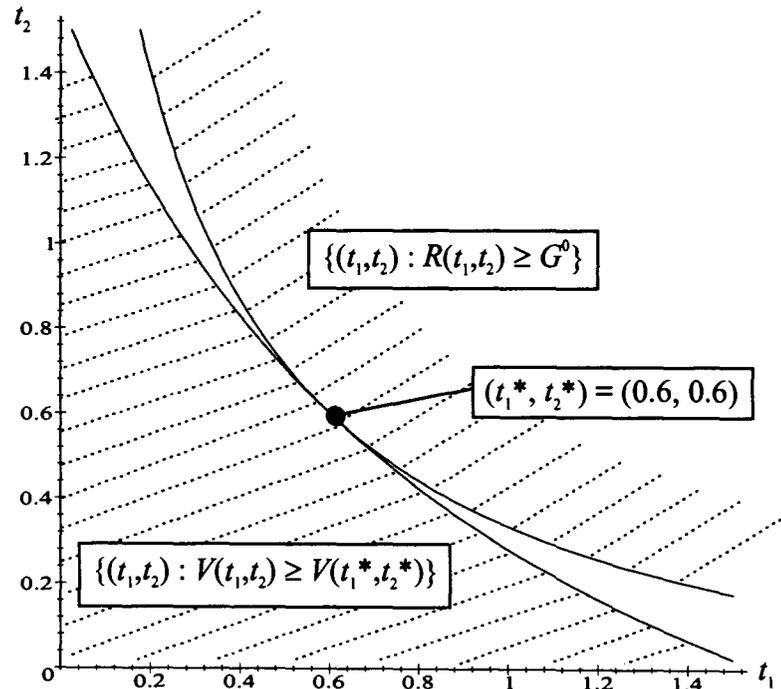


Figure 1. Cobb-Douglas preferences

In Figure 1, we see that the optimum,  $(t_1^*, t_2^*) = (0.6, 0.6)$ , is a point of tangency between a contour<sup>6</sup> of the indirect utility function and the border of the set  $\{(t_1, t_2) : R(t_1, t_2) \geq G^0\}$ . Since

<sup>4</sup> At  $\sigma = 1.0$ ,  $U(\cdot)$  is represented by the Cobb-Douglas function  $U(\cdot) = x_0^{1/3} x_1^{1/3} x_2^{1/3}$ .

<sup>5</sup> In general, Cobb-Douglas preferences  $x_0^a x_1^b x_2^c$  implies that  $\lim_{(t_1, t_2) \rightarrow \infty} R(t_1, t_2) = e_0(b+c)/(a+b+c)$

<sup>6</sup> The contours of  $V(\cdot)$  are indifference curves in the  $(P_1, P_2)$ -space.

the indirect utility function is quasi-convex in  $P$ , the «upper contour set»  $\{(t_1, t_2) : V(t_1, t_2) \geq V(t_1^*, t_2^*)\}$  is *not* a convex set. Thus, there exists no separating hyperplane between the upper contour set of the maximand and the tax income constraint set. In spite of this, we see that the desired separation property of a maximum is fulfilled: The optimum  $(t_1^*, t_2^*) = (0.6, 0.6)$  is the only common point between the upper contour set  $\{(t_1, t_2) : V(t_1, t_2) \geq V(t_1^*, t_2^*)\}$  and the constraint set  $\{(t_1, t_2) : R(t_1, t_2) \geq G^0\}$ . This separation property is precisely the requirement for having a global constrained maximum, see Dixit (1990), section 6.

The conclusion drawn on the basis of this particular example is that the quasi-convexity of the indirect utility function does not by itself represent a problem. If the border of the constraint set is more curved than the dual indifference curve for  $V(\cdot)$ , a tangency point such as the one shown in Figure 1 does represent the global constrained maximum. It is therefore the curvature of the contours of  $V(\cdot)$  *relative to* the curvature of the contours of  $R(\cdot)$  which determines the second order properties of optimal taxation. Evidently, the curvature of the contours of  $V(\cdot)$  and  $R(\cdot)$  cannot change independently of each other, since both  $V(\cdot)$  and  $R(\cdot)$  are derived from the same underlying preferences. What remains to be seen is whether the separation property depicted in Figure 1 is a special feature of this particular preference structure or a more general property.

Before we consider other preference structures, let us check whether the sign of the determinant of  $H$  can be derived for the Cobb-Douglas utility function without assuming specific parameter values. In this special case, the expressions for the indirect utility function and the demand functions are particularly simple;  $V(t_1, t_2) = e_0/(1+t_1)^b(1+t_2)^c$ ,  $x_0 = ae_0$ ,  $x_1 = be_0/(1+t_1)$ , and  $x_2 = ce_0/(1+t_2)$ , where  $a + b + c = 1$ . The determinant of  $H$  then becomes

$$\det(H) = \frac{bc(e_0)^3}{(1+t_1)^{(b+4)}(1+t_2)^{(c+4)}} \begin{pmatrix} -cb(t_2)^2 - b - 2bt_2 - b(t_2)^2 + 2b\mu(1+t_1)^b(1+t_2)^c \\ + 2b\mu t_2(1+t_1)^b(1+t_2)^c + 2bct_1t_2 - bc(t_1)^2 - c - 2ct_1 \\ -c(t_1)^2 + 2c\mu(1+t_1)^b(1+t_2)^c + 2c\mu t_1(1+t_1)^b(1+t_2)^c \end{pmatrix}.$$

Using the fact that  $t_1^*$  equals  $t_2^*$  with Cobb-Douglas preferences<sup>7</sup>, we define  $t = t_1^* = t_2^*$ , whereby  $\det(H)$  reduces into

$$\det(H) = \frac{(e_0)^3 bc(b+c)[2\mu(1+t)^{-a} - 1]}{(1+t)^{(6+b+c)}},$$

---

<sup>7</sup> This follows since  $x_1$  and  $x_2$  have the same degree of complementarity with  $x_0$ , cf. the Corlett-Hague rule.

such that  $\det(H)$  is positive whenever  $\Gamma \equiv (2\mu(1+t)^{-\alpha} - 1)$  is positive. Inserting the general solutions for  $t$  and  $\mu$  into  $\Gamma$ ,  $\mu = (e_0(b+c)/(e_0(b+c)-G)^\alpha$  and  $t = G/(e_0(b+c)-G)$ , we find that  $\Gamma = 1$ , whereby  $\det(H)$  is positive and  $H$  negative definite. The only requirement for this to hold is that  $G < e_0(b+c)$ , where  $e_0(b+c) = \lim_{t \rightarrow \infty} R(t)$ .

### 3.2 Reducing the elasticity of substitution

The curvature of the dual function  $V(P)$  and the curvature of the primal function  $U(x)$  are inversely related. With CES-utility, the higher is  $\sigma$ , the stronger is the curvature of the indifference curves of  $U(x)$ , and the flatter are the dual indifference curves of  $V(P)$ . Choosing  $\sigma = 0.5$ , we obtain Figure 2.

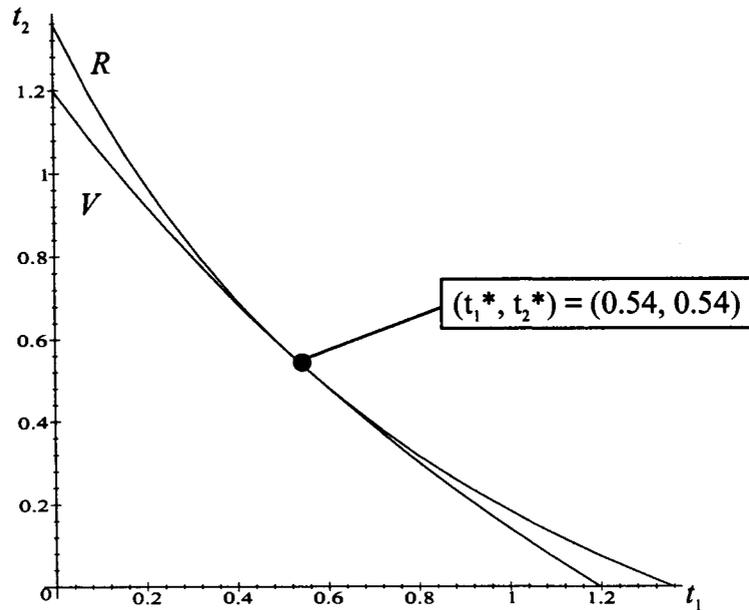


Figure 2. CES-utility with  $\sigma = 0.5$

Again, the solution of the first-order conditions is unique, and represents the global constrained maximum,  $(t_1^*, t_2^*) = (0.54, 0.54)$ . Since  $\sigma < 1.0$  in this case, all goods are necessities. This implies that the government's tax income grows asymptotically towards the full endowment income of  $e_0$ , i.e.,  $\lim_{(t_1, t_2) \rightarrow \infty} R(t_1, t_2) = e_0$ .

Figure 2 also indicates the consequence of approaching the limit case of  $\sigma = 0$ , i.e., Leontief preferences. As  $\sigma$  approaches zero from above, the indifference curves of  $U(x)$  approach  $L$ -shaped curves, and the indifference curves of  $V(\cdot)$  and the contours of  $R(\cdot)$

approach straight lines. Thus, with  $\sigma = 0$ , the borders of the constraint set  $\{(t_1, t_2) : R(t_1, t_2) \geq G^0\}$  and the upper contour set  $\{(t_1, t_2) : V(t_1, t_2) \geq V(t_1^*, t_2^*)\}$  are straight lines with identical slope. Consequently, the tax optimum is not unique with Leontief-preferences; any combination of tax rates such that the required revenue is exactly met give the same utility level.

### 3.3 Increasing the elasticity of substitution

Let us now increase the elasticity of substitution to  $\sigma = 2.0$ . The higher is  $\sigma$ , the higher is the degree of price responsiveness in the demand system. For the optimal tax problem this implies a more elastic tax base, and thus that the maximum available tax revenue is smaller the higher is  $\sigma$ . For  $\sigma > 1.0$ , the tax revenue constraint  $R(t_1, t_2) \geq G^0$  is a closed and bounded set with an interior global maximum. In the case of  $\sigma = 2.0$  we find that  $R(1.73, 1.73) = 26.79$  is the global maximum for  $R(\cdot)$ , implying that the tax revenue requirement  $R(\cdot) \geq 25$  is relatively close to the maximum available tax income. The tax-revenue function in this case – a three-dimensional Laffer surface – is illustrated in figure 3.

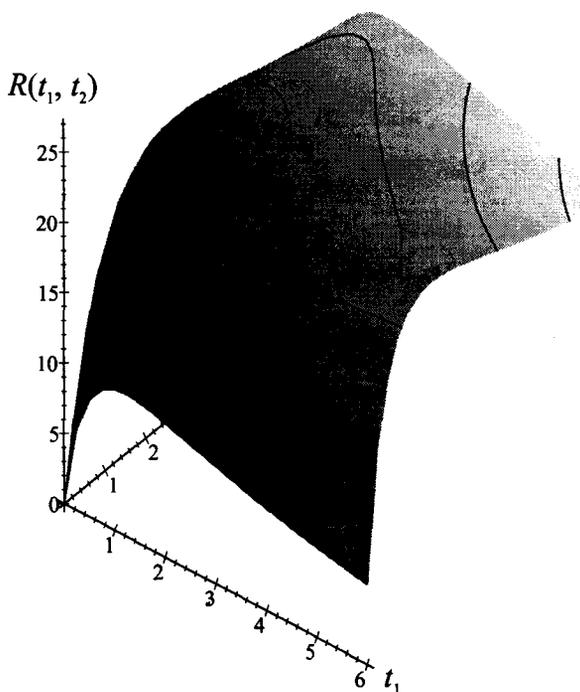
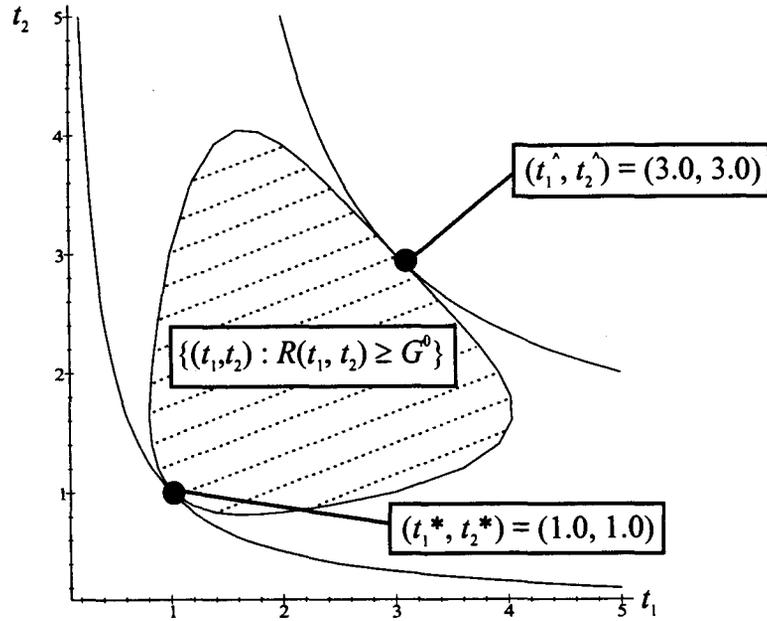


Figure 3.  $R(t_1, t_2)$  with CES-utility and  $\sigma = 2.0$

In this case, there are two points of tangency between contours of  $V(\cdot)$  and the border of the constraint set, see Figure 4.



**Figure 4. Tangency points representing global maximum and global minimum**

The two tangency points are  $(t_1^*, t_2^*) = (1.0, 1.0)$  and  $(t_1^{\hat{}}, t_2^{\hat{}}) = (3.0, 3.0)$ . Only the global maximum,  $(t_1^*, t_2^*) = (1.0, 1.0)$ , satisfies (3) and (4), however. This case demonstrates the significance of the complementary slackness condition (4). Since we require  $\partial L/\partial \mu \geq 0$ ,  $\mu \geq 0$ , and  $\mu(\partial L/\partial \mu) = 0$ , points where  $\mu < 0$  are ruled out. Rewriting (3), we obtain

$$\frac{\partial R}{\partial t_k} = \frac{\lambda}{\mu} x_k \quad \forall k \in [1, 2], \quad (9)$$

which, since  $\lambda$  and  $x_k$  are by definition non-negative, makes it clear that  $\mu < 0$  implies  $\partial R/\partial t_k < 0$ . Thus the non-negativity constraint on  $\mu$  rules out tangency points where more tax revenues could have been collected by decreasing the tax rates, e.g. the point  $(t_1^{\hat{}}, t_2^{\hat{}}) = (3.0, 3.0)$  in Figure 4.

If (4) is stated as a strict equality without a sign requirement for  $\mu$ , both the tangency points shown in Figure 4 will solve the first-order conditions. Going to the next step in the solution procedure in section 2, the point  $(t_1^{\hat{}}, t_2^{\hat{}}) = (3.0, 3.0)$  will be ruled out, however. Inserting the relevant numbers, we have that  $\det(H) = -6.78$  at  $(t_1^{\hat{}}, t_2^{\hat{}}) = (3.0, 3.0)$ ,

implying that  $H$  is positive definite at this point, and thus not a local maximum. Rather,  $(\hat{t}_1, \hat{t}_2) = (3.0, 3.0)$  is the global constrained minimum for  $V(\cdot)$  on the set  $\{(t_1, t_2) : R(t_1, t_2) \geq G^0\}$ .

Increasing  $\sigma$  even more, it turns out that an elasticity of substitution of  $\sigma = 2.12$  is the critical value where the maximum available tax income is exactly  $G^0 = 25$ . In this case,  $\partial R/\partial t_k = 0$  for  $k = 1, 2$  at  $R(t_1, t_2) = 25$ . Step (i) in the solution procedure above therefore generates no candidates, since there exists no solution of the first-order conditions, cf. (9). Therefore, the optimum is the stationary point for  $R(t_1, t_2)$ ,  $(t_1, t_2) = (1.52, 1.52)$ , which violates the constraint qualification, and is the only candidate found from steps (i)-(iii) in the solution procedure in section 2. Of course,  $\sigma > 2.12$  implies that no solution exists, since the set  $\{(t_1, t_2) : R(t_1, t_2) \geq G^0\}$  is empty as long as we stick to the requirement  $G^0 = 25$ .

#### 4. Two-level CES utility

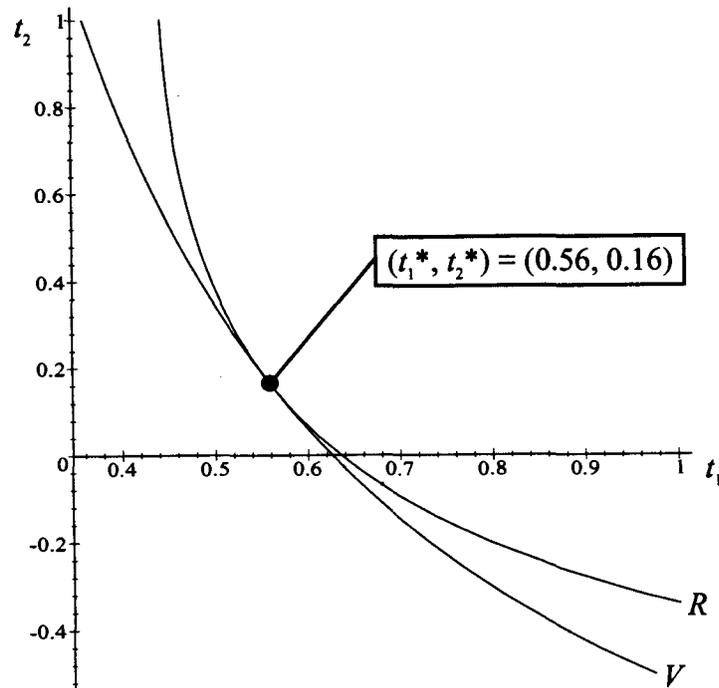
All tax optima studied so far have the property that optimality implies equal tax rates for  $t_1$  and  $t_2$ . With a one-level CES utility function, the compensated cross-elasticities between leisure and the two consumption goods are the same. From the Corlett-Hague rule we then have that  $t_1/P_1 = t_2/P_2$  at the optimum, cf. Corlett and Hague (1953-54) and Sandmo (1987). With a two-level CES-utility function we might choose a preference structure with a different degree of complementarity between leisure and the two consumption goods<sup>8</sup>. As an example, let the preference structure be  $U(x_0, x_1, x_2) = F(x_1, G(x_0, x_2))$ , where both  $F(\cdot)$  and  $G(\cdot)$  are CES functions with elasticities of substitution of  $\sigma^F$  and  $\sigma^G$  respectively. For example, let  $\sigma^F = 0.5$  and  $\sigma^G = 2.0$ . This particular choice implies that commodity 1 is a complement and commodity 2 a substitute for leisure.

In this case the contours of  $R(\cdot)$  and  $V(\cdot)$  are no longer symmetric around a 45<sup>0</sup>-line in the  $(t_1, t_2)$ -space, whereby the optimal tax rates are non-uniform;  $(t_1^*, t_2^*) = (0.56, 0.16)$ , cf. Figure 5.

---

<sup>8</sup> One could also employ a Generalised Leontief expenditure function,  $e(P, u) = u \sum_j \sum_k b_{jk} (P_j P_k)^{1/2}$ ,  $j, k = (0, \dots, n)$

(where each  $b_{jk} \geq 0$ ), and invert it into an indirect utility function. We then have one separate parameter  $b_{jk}$  for each own- and cross price elasticity. Some examples using Generalised Leontief preferences have been computed, but are not commented upon in the paper since the results apparently do not provide any further insight than the included examples.



**Figure 5. Non-symmetric tax optimum**

Although the position of the two contours is quite different compared to the previous examples, there is nothing new with respect to second order properties; the relative curvature of the contours of  $V(\cdot)$  and  $R(\cdot)$  ensures that the tangency point found by solving the first-order conditions is the global constrained maximum also in this case.

So far we have only found optima with strictly positive tax rates for all taxed commodities. Given the preference structure introduced above –  $U(x_0, x_1, x_2) = F(x_1, G(x_0, x_2))$  – and given the choice of  $\sigma^G = 2.0$ , we approach an optimal tax solution where  $t_2$  is zero by reducing  $\sigma^F$  towards zero. Choosing  $\sigma^F = 0.1$ , the tax optimum is  $(0.51, 0.03)$ . In the limit where  $\sigma^F = 0.0$ , the optimal tax scheme is to raise the whole amount of tax income by taxing  $x_1$  alone, i.e.,  $(t_1^*, t_2^*) = (0.5, 0)$ . Again, the second order properties of the optimal solution are not changed however. Note that the optimum  $(t_1^*, t_2^*) = (0.5, 0)$  is not a corner solution. Since we do not require non-negativity of the tax rates, a solution where a tax rate is optimally zero must occur at a point of tangency.

## 5. Optimal subsidies?

In the last paragraph of the above section we commented on a case where it was optimal to leave a taxable good untaxed. Could there also be cases where optimal taxation requires that some of the taxable goods are subsidised? Within the model structure studied so far – what we might think of as «Ramsey»-taxation – one would perhaps not expect that subsidising commodities may be the outcome of a second best optimum. From undergraduate textbooks, we are told that «the dead-weight loss rises with the square of the tax rate». Moreover, a negative commodity tax (a subsidy) is distortionary in its own right. It therefore seems to be a bad idea in terms of efficiency to a) subsidise one or several commodities, and b) increase the tax rates on one or several other commodities in order to finance the net tax revenue requirement of the public sector plus the subsidy payment.

The following example shows that this line of reasoning is not necessarily a valid one. In section 4 we had a case where optimal taxation required that  $t_1 > t_2$  (see Figure 5). Let us maintain all model assumptions from section 4 except for one; the choice of untaxed commodity. With section 4's tax system, the consumer's budget constraint was

$$(1 + t_1)x_1 + (1 + t_2)x_2 = (e_0 - x_0). \quad (10)$$

Dividing through by  $(1+t_1)$ , we obtain an alternative tax system,  $(1 - \tilde{t}_0) \equiv 1 / (1 + t_1)$  and  $(1 + \tilde{t}_2) \equiv (1 + t_2) / (1 + t_1)$ , where labour income and consumption of  $x_2$  are taxed, while  $x_1$  is the untaxed commodity, i.e.,

$$x_1 + (1 + \tilde{t}_2)x_2 = (1 - \tilde{t}_0)(e_0 - x_0). \quad (11)$$

Since  $t_1 > t_2$  in (10),  $\tilde{t}_2$  is negative in (11). In other words, changing the untaxed commodity into  $x_1$  implies that the tax optimum shown in Figure 5 is implemented by a combination of a positive labour income tax,  $\tilde{t}_0 = 0.36$ , and a subsidy on  $x_2$ ,  $\tilde{t}_2 = -0.26$ , cf. Figure 6.

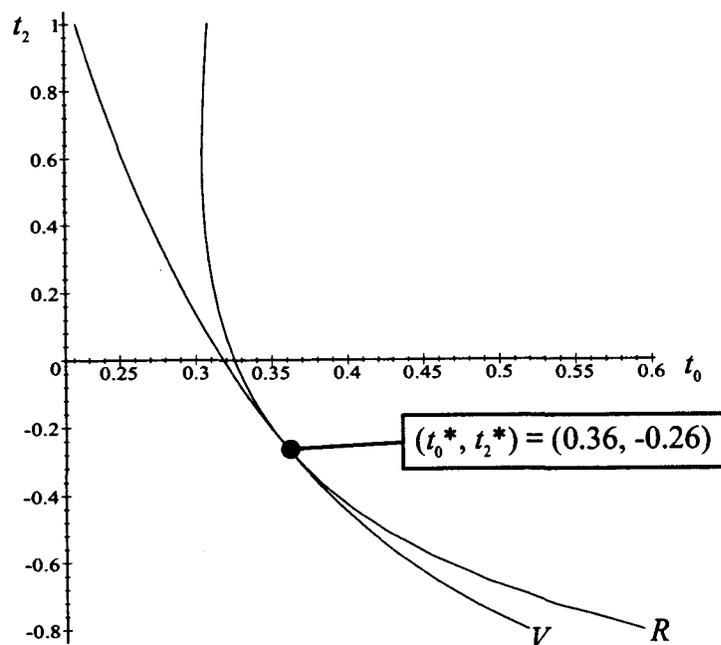


Figure 6. Tax optimum involving a subsidy

The insight provided by this example is of course that optimal taxation determines a specific set of *relative consumer prices*. Therefore, since the optimum requires a higher price of  $x_1$  relative to  $x_2$ ,  $x_2$  must be subsidised in the case where  $x_1$  is chosen as the untaxed commodity. Thus, since all relative prices and the level of public consumption are equivalent, the two alternative tax schemes shown in Figures 5 and 6 sustain the same equilibrium.

## 6. Concluding remarks

Throughout this paper a number of examples have shown that the quasi-convexity of the indirect utility function does not by itself impose a problem with respect to the second order properties of optimal taxation. One might of course ask what is learned about optimal taxation in general by solving a series of numerical examples. Provided that the examples are based on what we might denote as reasonable assumptions regarding the preference structure and corresponding demands, the answer is hopefully that we do learn something. The second order properties of the optimal tax solutions examined in this paper do not seem to be critically dependent on the specific preference structure. Although a number of different assumptions have been examined, the first-order conditions consistently describe the optimum correctly in all the studied examples. The examples therefore indicate that we can do better than «simply assume[d] that the first-order conditions will correctly describe the optima», cf. the quotation of Myles (1995) in the introduction.

All the examples investigated are based on highly stylised assumptions, with one endowment good, two final consumption goods, and a linear production technology. However, apart from making it impossible to illustrate the optima, it is not expected that increasing the number of endowments and final consumption goods would change the fundamental structure of the optimal tax problem. Whether or not more general assumptions regarding the production technology would change the second order properties of optimal tax problems is of course difficult to say without further analysis. This needs to be more closely looked into, and would be an interesting field for future research.

## References:

- Arrow, K.J. and A.C. Enthoven (1961): Quasi-concave programming, *Econometrica* **29**, 779-800.
- Auerbach, A.J. (1985): The theory of excess burden and optimal taxation, in Auerbach, A.J. and M. Feldstein (eds.): *Handbook of Public Economics*, North-Holland, Amsterdam.
- Corlett, W.J. and D.C. Hague (1953-54): Complementarity and the excess burden of taxation, *Review of Economic Studies* **21**, 21-30.
- Dixit, A.K. (1990): *Optimization in Economic Theory*, 2. ed., Oxford University Press.
- Sandmo, A. (1987): A reinterpretation of elasticity formulae in optimum tax theory, *Economica* **54**, 89-96.
- Simon, C.P. and L. Blume (1994): *Mathematics for Economists*, W.W. Norton & Co., New York.
- Sydsæter, K. (1990): *Matematisk Analyse*, Part II, 3. ed., Universitetsforlaget, Oslo.
- Takayama, A. (1994): *Analytical Methods in Economics*, Harvester Wheatsheaf.
- Varian, H.R. (1992): *Microeconomic Analysis*, 3. ed., W.W. Norton & Co., New York.
- Myles, G.D. (1995): *Public Economics*, Cambridge University Press.

# Chapter 2. An investigation into alternative representations of the marginal cost of public funds<sup># \*</sup>

## 1. Introduction

How much does it cost to raise an extra dollar of tax revenue? This is the fundamental question asked in the literature on the marginal cost of public funds (MCF). In a pure lump-sum tax system, tax revenue is transferred from the private to the public sector on a 1 : 1 basis. In other words, MCF is one and the marginal dead weight loss (or, synonymously, the marginal excess burden) is zero. Alternatively stated, lump-sum taxation only produces an income effect, such that the income gain for the public sector exactly equals the income loss for the private sector. These facts seem to be completely uncontroversial.

Unless there are externalities or demands are completely inelastic, any tax system in which there are price distortions is less efficient than a pure lump-sum (first best) tax system. Thus, second best taxation creates economic waste or dead weight losses. This implies that a second best tax regime is Pareto dominated by first best taxation. While this is straightforward, there is considerable controversy over the implications of these facts for MCF. It seems intuitively reasonable that the distortions created by second best taxes increase MCF relative to the level in the first best case. It is well known, however, that a commonly used definition of MCF (see e.g. Ballard and Fullerton (1992)) does not always show such a pattern; this MCF measure may indeed be below one with distortionary taxation. On the other hand, there are other definitions of MCF which are always greater or equal to one. This paper investigates the characteristics of several alternative MCF definitions and discusses how the different measures are related.

The rest of the paper is organised as follows: The next two sections of the paper briefly review a number of well known MCF measures from the existing literature. The fact that none of these measures are invariant to the choice of untaxed commodity motivates the

---

<sup>#</sup> Forthcoming in *International Tax and Public Finance*, 1998.

<sup>\*</sup> Earlier drafts of this paper have been presented at the Nordic Workshop on Tax Policy in Open Economies in Helsinki, (June 96), at seminars at the Universities of Bergen (Feb. 97) and Oslo (Apr. 97), at the Research Forum on Taxation in Oslo (June 97), and at the IIPF Congress in Kyoto (Aug. 97). For valuable comments and suggestions I am grateful to Kjetil Bjorvatn, Robin Boadway, Vidar Christiansen, Bev Dahlby, Kåre P. Hagen, Erling Holmøy, Toshihiro Ihuri, Agnar Sandmo, and two anonymous referees. Special thanks to Bjørn Sandvik, who made an observation which turned out to be very significant for the progress of the paper, and to Lars Mathiesen, whose detailed comments have greatly improved the paper. The usual disclaimer applies.

introduction of an alternative MCF measure in Section 4. In Section 5 it is shown how this measure is related to the dead weight loss measure defined by Kay (1980), Pazner and Sadka (1980), and Triest (1990). Section 6 briefly discusses some of the alternative MCF concepts in relation to the optimal provision of public goods, while Section 7 concludes the paper.

## 2. Commonly used MCF definitions in the existing literature

Ballard (1990) draws a distinction between two categories of MCF-calculations; i) differential analysis, and ii) balanced budget analysis. In a differential analysis, one compares alternative means of financing the same amount of government expenditure. Such an analysis typically has the form of increasing one particular tax rate and reducing another. This paper will focus on *balanced budget* tax increases, where the government's expenditure level is raised by a marginal unit, and the tax rates are changed in order to maintain budget balance. As a general assumption, the level of public goods is not an argument in the household's utility function, and does therefore not affect the tax base. (See section 6, however, where preferences for public goods are briefly introduced.) By this assumption, we only investigate the efficiency effects of *financing* the public expenditures, while the effects of public *spending* are of no concern. For an elaboration of the combined effects of both financing and spending the tax revenues, see e.g. Wildasin (1984), Mayshar (1991), Schöb (1994), and Snow and Warren (1996).

### 2.1 Measures derived from shadow prices

Consider an economy with a representative consumer<sup>1</sup> whose preferences are defined by the utility function  $U(H, C)$ . In  $U(\cdot)$ ,  $H$  denotes hours of leisure and  $C$  the consumption of a private consumption good<sup>2</sup>. The consumer is equipped with an endowment of time,  $E$ , which is optimally allocated between leisure consumption,  $H$ , and labour supply,  $L \equiv E - H$ . Labour is the only production factor in the economy, and there is a linear production constraint (1), where  $G$  denotes a commodity that is financed and consumed by the public sector. Without loss of generality, all productivity coefficients are normalised to one.

---

<sup>1</sup> Sandmo (1997) derives MCF in a setting where the government maximises a welfare function defined over the utilities of  $n$  individuals which differ with respect to their earnings capabilities.

<sup>2</sup>  $C$  might be thought of as an aggregate of  $n$  private consumption goods,  $C \equiv F(C_1, \dots, C_n)$ , where  $F(\cdot)$  is a sub-utility index. A sufficient but not necessary condition for uniform taxes on each consumption good  $C_i$  to be optimal, is that  $H$  and  $F(\cdot)$  are weakly separable and that  $F(\cdot)$  is homothetic, cf. Sandmo (1974). For more on the uniformity issue, see Myles (1995), Section 4.8.

$$L = C + G \quad (1)$$

Let lowercase letters ( $w, p_C, p_G$ ) denote the producer prices for labour, and the private and public consumption good, respectively. With competitive pricing and no taxes, (1) implies that  $w = p_C = p_G$ . Using the producer price of labour as the numeraire,  $w = 1$ , we have a price system where all producer prices equal one. Further, let  $P = 1+t_C$  denote the consumer price for the private consumption good,  $W = 1-t_L$  the net of tax wage rate on labour, and  $a$  the lump-sum tax. Then, the budget constraint for the representative consumer is given by

$$(1+t_C)C = (1-t_L)(E-H) - a. \quad (2)$$

The tax revenue function is  $R(t_L, t_C, a) = t_L L + t_C C + a$ , and the public sector's budget constraint is

$$R(t_C, t_L, a) = t_C C + t_L L + a = G^0, \quad (3)$$

where  $G^0$  is the exogenous output level for the commodity financed by the public sector.

The consumer maximises  $U(\cdot)$  subject to (2), producing the indirect utility function  $V(W, P, a)$ . The government's optimal tax problem<sup>3</sup> is to maximise  $V(\cdot)$  subject to (3), which yields the Lagrange function

$$\Lambda = V(W, P, a) + \mu [t_C C + t_L L + a - G^0]. \quad (4)$$

### First-best taxation

If lump-sum taxation is available, the optimal tax policy is  $\{a^* = G^0, t_C^* = 0, t_L^* = 0\}$ , i.e., the tax revenue is raised by the lump sum tax alone. With  $t_L$  and  $t_C$  equal to zero, solving  $\partial \Lambda / \partial a = 0$  yields

$$\left( \frac{\mu}{\lambda} \right)^{FB} = 1, \quad (5)$$

---

<sup>3</sup> Due to the highly simplified model, we do not need the Lagrangean (4) in order to solve the various cases considered. In e.g. the consumption tax case, one could simply set  $a$  and  $t_L$  equal to zero, and solve (3) to find the consumption tax rate which raises the required tax revenue. The Lagrangean is therefore only introduced in order to generate the same shadow-price ratios (e.g.  $\mu/\lambda$ ) as would appear in models with more endogenous variables.

where  $\lambda$  is the marginal utility of income<sup>4</sup>, and the superscript *FB* is shorthand for first-best.

### Labour income tax

In the second best case,  $a$  is restricted to be zero, leaving us with the two distortionary tax instruments  $t_L$  and  $t_C$ . Since there are no pure profits in this model, one of the two tax rates may without loss of generality be set to zero, cf. Munk (1978). Choosing  $C$  as the untaxed commodity,  $t_C = 0$ , we solve  $\partial\Lambda/\partial t_L = 0$ , obtaining

$$\left(\frac{\mu}{\lambda}\right)^{LIT} = \frac{L}{\left(L + t_L \frac{\partial L}{\partial t_L}\right)} = \frac{1}{1 + \varepsilon_L}, \quad (5)$$

where the superscript *LIT* is shorthand for labour income tax, and  $\varepsilon_L$  is the uncompensated elasticity of labour supply with respect to the income tax rate. Several authors associate  $(\mu/\lambda)^{LIT}$  with MCF, e.g., Ballard (1990), Ballard and Fullerton (1992), Bovenberg and van der Ploeg (1994), and Goulder (1995).

There are alternative measures, however. Following Diamond (1975) and Auerbach (1985), we insert the Slutsky equation,  $\partial L/\partial t_L = \partial L^C/\partial t_L - L(\partial L/\partial I)$ , into (5), and define

$$\alpha^{LIT} \equiv \lambda^{LIT} + \mu t_L \frac{\partial L}{\partial I} \quad (6)$$

as the *marginal social utility of income*<sup>5</sup>, obtaining

$$\left(\frac{\mu}{\alpha}\right)^{LIT} = \frac{L}{L + t_L \frac{\partial L^C}{\partial t_L}} = \frac{1}{1 + \eta_L}. \quad (7)$$

In (7),  $L^C$  is the compensated labour supply, and  $\eta_L$  the elasticity of compensated labour supply with respect to  $t_L$ . (Cf. for example Mayshar (1991), eq. (2) and Wildasin (1984) eq. (2).)

<sup>4</sup> Since the consumer's full income,  $I$ , is  $I = (1-t)E - a$ ,  $\partial V/\partial a$  equals  $-\partial V/\partial I \equiv -\lambda$ .

<sup>5</sup> The marginal social utility of income is the private marginal utility of income plus the income effect on the tax base multiplied with the shadow price on the public sector's budget constraint. According to Auerbach (1985) p. 88,  $(\mu-\alpha)$  represents the marginal excess burden of the tax. An alternative definition of the marginal excess burden (dead weight loss) is provided in Sections 4 and 5 in the present paper.

### Indirect consumption tax

Let us alternatively choose labour as the untaxed commodity,  $t_L = 0$ , and finance the public sector's revenue requirement by means of the consumption tax,  $t_C$ .  $\partial \Lambda / \partial t_C = 0$  then gives

$$\left(\frac{\mu}{\lambda}\right)^{CT} = \frac{C}{\left(C + t_C \frac{\partial C}{\partial t_C}\right)} = \frac{1}{1 + \varepsilon_C}, \quad (8)$$

where superscript  $CT$  denotes «consumption tax», and  $\varepsilon_C$  is the elasticity of (uncompensated) demand for the consumption good with respect to  $t_C$ . The analogue of (6) for the consumption tax case is

$$\alpha^{CT} = \lambda^{CT} + \mu t_C \frac{\partial C}{\partial I}. \quad (9)$$

Using (9) and the Slutsky equation,  $\partial C / \partial t_C = \partial C^C / \partial t_C - C(\partial C / \partial I)$  in (8), we obtain

$$\left(\frac{\mu}{\alpha}\right)^{CT} = \frac{C}{C + t_C \frac{\partial C^C}{\partial t_C}} = \frac{1}{1 + \eta_C}, \quad (10)$$

where  $C^C$  denotes the compensated demand for  $C$ , and  $\eta_C$  the compensated demand elasticity. (Cf. Wildasin (1984), eq. (1).)

### 2.2 MCF defined by means of the equivalent variation

Following Ballard (1990), MCF may be defined as

$$\text{MCF} = - \frac{[\text{change in consumer welfare}]}{[\text{change in government revenue}]}. \quad (11)$$

For a consistent comparison with the money-term in the denominator, the «change in consumer welfare» is measured by means of some money-metric utility concept, e.g. equivalent ( $EV$ ) or compensating ( $CV$ ) variation. For marginal changes,  $CV$  and  $EV$  are identical measures and may be used interchangeably (cf. Mayshar (1990) and Fullerton (1991)). Choosing  $EV$ , and using the notation  $EV^{CP}$  for the case where  $EV$  is computed at the current consumer prices, we obtain the measure

$$\text{MCF} = \frac{-EV^{CP}}{dR}, \quad (12)$$

where  $dR$  is a marginal increase in  $R(\cdot)$ . This is the *MCF*-measure adopted also by Mayshar (1990), section 2.2. It turns out that  $-EV^{CP}/dR$  and  $\mu/\lambda$  is the same measure for marginal increases of optimally financed tax revenues<sup>6</sup>. An alternative definition of  $EV$ , with the producer prices as the reference price vector, is discussed in Section 5.

To show that  $-EV^{CP}/dR$  equals  $\mu/\lambda$ , consider the indirect consumption tax case. Let  $dt_C$  be a marginal change in  $t_C$ , and  $P^0 = 1+t_C^0$  the price level before the marginal tax increase.  $EV^{CP}$  is defined as the amount of lump-sum income such that  $V(W^0, P^0, a^0+EV^{CP}) = V(W^0, P^0+dt_C, a^0)$ . Setting  $(\partial V/\partial I)EV^{CP}$  equal to  $(\partial V/\partial t_C)dt_C$  yields

$$\lambda EV^{CP} = -\lambda C dt_C. \quad (13)$$

Using  $t_L = 0$  and solving  $\partial \Lambda/\partial t_C = 0$  we obtain  $\lambda C = \mu \partial R/\partial t_C = \mu(C + t_C \partial C/\partial t_C)$ . Finally, defining  $dR \equiv (\partial R/\partial t_C)dt_C$ , (13) may be rewritten as

$$\left(\frac{\mu}{\lambda}\right)^{CT} = \left(\frac{-EV^{CP}}{dR}\right)^{CT}. \quad (14)$$

The proofs for  $(\mu/\lambda)^{LIT} = (-EV^{CP}/dR)^{LIT}$  and  $(\mu/\lambda)^{FB} = (-EV^{CP}/dR)^{FB}$  follow by analogy.

### 3. Comparing the alternative measures

So far, four alternative second best measures have been introduced,

- a)  $(\mu/\lambda)^{LIT} = (-EV^{CP}/dR)^{LIT} = 1/(1+\varepsilon_L)$ ,
- b)  $(\mu/\lambda)^{CT} = (-EV^{CP}/dR)^{CT} = 1/(1+\varepsilon_C)$ ,
- c)  $(\mu/\alpha)^{LIT} = 1/(1+\eta_L)$ , and
- d)  $(\mu/\alpha)^{CT} = 1/(1+\eta_C)$ .

Of course, these measures are not identical. For example, if  $U(C,H)$  is a Cobb-Douglas function,  $\varepsilon_L = 0$ , while  $\varepsilon_C < 0$ . The simplest and most illuminating way to compare these measures is probably by use of a numerical specification of the model. To this end, let us employ the CES utility function

---

<sup>6</sup>  $EV^{CP}/dR$  may also be computed for non-optimal (arbitrary) and discrete tax increases, where one for example increases only one tax rate while keeping all other tax rates constant. The distinction between optimal (second

$$U(C, H) = \left( \alpha_C \frac{1}{\sigma} C^{\frac{\sigma-1}{\sigma}} + \alpha_H \frac{1}{\sigma} H^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \quad (15)$$

where  $\sigma$  denotes the elasticity of substitution between  $C$  and  $H$ . The corresponding indirect utility function is given by (16), from which the demand functions can be derived by using Roy's identity.

$$V(W, P, a) = \frac{WE - a}{\left( \alpha_H W^{1-\sigma} + \alpha_C P^{1-\sigma} \right)^{\frac{1}{1-\sigma}}}. \quad (16)$$

Choosing a time endowment  $E = 100$ , share parameters  $\alpha_H = 1$ ,  $\alpha_C = 2$ , and a public sector consumption level  $G^0 = 20$ , completes the numerical specification of the model. Thus, in the absence of taxes, the consumer would spend 2/3 of his time endowment working and allocate the rest of his available time for leisure consumption.

Figure 1 shows the four measures for five different values<sup>7</sup> of the substitution parameter,  $\sigma = \{0.0, 0.5, 1.0, 1.5, 2.0\}$ . In addition, the curve  $(\mu/\lambda)^{FB} = 1.0$  is included as a reference. The fact that  $(\mu/\lambda)^{FB}$  is independent of  $\sigma$  simply follows since the government may collect taxes without changing relative prices in this case.

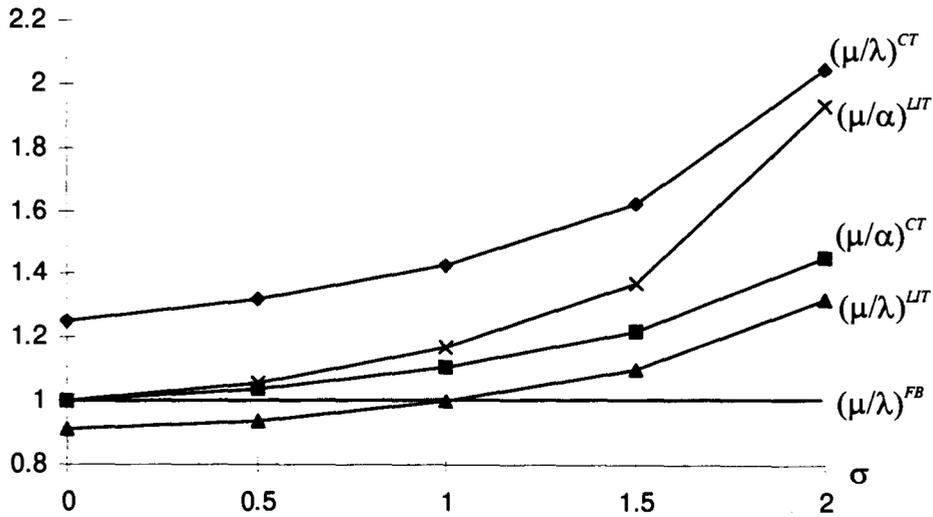


Figure 1.  $(\mu/\lambda)^{LIT}$ ,  $(\mu/\alpha)^{LIT}$ ,  $(\mu/\lambda)^{CT}$ ,  $(\mu/\alpha)^{CT}$ , and  $(\mu/\lambda)^{FB}$  as  $\sigma$  varies from zero to two.

best) and non-optimal (third best) tax increases would only be relevant in richer models than the one used in this paper.

<sup>7</sup>  $\sigma = 1.0$  is represented by the Cobb-Douglas utility function  $U(H, C) = H^{1/3}C^{2/3}$ , while  $\sigma = 0$  is represented by the Leontief utility function  $U(H, C) = \min\{H, 2C\}$ .

No real variables (the consumed quantities and thus the utility level) are affected by the choice of untaxed commodity. From the envelope theorem, the interpretation of the Lagrangean multiplier  $\mu$  is the decrease in the maximand (the utility level) following from a marginal increase in the level of public sector consumption. Since the decrease in utility level by definition is independent of whether the labour income tax or the consumption tax raise the tax revenues, we have that  $\mu^{LIT} = \mu^{CT}$ . Therefore, the fact that  $(\mu/\lambda)^{LIT} \neq (\mu/\lambda)^{CT}$  and  $(\mu/\alpha)^{LIT} \neq (\mu/\alpha)^{CT}$  is due to  $\lambda^{LIT} \neq \lambda^{CT}$  and  $\alpha^{LIT} \neq \alpha^{CT}$  respectively.

To see this, consider first (16) and observe that<sup>8</sup>  $\lambda = \partial V(\cdot)/\partial I = \frac{1}{(\alpha_H W^{1-\sigma} + \alpha_C P^{1-\sigma})^{1/\sigma}} \equiv \frac{1}{e(W, P, 1)}$ , where  $e(W, P, 1)$  is the unit expenditure function. Thus,

the indirect tax case where  $W = 1$  and  $P > 1$ , implies a higher  $e(W, P, 1)$  and a smaller  $\lambda$  than the labour income tax case where  $W < 1$ , and  $P = 1$ .

Turning to  $\alpha^{CT} = \lambda^{CT} + \mu t_C \partial C / \partial I$  and  $\alpha^{LIT} = \lambda^{LIT} + \mu t_L \partial L / \partial I$ , these two differ partly because  $\lambda^{CT} \neq \lambda^{LIT}$ , and partly because  $t_C \partial C / \partial I \neq t_L \partial L / \partial I$ <sup>9</sup>. Provided normality,  $\partial C / \partial I > 0$  while  $\partial L / \partial I < 0$ . Therefore, while  $\lambda^{CT} < \lambda^{LIT}$ , we have in our example that  $\alpha^{CT} > \alpha^{LIT}$ , which explains the result that the relative positions of  $\mu/\lambda$  and  $\mu/\alpha$  are reversed;  $(\mu/\lambda)^{CT} > (\mu/\lambda)^{LIT}$  while  $(\mu/\alpha)^{CT} < (\mu/\alpha)^{LIT}$ .

#### 4. An alternative MCF measure

The MCF definition in (11) is intuitively appealing; it makes clear that the fundamental cost of increased tax revenues is the loss in consumer welfare. According to this measure, two alternative marginal tax increases would yield the same MCF if the change in consumer welfare is the same. None of the four second best MCF measures studied so far has this property. This section presents a shadow price based MCF measure which is uniquely defined in terms of real variables (the economic fundamentals), i.e., invariant to alternative price normalisations.

The discussion in Section 3 made clear that the Lagrangean multiplier  $\mu$  is invariant with respect to the choice of untaxed commodity. Let us denote the common value for  $\mu$  in the

<sup>8</sup> The result that  $\partial V / \partial I = 1/e(P, W, 1)$  only applies in the case where the marginal utility of income is constant or, alternatively stated, for the case where  $U(\cdot)$  is homogeneous of degree one.

<sup>9</sup>  $\partial L / \partial I < 0$  and  $\partial C / \partial I > 0$  provided that  $C$  and  $H$  are normal goods.

second best case by  $\mu^{SB} \equiv \mu^{LIT} = \mu^{CT}$ . The ratio of Lagrangean multipliers in the second and first best cases,  $\mu^{SB}/\mu^{FB}$ , suggests itself naturally in a search for MCF measures which do not respond to alternative price normalisations. Contrary to the previously considered alternatives<sup>10</sup>, however,  $\mu^{FB}/\mu^{SB}$  is only invariant to positive affine transformations of the utility index. Since a cardinal measure is not particularly interesting in an otherwise ordinal context, this measure is discarded. For further details concerning the properties of  $\mu^{SB}/\mu^{SB}$ , see Håkonsen (1997).

By formulating the dual to the traditional formulation of the optimal tax problem,  $\max R(\cdot) \text{ s.t. } V(\cdot) = U$ , it turns out that we obtain a measure which is invariant to both the choice of untaxed commodity and to positive monotone transformations of the utility index. We therefore define the Lagrangean

$$\Gamma = R(t_C, t_L, a) + \beta[V(t_C, t_L, a) - U], \quad (17)$$

where  $\beta$  is the multiplier associated with the constraint  $V(\cdot) = U$ .

In the first best case (where  $t_C = t_L = 0$ ), let  $\{a(U), \beta^{FB}(U)\}$  denote the solution to  $\partial\Gamma/\partial a = 0$ ,  $\partial\Gamma/\partial\beta = 0$ , and let the corresponding maximum value function (the first best maximum tax revenue function) be  $R^{FB}(U) = a(U)$ . From the envelope theorem,  $dR^{FB}/dU = -\beta^{FB}$ .

In the labour income tax case (where  $a = t_C = 0$ ), let  $\{t_L(U), \beta^{LIT}(U)\}$  be the solution to  $\partial\Gamma/\partial t_L = 0$ ,  $\partial\Gamma/\partial\beta = 0$ . Similarly, with the consumption tax (where  $a = t_L = 0$ ), let  $\{t_C(U), \beta^{CT}(U)\}$  be the solution to  $\partial\Gamma/\partial t_C = 0$ ,  $\partial\Gamma/\partial\beta = 0$ . Since the maximum tax revenue by definition is the same in the two alternative second best regimes, we define the second best maximum tax revenue function by  $R^{SB}(U) \equiv t_L(U) \cdot L(t_L(U)) = t_C(U) \cdot C(t_C(U))$ , where  $\partial R^{SB}/\partial U = -\beta^{LIT} = -\beta^{CT} \equiv -\beta^{SB}$ .

The maximum value function for the second best case in the primal formulation,  $\max V(\cdot) \text{ s.t. } R(\cdot) = G$ , is defined by

---

<sup>10</sup> The fact that  $(\mu/\lambda)^{LIT} = 1/(1+\varepsilon_L)$ ,  $(\mu/\lambda)^{CT} = 1/(1+\varepsilon_C)$ ,  $(\mu/\alpha)^{LIT} = 1/(1+\eta_L)$  and  $(\mu/\alpha)^{CT} = 1/(1+\eta_C)$  are invariant to positive monotone transformations of the utility index follows since the elasticities  $\varepsilon_L$ ,  $\varepsilon_C$ ,  $\eta_L$ , and  $\eta_C$  are not affected by such transformations.

$$\begin{aligned}
V^{SB}(G) &\equiv \\
&\max_{t_L} \{V(t_L) \text{ s.t. } R(t_L) = G, t_C = a = 0\} = \\
&\max_{t_C} \{V(t_C) \text{ s.t. } R(t_C) = G, t_L = a = 0\}.
\end{aligned} \tag{18}$$

Defining the argument  $U$  in  $R^{FB}(U)$  and  $R^{SB}(U)$  as  $U \equiv V^{SB}(G)$ , we obtain the composite functions  $\rho^{FB}(G) \equiv R^{FB}(V^{SB}(G))$  and  $\rho^{SB}(G) \equiv R^{SB}(V^{SB}(G)) = G$ , where  $\rho^{SB}(G) = G$  follows by definition of the primal-dual relationship between  $V^{SB}(G)$  and  $\rho^{SB}(U)$ .

Armed with these definitions, it is easy to establish central results regarding the inefficiency of second best taxation relative to the first best case. The fact that first best taxation Pareto dominates second best taxation implies that  $\rho^{FB}(G) \geq \rho^{SB}(G)$ , with strict equality for all cases where there is a non-zero dead weight loss. This is illustrated in Figure 2.

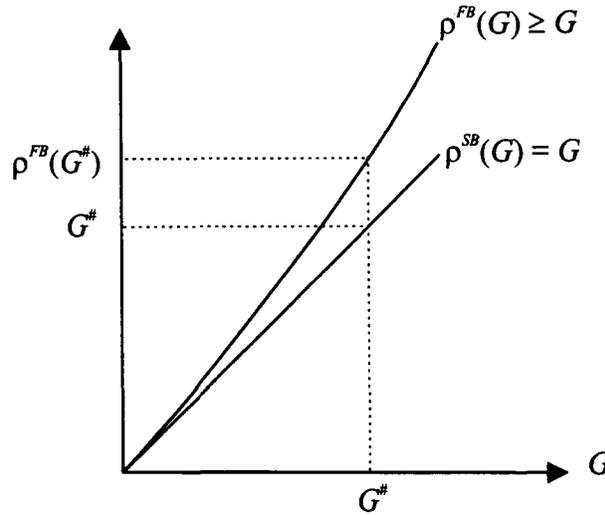


Figure 2.  $\rho^{FB}(G)$  and  $\rho^{SB}(G)$

Suppose now that the amount  $G^\#$  stipulated in Figure 2 is collected by means of distortionary taxation. The maximum tax revenue which could have been collected in the first best case while keeping the utility level constant at  $U = V^{SB}(G^\#)$  is  $\rho^{FB}(G^\#)$ . The vertical distance  $(\rho^{FB}(G^\#) - G^\#)$  represents the Pareto improvement gained by going from second to first best taxation, i.e., the total dead weight loss created by the tax distortions<sup>11</sup>.

<sup>11</sup> Note that  $\rho^{FB}(G) - G$  represents an alternative formulation of the dead weight loss measure in Kay (1980), Pazner and Sadka (1980) and Triest (1990). For further details, see the next Section.

The maximum value functions derived above are related to the cost of public funds by noting that  $\rho^{FB}(G)$  may be interpreted as a measure of the *total* cost of funds. To see this, consider first the case of first best taxation. From the second fundamental theorem of welfare economics, any Pareto optimum may be realised as a competitive equilibrium by transfers of units of the initial endowment,  $E$ . Lump sum taxation does exactly this; it works as if the government was in a position to access the consumer's endowment directly. Therefore, the opportunity cost of  $G$  units of public consumption simply equals the value of the reduction in the endowment, i.e.,  $G$ .

In the second best case, however, a positive level of tax financed public consumption is only possible by violating the conditions for Pareto optimality. The economy is therefore no longer at the efficiency-frontier, and the resulting dead weight loss must be included as an extra cost component in addition to the direct resource cost  $G$ . We therefore define the total cost of funds (TCF), interpreted as an opportunity cost concept, as

$$\text{TCF}(G) = \rho^{FB}(G), \quad (18)$$

i.e., the direct resource cost plus the total dead weight loss. Taking the derivative of TCF yields the corresponding marginal cost of funds measure,

$$\text{MCF} = \text{TCF}' = \frac{d\rho^{FB}}{dG} = \frac{dR^{FB}}{dU} \frac{dV^{SB}}{dG} = \beta^{FB} \mu^{SB} = \frac{\beta^{FB}}{\beta^{SB}}, \quad (19)$$

where the last equality follows from the primal-dual relationship between the multipliers,  $\mu^{SB} = 1/\beta^{SB}$ . Since  $\rho^{FB}(G) - G$  is the total dead weight loss, the marginal dead weight loss becomes  $d\rho^{FB}/dG - 1 = \beta^{FB}/\beta^{SB} - 1$ . The MCF measure in (19) therefore has the interpretation «one plus the marginal dead weight loss»<sup>12</sup>.

Let us now relate  $\beta^{FB}/\beta^{SB}$  in (19) to the measures  $(\mu/\lambda)^{LIT}$  and  $(\mu/\lambda)^{CT}$  in the previous sections. Solving  $\partial\Gamma/\partial t_L = 0$  (where  $t_C = a = 0$ ) yields  $\beta^{LIT} = (1+\epsilon_L)/\lambda^{LIT}$ . Similarly, for the consumption tax case, we have that  $\partial\Gamma/\partial t_C = 0$  (where  $t_L = a = 0$ ) yields  $\beta^{CT} = (1+\epsilon_C)/\lambda^{CT}$ . Finally, in the first best case (where  $t_L = t_C = 0$ ), solving  $\partial\Gamma/\partial a = 0$  yields  $\beta^{FB} = 1/\lambda^{FB}$ . Since  $\beta^{SB} = \beta^{LIT} = \beta^{CT}$ , we have that

---

<sup>12</sup> For a proof that  $\beta^{FB}/\beta^{SB}$  is invariant with respect to positive monotone transformations of the utility index, see Håkonsen (1997).

$$\frac{\beta^{FB}}{\beta^{SB}} = \left(\frac{\mu}{\lambda}\right)^{CT} \frac{\lambda^{CT}}{\lambda^{FB}} = \left(\frac{\mu}{\lambda}\right)^{LIT} \frac{\lambda^{LIT}}{\lambda^{FB}}. \quad (20)$$

Alternatively, cancel the terms  $\lambda^{CT}$  and  $\lambda^{LIT}$  in (20), and use the notation  $\mu^{SB} = \mu^{CT} = \mu^{LIT}$ , to obtain

$$\frac{\beta^{FB}}{\beta^{SB}} = \frac{\mu^{SB}}{\lambda^{FB}}. \quad (20')$$

Thus, the ratio of multipliers in the dual formulation,  $\beta^{FB}/\beta^{SB}$ , equals the multiplier  $\mu^{SB}$  in the primal formulation divided by the marginal utility of income when the consumer faces the first best (producer) prices.

Using the numerical example introduced in Section 3, Figure 3 illustrates how  $\beta^{FB}/\beta^{SB}$  compares to the MCF-measures developed in Section 2.

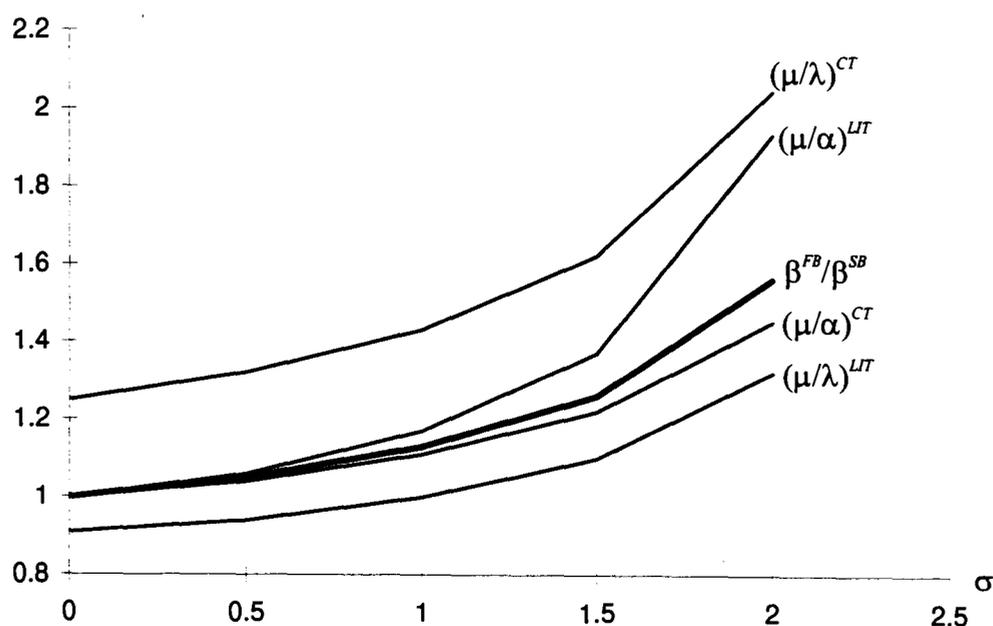


Figure 3.  $\beta^{FB}/\beta^{SB}$ ,  $\mu/\lambda$  and  $\mu/\alpha$

The two maximum value functions  $\rho^{FB}(G)$  and  $\rho^{SB}(G)$  shown in Figure 2 coincide in the case where  $\sigma = 0$ . Thus there is no dead weight loss, totally or at the margin, and  $\beta^{FB}/\beta^{SB}$  equals one. For all strictly positive values of  $\sigma$ , however, second best taxation moves the economy

away from the efficiency frontier, making second best financed public sector consumption more expensive than in the first best case, whereby  $\beta^{FB}/\beta^{SB} > 1$ .

## 5. An alternative derivation of $\rho^{FB}(G)$ and $\beta^{FB}/\beta^{SB}$

Triest (1990) defines a total dead weight loss measure using the equivalent variation evaluated at producer prices. Triest's measure is a reformulation of the dead weight loss definition in Kay (1980) and Pazner and Sadka (1980). To relate the Kay-Pazner-Sadka-Triest measure to the approach used in the present paper, let  $e(W,P,U)$  denote the expenditure function, and let  $U^0$  and  $V^{SB}(G)$  denote the pre- and after-tax utility levels, respectively. The equivalent variation evaluated at producer prices  $(w, p) = (1, 1)$  is defined by<sup>13</sup>

$$EV^{PP} = e(1, 1, V^{SB}(G)) - e(1, 1, U^0). \quad (21)$$

$EV^{PP}$  is the change in lump sum income which for a consumer facing the producer prices would realise the same utility level as in the case of second best taxation,  $V^{SB}(G)$ . By definition, we therefore have that  $-EV^{PP} = \rho^{FB}(G)$ . Taking the derivative of  $-EV^{PP}$ , we obtain<sup>14</sup>

$$\frac{d(-EV^{PP})}{dG} = -\frac{\partial e}{\partial U} \frac{dV^{SB}}{dG} = \frac{\mu^{SB}}{\lambda^{FB}} = \frac{\beta^{FB}}{\beta^{SB}}, \quad (22)$$

where the last step follows from (20').

## 6. Relation to optimal public goods provision

A well known example of the importance of  $\mu/\lambda$  and  $\mu/\alpha$  is the optimality conditions for public goods provision. To briefly address this branch of literature, consider a slight amendment to the model in Section 2, where the utility function  $U(\cdot)$  now includes preferences over the public good  $G$ , i.e.,  $U = U(G, H, C)$ . Assume further that there are  $n$  identical consumers, and that each consumer takes the level of  $G$  as given. Since  $G$  enters into each consumer's utility function,  $G$  is a pure public good. Consider a tax system where tax revenue is financed by means of a labour income tax. All model relations are otherwise retained from

<sup>13</sup> Observe that (21) equals the negative of Triest's eq. (2). This makes  $EV^{PP}$  in (21) a negative number, in line with the definition of  $EV^{CP}$  used in section 2.2 and in Varian (1992), eq. (10.2).

<sup>14</sup> In the step from the second to the third term in (22) we use that  $\partial e(1, 1, V^{SB}(G))/\partial U = 1/[\partial V(1, 1, EV)/\partial \Pi] = 1/\lambda^{FB}$ , cf. Triest (1990), Appendix.

Section 2. We may then express each consumer's indirect utility function by  $V(t_L, G)$ , and formulate the Lagrange function

$$\Omega = nV(t_L, G) + \mu[t_L nL - G] \quad (23)$$

Setting  $\partial\Omega/\partial G$  and  $\partial\Omega/\partial t_L$  equal to zero, dividing the former by the latter, and utilising that  $(\mu/\lambda)^{LIT} = 1/(1+\varepsilon_L)$ , we obtain

$$\frac{n}{\lambda^{LIT}} \frac{\partial V}{\partial G} = \left(\frac{\mu}{\lambda}\right)^{LIT} \left(1 - nt_L \frac{\partial L}{\partial G}\right). \quad (24)$$

Defining  $MRS^{LIT} \equiv (\partial V/\partial G)/\lambda^{LIT} = (\partial U/\partial G)/(\partial U/\partial C)$  as the marginal rate of substitution between the public good and the untaxed good,  $C$ , and  $MRT$  as the marginal rate of transformation between the same two goods, which from (1) equals one, we obtain

$$\sum_n MRS^{LIT} = \left(\frac{\mu}{\lambda}\right)^{LIT} \left(MRT - nt_L \frac{\partial L}{\partial G}\right), \quad (25)$$

which corresponds to Atkinson and Stern (1974), eq. (3), or Wildasin (1984), eq. (7). In the special case of «ordinary independence»,  $\partial L/\partial G = 0$  (cf. Wildasin (1984), eq. 10), (25) simplifies into

$$\sum_n MRS^{LIT} = \left(\frac{\mu}{\lambda}\right)^{LIT} MRT, \quad (26)$$

which is the case focussed on in Ballard and Fullerton (1992)<sup>15</sup>. Ballard and Fullerton give special attention to the case of Cobb-Douglas preferences, where the labour supply curve is vertical,  $\varepsilon_L$  equals zero, and  $(\mu/\lambda)^{LIT}$  equals one. Thus, the optimum provision formula appears to be the same as in the first best case, although the tax system is second best. Ballard and

<sup>15</sup> If the consumption tax replaces the income tax, the equivalent of (26) would be  $\sum MRS^{CT} = (\mu/\lambda)^{CT} MRT$ , where  $MRS^{CT}$  is defined by  $(\partial V/\partial G)/\lambda^{CT} = (\partial U/\partial G)/(\partial U/\partial H)$ . If we alternatively assume that *compensated* demands are independent of the level of  $G$ , the optimality condition corresponding to (26) would change into  $\sum MRS^{LIT} = (\mu/\alpha)^{LIT} MRT$ , see e.g. Wildasin (1984), eq. (9). In other words, all the four measures  $(\mu/\lambda)^{LIT}$ ,  $(\mu/\lambda)^{CT}$ ,  $(\mu/\alpha)^{LIT}$ , and  $(\mu/\alpha)^{CT}$  may be the relevant term in between  $\sum MRS$  and  $MRT$  in the optimality condition for public goods provision, depending on the choice of untaxed commodity and on the way the public good  $G$  enters into the preference structure.

Fullerton discuss this apparent paradox by referring to a survey containing the following two questions (reformulated to our setting):

1. Is a labour income tax  $t_L$  of 50% distortionary in the case of Cobb-Douglas preferences?
2. In the same model, suppose a public project with production costs (MRT) of one, and benefits ( $\sum MRS$ ) of slightly more than one, could be funded by a 1% increase in the wage tax. Would this be desirable?

As pointed out by Ballard and Fullerton, the correct answer to both these questions is «yes». What is only implicit in Ballard and Fullerton’s discussion, however, is a measure of the extra costs created by the labour income tax compared to first best taxation. The measure developed in Section 4 in this paper,  $\beta^{FB}/\beta^{SB}$ , represents exactly this «missing link» in Ballard and Fullerton’s discussion. In their concluding remarks, they allow for some vagueness concerning the answer to the first question:

«However, we recognise that the “yes” answer to the first question is subject to semantic interpretation.» [Ballard and Fullerton (1992), p. 128]

Let us confront this remark with the results of our previously used numerical model. We assume that  $U(C,H)$  is a Cobb-Douglas function, and let  $G$  vary from zero to 50 per cent of the full endowment income. Figures 4 and 5 show the maximum value functions  $\rho^{FB}(G)$  and  $\rho^{SB}(G)$ , and the MCF measures  $\beta^{FB}/\beta^{SB}$  and  $(\mu/\lambda)^{LIT}$ , respectively.

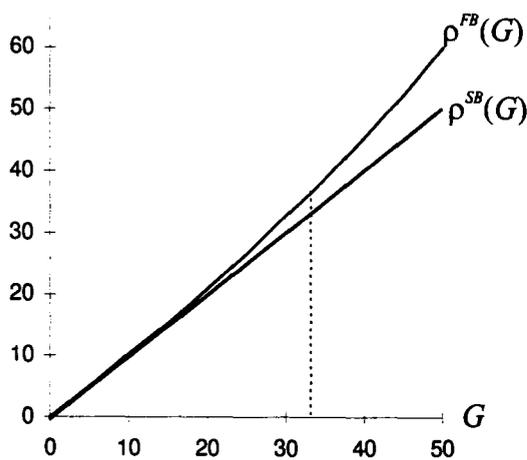


Figure 4.  $\rho^{FB}(G)$  and  $\rho^{SB}(G) = G$  with Cobb-Douglas preferences.

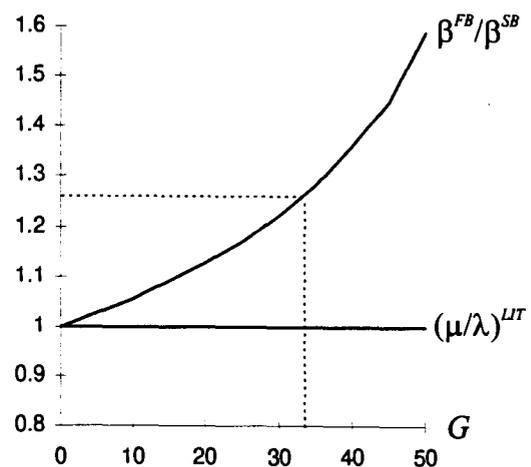


Figure 5.  $(\mu/\lambda)^{LIT}$  and  $\beta^{FB}/\beta^{SB}$  with Cobb-Douglas preferences.

The dotted lines indicate the level of public consumption,  $G = 33.33$ , which corresponds to the tax rate mentioned in question one;  $t_L = 0.5$ . At this level, the total dead weight loss is 11 per cent of the direct resource cost, the marginal dead weight loss is 26 per cent,  $\beta^{FB}/\beta^{SB}$  is 1.26, and  $(\mu/\lambda)^{LIT}$  is one. Based on these two Figures, there is little scope for semantic interpretation regarding whether or not the labour income tax is distortionary. The correct answer to question one must be a perfectly clear «yes». At the same time, the correct answer to question 2 is an equally unambiguous «yes».

The important observation is that there is no inherent conflict between these two answers. The main source of confusion is – presumably – that  $(\mu/\lambda)^{LIT}$  has been interpreted as an indicator of the degree of inefficiency created by the tax distortions. Figure 5 makes clear that it is not. To see why, it is illuminating to write  $(\mu/\lambda)^{LIT} = (\beta^{FB}/\beta^{SB}) \cdot (\lambda^{FB}/\lambda^{LIT})$ , cf. (20). The first term,  $\beta^{FB}/\beta^{SB}$ , is an indicator of the degree of inefficiency relative to the first best case, interpreted as one plus the marginal dead weight loss, and is strictly increasing in the tax rate  $t_L$ . Since the labour income tax reduces the consumer price of leisure, however, the second term is strictly decreasing in  $t_L$ , exactly outweighing the first effect in the special case of Cobb-Douglas preferences.

## 7. Concluding remarks

This paper has studied alternative measures of the marginal cost of public funds. In a range of articles it seems that  $\mu/\lambda = -EV^{CP}/dR$  is recognised as representing MCF. The fact that  $\mu/\lambda$  may be less than one even though distortionary taxation is used has therefore caused considerable confusion. The observation that  $\mu/\lambda$  is less than one when taxing a backward-bending labour supply was shown already by Atkinson and Stern (1974). However, unless the compensated labour supply is completely inelastic, labour income taxation is strictly Pareto dominated by lump sum taxation, implying that the real costs of financing tax revenues with labour income taxation is strictly higher than in the lump sum tax case. Despite this, many authors seem to argue that MCF should be defined as «the term by which MRT is multiplied» in the optimum public goods provision formula. If this is the preferred definition, one should be aware that MCF does not represent an indicator of the inefficiencies created by the tax distortions.

To briefly recollect a central point, the main motivation for the total and marginal cost of funds concepts introduced in Sections 4 and 5 is the following line of reasoning:

Second best taxation creates tax distortions and first best taxation does not. The dead weight loss resulting from the tax distortions in a second best tax system represents an extra cost component which comes in addition to the direct resource use of the public sector. According to this line of reasoning, it is a fundamental fact that second best financed tax revenue is more expensive (costs more) than lump sum financed tax revenue.

Depending on which interpretation of MCF one has in mind, and on which price normalisation one chooses, any of the cost of funds concepts described in this paper may indeed be the relevant one. The problem with such a richness of alternative measures is of course that several different concepts are denoted «MCF». This paper, deriving and comparing several alternative MCF measures hopefully provides a useful taxonomy which might clarify the sense in which the different measures ought to be used and interpreted in future writings on this field.

## References:

- Atkinson, A.B. and N.H. Stern (1974): Pigou, taxation and public goods, *Review of Economic Studies* 41, 119-128.
- Auerbach, A.J. (1985): The theory of excess burden and optimal taxation, chapter 2 in Auerbach, A.J. and M. Feldstein (eds.): *Handbook of Public Economics*, vol. I, North-Holland.
- Ballard, C.L. (1990): Marginal welfare cost calculations. Differential analysis vs. balanced-budget analysis, *Journal of Public Economics* 41, 263-276.
- Ballard, C.L. and D. Fullerton (1992): Distortionary taxation and the provision of public goods, *Journal of Economic Perspectives* 6, 117-131.
- Bovenberg, A.L. and F. van der Ploeg (1994): Environmental policy, public finance and the labour market in a second-best world, *Journal of Public Economics* 55, 349-390.
- Diamond, P.A. (1975): A many-person Ramsey rule, *Journal of Public Economics* 4, 335-342.
- Fullerton, D. (1991): Reconciling recent estimates of the marginal welfare cost of taxation, *American Economic Review* 81, 302-308.
- Goulder, L.H. (1995): Environmental taxation and the double dividend: A reader's guide, *International Tax and Public Finance* 2, 157-183.
- Håkonsen, L. (1997): An investigation into alternative representations of the average and marginal cost of public funds, Discussion Paper 11/97, Norwegian School of Economics and Business Administration, Bergen, Norway.
- Kay, J.A. (1980): The deadweight loss from a tax system, *Journal of Public Economics* 13, 111-119.
- Mayshar, J. (1990): On measures of excess burden and their application, *Journal of Public Economics* 43, 263-289.
- Mayshar, J. (1991): On measuring the marginal cost of funds analytically, *American Economic Review* 81, 1329-1334.
- Munk, K.J. (1977): Optimal taxation and pure profit, *Scandinavian Journal of Economics* 79, 1-19.
- Myles, G.D. (1995): *Public Economics*, Cambridge University Press.
- Pazner, E.A. and E. Sadka (1980): Excess-burden and economic surplus as consistent welfare indicators, *Public Finance* 35, 439-49.
- Sandmo, A. (1974): A note on the structure of optimal taxation, *American Economic Review* 64, 701-706.
- Sandmo, A. (1997): Redistribution and the marginal cost of public funds, Discussion Paper 8/97, Norwegian School of Economics and Business Administration, Bergen, Norway. [Forthcoming in *Journal of Public Economics*]
- Schöb, R. (1994): On marginal cost and marginal benefit of public funds, *Public Finance* 49, 87-106.
- Snow, A. and R. S. Warren Jr. (1996): The marginal welfare cost of public funds: theory and estimates, *Journal of Public Economics* 61, 289-305.
- Triest, R. K. (1990): The relationship between the marginal cost of public funds and the marginal excess burden, *American Economic Review* 80, 557-566.

Varian, H. (1992): *Microeconomic Analysis*, W.W. Norton & Co., 3. edition.

Wildasin, D.E. (1984): On public good provision with distortionary taxation, *Economic Inquiry* 22, 227-243.  
[Errata, *Economic Inquiry* 23, p. 185.]

## Chapter 3. On green tax reforms and double dividends\*

### 1. Introduction

The idea of a «double dividend» (DD) from introducing environmental taxes is intuitively appealing. The intuition goes as follows: Suppose that policy makers have been unaware that consumption and/or production of certain goods lead to a negative externality in terms of reduced environmental quality. Taxing these particular goods is a corrective device for a fundamental failure of the price system – missing markets for environmental quality – and is thus efficiency-improving. On the other hand, taxes on non-externality-producing goods create dead weight losses, and violate the conditions for having a first best efficient allocation. Suppose that the tax revenue requirement of the public sector is unaffected. The introduction of environmental taxes then leads to two effects: i) improved environmental quality, and ii) reductions in existing distortionary tax rates. The former effect is, and has always been, the primary goal of environmental taxes. The latter effect, however, indicates that the presence of distortionary taxes gives an extra benefit (dividend) from environmental taxation.

In the economics literature, the idea of a DD from environmental taxes was sketched already by Terkla (1984) and Lee and Misiolek (1986), and, to my knowledge, Pearce (1991) was the first to use the term «double dividend» explicitly. Pearce does not give a precise definition of a double dividend, however, and provides no analytical support for his suggestion. Bovenberg and de Mooij (1994) use a simple analytical model in order to examine the idea further. Goulder (1995), Parry (1995), Schöb (1996), and Fullerton (1997) refer to and extend Bovenberg and de Mooij's results, and Oates (1995) provides an overview. While the literature on double dividends thus has become quite large, there are still a few loose ends. This is partly because the widely cited paper by Bovenberg and de Mooij seems to generalise results of restricted validity, and partly because of misconceptions and explicit errors in other writings on this field.

---

\* This paper has been presented at the 19. National Research Meeting at the University of Bergen, Jan. 97, and at a seminar at the department of economics, NHH. Comments and suggestions from Kåre Petter Hagen, Lars Mathiesen, Agnar Sandmo, Bjørn Sandvik, Guttorm Schjelderup, and Fred Schroyen are gratefully acknowledged. As usual, the responsibility for remaining errors is mine.

## 2. Alternative definitions of double dividends

Let us introduce and briefly explain two alternative double dividend claims<sup>1</sup> that will be investigated in this paper. More precise definitions will be stated in sections 4 and 5.

### **Welfare double dividend (WDD)**

A WDD is realised if a revenue neutral tax reform i) improves environmental quality, and ii) increases «welfare exclusive of environmental quality».

WDD presupposes that «welfare» may be decomposed into two separate arguments; i) «welfare exclusive of environmental quality» (i.e., preferences over leisure and private consumption goods), and ii) «environmental quality». Given this assumption, the relevance of WDD has been motivated as follows<sup>2</sup>:

«If the strong double dividend obtained, it would be necessary only to establish the positive sign of environmental benefits to justify a given revenue-neutral environmental tax on benefit-cost ground. In light of the substantial uncertainties surrounding the magnitudes of environmental benefits from green taxes, the appeal of the strong form is quite understandable.»  
[Goulder (1995), p. 162]

### **Labour double dividend (LDD)**

A LDD is realised if a revenue neutral tax reform i) improves environmental quality, and ii) increases employment.

The interest in whether or not a LDD will materialise may be based on the following line of reasoning: Labour income taxes and/or indirect consumption taxes on non-polluting goods reduce the real wage rate, and distort the labour-leisure choice of consumers. If a revenue neutral green tax reform boosts labour supply, one might think of this as a double dividend since both improved environmental quality and reduced labour-leisure distortions are obtained<sup>3</sup>.

---

<sup>1</sup> These two definitions of DD are by no means the only ones to be found in the literature. A third definition is that it is better in terms of efficiency to recycle the tax revenue generated by an increase in the tax rates on polluting goods by cutting existing distortionary tax rates rather than by a lump sum transfer. (Goulder (1995) denotes this version the «weak form».) A fourth DD issue is the long-term effects of green tax reforms (e.g. steady state effects on private consumption and welfare), see e.g. Bye (1997)

<sup>2</sup> Goulder's «strong form» corresponds to my definition of WDD.

<sup>3</sup> Another motivation for studying LDD is the existence of unemployment, e.g. because of labour unions (Schöb (1997)) or rigid and too high wages (Bovenberg and van der Ploeg (1996)).

### 3. MODEL

The model is a slightly modified version<sup>4</sup> of the model adopted by Bovenberg and de Mooij (1994), hereafter referred to as BdM. The only production factor in the economy is labour, of which the volume is denoted  $L$ . The production possibilities are described by the following linear aggregate technology (where all productivity coefficients are normalised),

$$L = C + D + G, \quad (1)$$

where,  $C$ ,  $D$ , and  $G$  denote the volumes of a clean (non-polluting) consumption good, a dirty (polluting) consumption good, and the good consumed by the public sector, respectively. The preferences of the representative household are described by the utility function  $U = U(E, V, C, D)$ , where  $E$  is the level of environmental quality and  $V$  is the amount of leisure.  $E$  is a function of the consumption of the dirty good,  $E = e(D)$ ,  $e' < 0$ . This relationship is ignored by the household. Rather the household assumes that its consumption has no influence upon  $E$ . Thus there is a market failure which calls for the introduction of an environmental tax. The household's budget constraint is given by

$$P_C C + P_D D = I = W(T - V) = WL, \quad (2)$$

where,  $I$  denotes income,  $T$  the time endowment, and  $P_C$ ,  $P_D$ , and  $W$  are consumer (net of tax) prices. From (1) the producer prices on  $L$ ,  $C$ ,  $D$ , and  $G$ , all equal one. Choosing  $C$  as the untaxed commodity, the tax instruments are represented by the tax vector  $t = (t_L, t_D)$ , where  $t_L$  is a labour income tax and  $t_D$  is an indirect tax on commodity  $D$ . We may then rewrite the household's budget constraint as follows:

$$C + (1+t_D)D = I = (1-t_L)L. \quad (2')$$

Finally, the government's budget constraint is given by

$$t_D D + t_L L = G^0, \quad (3)$$

where  $G^0$  is the exogenous level of public consumption.

---

<sup>4</sup> Two modifications are introduced: i) There is one representative household instead of  $n$  identical households. ii) The supply of the public good,  $G$ , is not included as an argument in the household's utility function. Since the purpose of the analysis is to study revenue neutral tax reforms, the level of  $G$  will in any case be kept constant.

Following BdM, let us introduce some more structure on the consumer's preferences. In particular assume that preferences are represented by

$$U(E, H(V, C, D)). \quad (4)$$

(4) implies that environmental quality ( $E$ ) is assumed to be weakly separable from the sub-utility function  $H(\cdot)$ , which represents «welfare exclusive of environmental quality» (i.e., leisure and private consumption goods). This assumption is crucial for being able to define WDD in a meaningful way.

We may now express the private demands for  $V$ ,  $C$ , and  $D$  as functions of the tax rates, i.e.,  $V(t)$ ,  $C(t)$ , and  $D(t)$ . Since the level of environmental quality is a function of the consumption of the dirty good, environmental quality is also a function of  $t$ ,  $E(t) = e(D(t))$ . Further, the total tax revenue is defined by the tax revenue function  $R(t) = t_L L(t) + t_D D(t)$ . Finally, we define two indirect utility functions,  $H(t) = H(V(t), C(t), D(t))$  (welfare exclusive of environmental quality), and  $U(t) = U(E(t), H(t))$  (welfare inclusive of environmental quality).

#### 4. Some definitions

In terms of the variables defined above, we introduce the following definitions and sets.

**Revenue neutrality:**  $RN \equiv \{t : R(t) = G^0\}$ .

This is the set of tax rates which ensures budget balance.

For arbitrary initial tax rates  $t^0 = (t_L^0, t_D^0) \in RN$ , we further define the following sets:

**Environmental dividend:**  $E^+ \equiv \{t : E(t) > E(t^0)\}$

**Sub-utility dividend:**  $H^+ \equiv \{t : H(t) > H(t^0)\}$

**Labour dividend:**  $L^+ \equiv \{t : L(t) > L(t^0)\}$

Suppose that a tax reform from  $t^0$  to  $t^1$  takes place.

### Welfare double dividend (WDD):

The tax reform leads to a WDD if and only if  $t^l \in RN \cap E^+ \cap H^+$ .

### Labour double dividend (LDD):

The tax reform leads to a LDD if and only if  $t^l \in RN \cap E^+ \cap L^+$ .

### Fiscal optimum

A fiscal optimum is defined as<sup>5</sup>  $t^R = \operatorname{argmax}_{t \in RN} H(t)$ .

For later reference, we define the Lagrangean  $\Lambda = H(t) + \gamma(R(t) - G^0)$ , and derive the corresponding first-order conditions,

$$\frac{\partial \Lambda}{\partial t_L} = \frac{\partial H}{\partial t_L} + \gamma \frac{\partial R}{\partial t_L} = 0, \quad (5)$$

$$\frac{\partial \Lambda}{\partial t_D} = \frac{\partial H}{\partial t_D} + \gamma \frac{\partial R}{\partial t_D} = 0, \quad (6)$$

$$\frac{\partial \Lambda}{\partial \gamma} = R(t) - G^0 = 0. \quad (7)$$

### Environmental second best optimum

An environmental second best optimum is defined as<sup>6</sup>  $t^S = \operatorname{argmax}_{t \in RN} U(t)$ .

Again for later reference, we define the Lagrangean  $\Gamma = U(t) + \mu(R(t) - G^0)$  and derive the first-order conditions (8)-(10)<sup>7</sup>, where  $\lambda$  denotes the marginal utility of income,  $\lambda = \partial U / \partial I$ .

$$\frac{\partial \Gamma}{\partial t_L} = -\lambda L + \frac{\partial U}{\partial E} e' \frac{\partial D}{\partial t_L} + \mu \left[ L + t_L \frac{\partial L}{\partial t_L} + t_D \frac{\partial D}{\partial t_L} \right] = 0, \quad (8)$$

<sup>5</sup> The superscript  $R$  is shorthand for "Ramsey". We assume that  $t^R$  is unique.

<sup>6</sup> The superscript  $S$  is shorthand for «Sandmo», who first analysed the consequences of adding a negative externality to the optimal tax problem, cf. Sandmo (1975). We assume that  $t^S$  is unique.

<sup>7</sup> (8) and (9) correspond to equations (12) and (15) in Bovenberg and van der Ploeg (1994). There is one difference, however: Since environmental quality,  $E$ , is assumed to be weakly separable from the sub-utility-function  $H(V, Q(\cdot))$ , cf. (4), the level of  $E$  does not (contrary to Bovenberg and van der Ploeg's formulation) affect the private demands for  $V$ ,  $C$  and  $D$  in (8) and (9).

$$\frac{\partial \Gamma}{\partial t_D} = -\lambda D + \frac{\partial U}{\partial E} e' \frac{\partial D}{\partial t_D} + \mu \left[ D + t_L \frac{\partial L}{\partial t_D} + t_D \frac{\partial D}{\partial t_D} \right] = 0, \quad (9)$$

$$\frac{\partial \Gamma}{\partial \mu} = R(t) - G^0 = 0. \quad (10)$$

## 5. Marginal tax reforms

Marginal tax changes are of course included as special cases of the definitions in section 4. It is worthwhile, however, to state the particular results for marginal tax changes, since these results provide additional insight into the economics behind the definitions and sets stated above.

Suppose that initial tax rates are  $t^0$  and that a marginal tax reform  $dt = (dt_L, dt_D)$  takes place. From the *RN* condition we have that

$$dR = \frac{\partial R}{\partial t_L} dt_L + \frac{\partial R}{\partial t_D} dt_D = 0. \quad (11)$$

Assuming Laffer-efficiency<sup>8</sup>, i.e.,  $\partial R/\partial t_L > 0$  and  $\partial R/\partial t_D > 0$ ,  $dt_D > 0$  implies  $dt_L < 0$ . The impact on environmental quality is  $dE = e' dD$ , and since  $e' < 0$ , we need to have  $dD < 0$  in order to realise an environmental dividend,

$$dD = \frac{\partial D}{\partial t_L} dt_L + \frac{\partial D}{\partial t_D} dt_D < 0. \quad (12)$$

Combining (11) and (12), we see that  $dD$  may be rewritten as (cf. Schöb (1996), eq. (14))

$$dD = \left( \frac{\partial D / \partial t_D}{\partial D / \partial t_L} - \frac{\partial R / \partial t_D}{\partial R / \partial t_L} \right) dt_D. \quad (13)$$

---

<sup>8</sup> In the no-externality case  $\partial R/\partial t_L > 0$  and  $\partial R/\partial t_D > 0$  follows directly from (5) and (6). In the case with negative externalities, however, it is not evident that the partials of the top level utility with respect to the tax rates are negative. This makes it theoretically possible that an optimum may push tax rates into regions where (one of) the partials of  $R(\cdot)$  are negative. For relatively high levels of the tax revenue requirement, however, one would expect that the fiscal concerns are sufficiently important for the assumption  $\partial R/\partial t_L > 0$  and  $\partial R/\partial t_D > 0$  to be fulfilled.

Assuming that  $\partial D/\partial t_D < 0$  ( $D$  is not a Giffen-good), we obtain that a sufficient condition for the reform  $dt_D > 0$ ,  $dt_L < 0$  to increase  $E$ , is that  $\partial D/\partial t_L \geq 0$ . If  $\partial D/\partial t_L < 0$ , however, the ratio of own and cross price effects on  $D$ ,  $\frac{\partial D/\partial t_D}{\partial D/\partial t_L}$ , must be smaller than the ratio of the marginal tax revenues,  $\frac{\partial R/\partial t_D}{\partial R/\partial t_L}$ . Cf. Schöb (1996), Proposition 1.

The impact on welfare exclusive of environmental quality,  $dH$ , is found similarly to  $dD$ , obtaining

$$dH = \left( \frac{\partial H/\partial t_D}{\partial H/\partial t_L} - \frac{\partial R/\partial t_D}{\partial R/\partial t_L} \right) dt_D. \quad (14)$$

In other words, the tax reform increases  $H(\cdot)$  if and only if the ratio of the partials of  $H(\cdot)$  w.r.t.  $t_D$  and  $t_L$  is larger than the ratio of the marginal tax revenues.

Finally, the corresponding expression for the impact on employment is given by

$$dL = \left( \frac{\partial L/\partial t_D}{\partial L/\partial t_L} - \frac{\partial R/\partial t_D}{\partial R/\partial t_L} \right) dt_D. \quad (15)$$

For the tax reform to boost employment, the first ratio in the parenthesis in (15) must be positive and greater than the second. It follows that if the labour supply curve is upward (downward) sloping,  $\partial L/\partial t_D$  must be negative (positive) to prevent employment from unambiguously falling.

## 6. Two fundamental results

Suppose that the initial tax rates  $t^0$  are the fiscal (Ramsey) optimum,  $t^0 = t^R$ .

### Proposition 1

If the initial tax rates  $t^0 = t^R$ , there exists no tax reform satisfying WDD.

Proof:  $t^R$  maximises  $H(t) : t \in \mathbb{R}^N$ . It follows that the set  $H^+$  is empty. QED.

Next, consider the impact on top level utility of a marginal tax reform,  $dt_D > 0$ ,  $dt_L < 0$ , starting at  $t^0 = t^R$ ,

$$dU = (\partial U/\partial E)dE + (\partial U/\partial H)dH. \quad (16)$$

Assume that utility is strictly increasing in  $E$  and  $H$  and that  $dD < 0$  in (13) (whereby  $dE > 0$ ). Then  $dU > 0$  at  $t^R$  since  $dH = 0$ .<sup>9</sup> By definition of a maximum,  $dU = 0$  at the Sandmo-optimum,  $t^S$ , whereby  $(\partial U/\partial E)dE = -(\partial U/\partial H)dH$ . Therefore, starting from  $t^R$ , increasing  $t_D$  and reducing  $t_L$  along the iso-tax-revenue curve  $R(t) = G^0$ ,  $U$  increases as long as  $(\partial U/\partial E)dE > -(\partial U/\partial H)dH$ , i.e., until  $t^S$  is reached. Thus, at the Sandmo-optimum  $t^S$ ,  $E(t^S) > E(t^R)$  and  $U(t^S) > U(t^R)$ , but  $H(t^S) < H(t^R)$ . This is summarised in Proposition 2.

**Proposition 2.**

At the Sandmo-optimum  $t^S$ , utility inclusive of environmental quality is maximised, optimally balancing the utility from environmental quality against the utility from private consumption goods and leisure. It follows that the utility from private consumption goods and leisure is strictly lower than at the Ramsey-optimum  $t^R$ .

This result is rather obvious, but nevertheless crucial for the WDD debate. Assuming that the tax authorities are welfare maximising, their ultimate goal is to reform the tax system into  $t^S$ . This tax optimum implies a state of the world where the gain from a marginal increase in environmental quality equals the loss from a marginal decrease in welfare exclusive of environmental quality. In other words, there is no «free lunch» or welfare double dividend when reforming the tax system from  $t^R$  to  $t^S$ . On the contrary, there is an explicit trade-off between environmental quality and «other desirable things», viz. private consumption and leisure.

## 7. Some results and conclusions in the existing literature

### 7.1 Bovenberg and de Mooij (1994)

An important and widely cited result in the DD literature is Bdm's point that

«The main contribution of this note is to show that environmental taxes typically exacerbate, rather than alleviate, preexisting distortions — even if revenues are employed to cut preexisting distortionary taxes.» [BdM p. 1085]

The exact meaning of this statement is not obvious since the authors do not provide an explicit definition of their meaning of «preexisting distortions». What they do provide, however, is an analysis of the effect of a marginal revenue neutral tax reform,  $dt_D > 0$ ,  $dt_L < 0$ , on the level of employment. Given the following preference structure,

<sup>9</sup> The result that  $dH = 0$  at  $t^R$  follows from (5), (6), and (14). See also Schöb (1996) Corollary 1.

$$U = U(E, H(V, Q(C, D))) \quad (17)$$

where  $Q(\cdot)$  is homothetic, they conclude that

«An increase in the pollution tax from a positive initial level (i.e.  $t_D > 0$ ) reduces employment if the uncompensated wage elasticity of labour supply,  $\theta_l$ , is positive.» [BdM p. 1087]

They further state that

«Intuitively, as an instrument to finance public spending with the least costs to after-tax wages, the environmental tax, which amounts to a narrow-based tax, is less efficient than a broad-based labour because, in contrast to a labour tax, it «distorts» the composition of the consumption basket.» [BdM p. 1088]

The fact that the environmental tax  $t_D$  distorts the relative price between commodities  $C$  and  $D$  while the labour income tax does not, therefore seems to be their meaning of «exacerbate, rather than alleviate, preexisting distortions». These quotations may easily be misunderstood, since they appear to be more general than they actually are. It is not a general property of optimal tax solutions that different consumption goods should not be taxed differently. Therefore, the «narrow-based tax» ( $t_D$ ) is not in general less efficient than the labour income tax. This basic and important point is illustrated and discussed further in sections 8 and 9.

## 7.2 Goulder (1995)

In an overview, Goulder (1995) interprets BdM's results as follows:

«[Bovenberg and de Mooij] then consider the effects of a revenue-neutral policy change in which a tax is imposed on the dirty consumption commodity and the revenues are devoted to a reduction in the labor tax rate. The strong double-dividend claim is that this policy would yield an increase in nonenvironment-related welfare—that is, in the utility from the composite of consumption and leisure enjoyed by the representative household. These authors find that this claim is substantiated if and only if the uncompensated wage elasticity of labour supply is negative» [Goulder (1995) p. 162]

Since Goulder's term «the strong double divided claim» corresponds to WDD in my usage, Goulder clearly misrepresent BdM's results. First, the issue of whether or not a WDD occurs is not explicitly treated by BdM<sup>10</sup>. Second, there exists no result stating that a WDD is realised if and only if the uncompensated elasticity of labour supply is negative.

In fact, BdM's assumed preference structure is sufficient (but not necessary) for a uniform tax rate on all consumption goods to be optimal in the fiscal optimisation problem

stated above<sup>11</sup>. Since commodity  $C$  is chosen as numeraire, the Ramsey optimum is defined by having a tax on labour income only. It immediately follows from Proposition 1 that a WDD cannot be obtained if the starting point is a tax system with only a labour income tax, irrespective of whether or not the labour supply curve has a positive or negative slope. Goulder therefore seems to mistake the conditions for realising WDD and those for LDD. An illustrative example of these points is provided in section 8.

Goulder also refers to Bovenberg and van der Ploeg (1994), hereafter BvdP. To restate Goulder's argument, we solve (8) for  $\mu/\lambda \equiv \eta$  yielding (cf. BvdP eq. 18)

$$\eta = \frac{1}{1 + \frac{\partial L}{\partial t_L} \frac{t_L}{L} + (t_D + \frac{\partial U}{\partial E} e' \frac{1}{\mu}) \frac{\partial D}{\partial t_L} \frac{1}{L}} \quad (18)$$

Defining the marginal environmental damage  $MED = (\partial U/\partial E)e'(1/\lambda)$  (cf. BvdP eq. (5)), and the second best Pigouvian tax rate by  $t_{DP} \equiv -MED(1/\eta)$  (cf. BvdP eq. (14)), we may rewrite (18) as

$$\eta = \frac{1}{1 + \frac{\partial L}{\partial t_L} \frac{t_L}{L} + (t_D - t_{DP}) \frac{\partial D}{\partial t_L} \frac{1}{L}} \quad (18b)$$

This expression is commented upon by Goulder as follows:

«In their model,  $[\eta]$  exceeds unity if and only if the uncompensated wage elasticity of labour supply is positive. This is precisely the condition that denies the double dividend in its strong form.» [Goulder (1995) p. 173]

First of all, the «if and only if» in this quotation is only valid when  $t_D = t_{DP}$ , which it in general need not be<sup>12</sup>. Second, the size of  $\eta$  at the Sandmo-optimum cannot be related to the existence or non-existence of a WDD, since  $dU$  in (16) is zero at the optimum, making a WDD impossible, cf. Proposition 2. Whether or not  $\eta$  is greater or less than one in the solution of the first-order conditions (8)-(10) is therefore simply irrelevant for this issue.

<sup>10</sup> BdM show that  $dL < 0$  if the labour supply curve is upward-sloping, and that the optimal pollution tax is less than the marginal social damage from pollution, but they do not discuss effects on the sub-utility level  $H(\cdot)$ .

<sup>11</sup> This fact is stated by BdM and is referred to also by Goulder (footnote 24). For more details on the uniformity-issue, see e.g. Myles (1995), chapter 4.8, and the references therein.

## 8. Some illustrations

In this section we adopt a numerical specification of the model developed in section 3. The purpose is twofold: i) to show how alternative assumptions about the preference structure affects the existence of WDD and LDD, and ii) to relate the insight this provides to the quotations in section 7.

Suppose that the reference equilibrium has the following production and consumption pattern:  $C = D = G = 100$ ,  $L = 300$ ,  $V = 300$ ,  $E = 600$ . This equilibrium is supported by all producer prices being one, and (choosing commodity  $C$  as the numeraire good) the consumer prices being  $(W, P_C, P_D) = (1 - t_L, 1 + t_D, 1)$ , where  $t_L = 1/3$  and  $t_D = 0$  at the benchmark. In other words, the representative consumer spends half his/her available time on leisure, supplies the other half to the labour market, receives a net of tax labour income of  $L(1 - t_L) = 200$ , and spends this income on the two private consumption goods, buying 100 units of each at prices 1.0. The government receives a tax revenue of  $t_L L = 100$ , which is spent on buying hundred units of the public sector consumption good,  $G = 100$ , at price one.

### 8.1. Bovenberg and de Mooij's assumed preference structure

As stated above, BdM assume that the sub-utility function  $H(V, C, D)$  is weakly separable in  $V$  and  $(C, D)$ , i.e.,  $H(V, Q(C, D))$ . In this section we postulate the following numerical specification,

$$H(V, C, D) = \left( \alpha^{1/\sigma} V^{(\sigma-1)/\sigma} + (1-\alpha)^{1/\sigma} \left( C^\beta D^{(1-\beta)} \right)^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)}, \quad (19)$$

where  $\sigma$  is the elasticity of substitution between  $V$  and a Cobb-Douglas aggregate  $C^\beta D^{(1-\beta)}$ , and the share parameters  $\alpha$  and  $\beta$  are calibrated according to the benchmark equilibrium quantities<sup>13</sup>. Choosing  $\sigma = 1.5$  implies that the labour supply curve is upward-sloping, with a wage elasticity of supply of 0.25 at the benchmark. We may now illustrate the benchmark tax rates  $t^0 = (1/3, 0)$ , and the sets  $RN$ ,  $H^+$ ,  $E^+$  and  $L^+$  from Section 4.

---

<sup>12</sup> The result  $t_D = t_{DP}$  would occur if the negative externality is the only reason to tax commodity  $D$ . This will only be the case if it is optimal to have only a proportional labour tax in the Ramsey tax problem (the case without negative externalities). Cf. footnotes 8 and 12.

<sup>13</sup> Of course, we could adopt a CES-function also for the private consumption aggregate, but as a first example, the Cobb-Douglas case seems to be a good starting point.

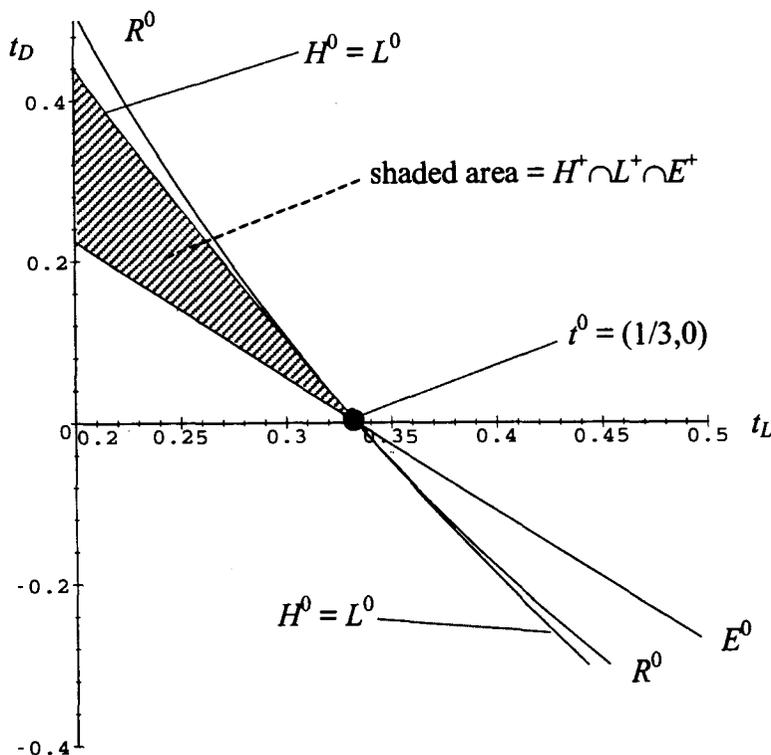


Figure 1. The contours  $H^0$ ,  $L^0$ ,  $R^0$  and  $E^0$  when the elasticity of substitution between leisure and consumption goods is 1.5.

In Figure 1 there are four contours:  $H^0 \equiv \{t : H(t) = H(t^0)\}$ ,  $R^0 \equiv \{t : R(t) = R(t^0)\}$ ,  $L^0 \equiv \{t : L(t) = L(t^0)\}$ , and  $E^0 \equiv \{t : E(t) = E(t^0)\}$ . The set  $H^+$  consists of tax rates below and to the left of  $H^0$ ,  $L^+$  of tax rates below and to the left of  $L^0$ , and  $E^+$  of tax rates above and to the right of  $E^0$ . Thus, the shaded area represents the set  $H^+ \cap L^+ \cap E^+$ .

For a revenue neutral tax reform to improve environmental quality compared with  $E(t^0)$ , there must be a movement upwards and to the left along  $R^0$ . However, no tax rate along  $R^0$  belongs to the shaded area, viz. the set  $H^+ \cap L^+ \cap E^+$ . Thus, tax reforms yielding WDD or LDD do not exist. The reason is that the preference structure in (19) implies that the starting point  $t^0$  is the constrained maximum for  $H(\cdot)$  s.t.  $t \in RN$ , i.e.,  $t^0 = t^R$ . Since  $L^+$  equals  $H^+$  in this particular example, the same is obviously the case for the employment level,  $L(\cdot)$ . Nevertheless, increasing  $t_D$  and reducing  $t_L$  marginally from  $t^0$  along  $R^0$  will be welfare improving (cf. Proposition 2). Exactly how far one has to go to reach the Sandmo-optimum  $t^S$ , depends on the exact specification of the top level utility function  $U(E, H(\cdot))$  and the environmental damage function  $e(D)$ .

How general are these results? That is the central theme for the remainder of the paper. Consider first the utility function adopted above, with a reduced elasticity of substitution between leisure and the private consumption aggregate from 1.5 to 0.5. The slope

of the labour supply curve then becomes negative, with a supply elasticity of -0.25 at the benchmark. Figure 2 displays the results.

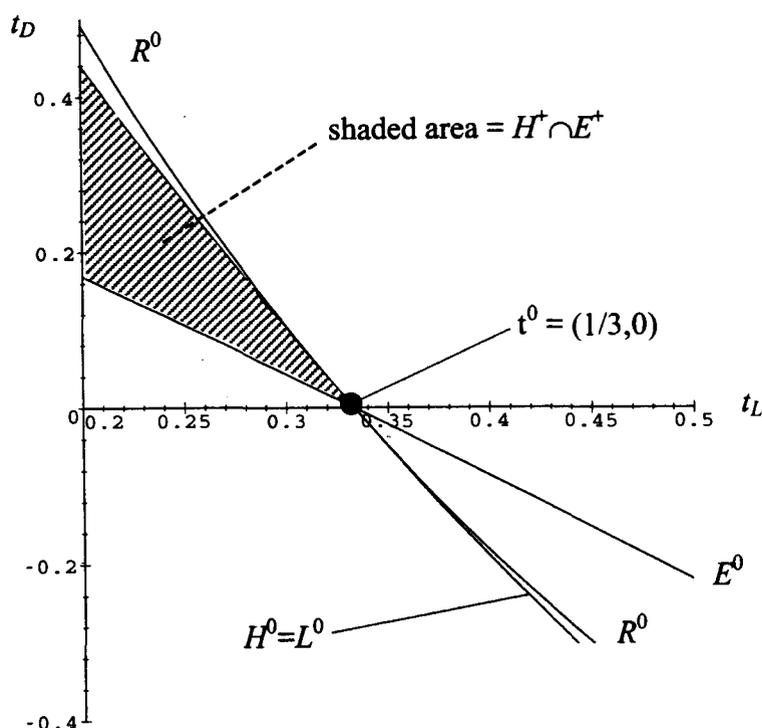


Figure 2. Reducing the elasticity of substitution between leisure and consumption goods to 0.5.

At first glance Figure 2 is hardly distinguishable from Figure 1. The fundamental difference from Figure 1, however, is that  $L^+$  now is located *above and to the right* of  $L^0$ . Therefore, the two sets  $H^+$  and  $L^+$  are disjoint,  $H^+ \cap L^+ = \emptyset$ , and the shaded area is the set  $H^+ \cap E^+ \cap L^+$ , where  $L^+$  is the complement to  $L^+$ , i.e.,  $L^+ = \{t : L(t) < L(t^0)\}$ . Thus, while  $H^0$ , as in Figure 1, is the global constrained maximum for  $H(\cdot)$ ,  $L^0$  is now the global constrained *minimum* for the employment level  $L(\cdot)$ . Thus, any movement along  $R^0$  upwards and to the left leads to a LDD, but no such movement may lead to a WDD.

Observe that since there are no *fiscal* arguments for introducing a tax on the polluting good with the assumed preference structure, the optimal tax rate on the polluting good,  $t_D^S$ , equals  $MED(1/\eta)$  at the Sandmo optimum, (8)-(10). Therefore,  $\eta$  in (18b) reduces into

$$\eta = \frac{1}{1 + \frac{\partial L}{\partial t_L} \frac{t_L}{L}} \equiv \frac{1}{1 + \varepsilon}, \quad (18c)$$

whereby  $\eta$  will be larger (less) than one whenever  $\sigma$  is larger (less) than one in the model specification adopted in this section. Combining this result with the insight provided by Figures 1 and 2, we may sum up this section as follows:

If  $\sigma$  is larger than one, we do not get a LDD, and the second best Pigouvian tax element  $t_{DP}$  is less than  $MED$  at  $r^S$ . On the other hand, if  $\sigma$  is less than one<sup>14</sup>, we *do* get a LDD, and  $t_{DP}$  exceeds  $MED$  at  $r^S$ . There is no WDD for *any* value of  $\sigma$ , however, regardless of whether or not there is a LDD and regardless of whether or not the second best optimal pollution tax is above or below  $MED$ . The fundamental reason why a WDD cannot occur is i) the assumed preference structure in (17), and ii) the assumed initial tax system with a labour income tax only. These results demonstrate that the two quotations from Goulder (1995) in section 7.2 are incorrect.

## 8.2. Different degrees of substitutability between leisure and the two private consumption goods

In search of more general results, we need to reconsider the assumption that leisure is weakly separable from private consumption goods. Suppose that preferences over leisure and private consumption have the form  $H(V,C,D) = H(D,F(V,C))$ , where both  $H(\cdot)$  and  $F(\cdot)$  are CES functions with elasticities of substitution  $\sigma^H$  and  $\sigma^F$ , respectively. With this functional form, we are free to specify different degrees of substitutability between leisure and the two consumption goods. Consider Figure 3, which is based on the same benchmark quantities and initial tax system as above, and where<sup>15</sup>  $\sigma^H = 0.75$  and  $\sigma^F = 1.25$ . The set  $H^+$  is below and to the left of  $H^0$ ,  $L^+$  above and to the left of  $L^0$ , and  $E^+$  above and to the right of  $E^0$ .

---

<sup>14</sup> While Figures 1 and 2 only show what happens if  $\sigma = 1.5$  and  $0.5$  respectively, we may readily generalise these results to include all values of  $\sigma$ , since the critical value obviously is the case with a vertical labour supply curve, i.e., the case where  $H(V,Q(\cdot))$  is a Cobb-Douglas function. This remains true also if allowing for different specifications of the consumption aggregate  $Q(C,D)$ . For example, figures similar to Figures 1 and 2 have been constructed also for cases where  $Q(C,D)$  is a CES-function, with elasticities of substitution both greater and less than one. No qualitative conclusions are influenced by such respecifications, however.

<sup>15</sup> Admittedly, this choice is ad hoc. For the existence of WDD, the crucial thing in this example is to have  $\sigma^H < \sigma^F$ , while the absolute magnitudes of  $\sigma^H$  and  $\sigma^F$  are of less importance. The combined choice of  $H(V,C,D) = H(D,F(V,C))$  and  $\sigma^H < \sigma^F$  yields the same qualitative results (in terms of WDD) as having  $H(V,C,D) = H(C,F(V,D))$  and  $\sigma^H > \sigma^F$ .

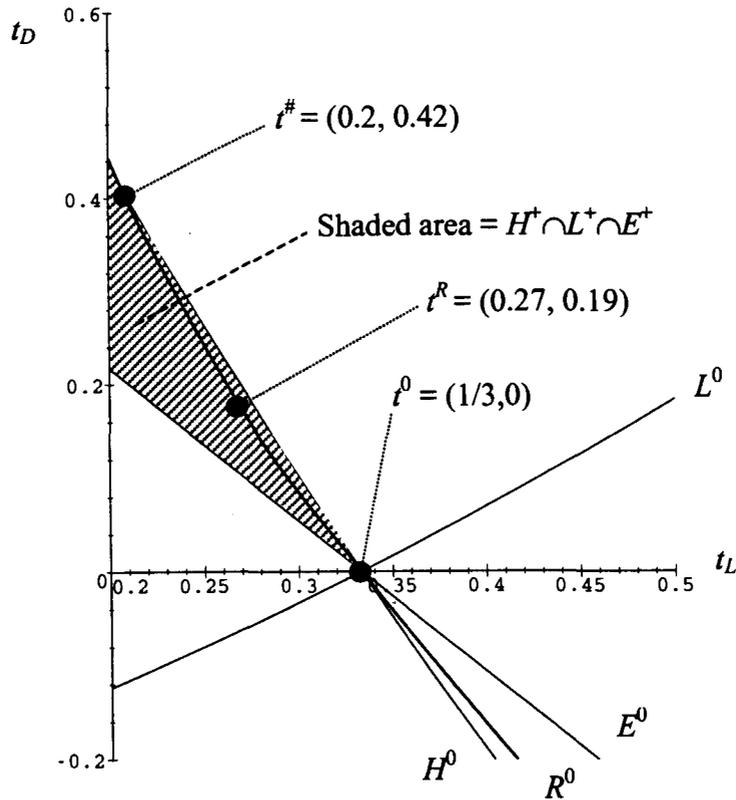


Figure 3. Higher degree of substitutability between  $V$  and  $C$  than between  $V$  and  $D$ .

The Corlett-Hague rule (see Corlett and Hague (1953-54) and Sandmo (1987)) tells us that we should levy the highest tax rate on the commodity with the highest degree of (compensated) complementarity with leisure (if both consumption goods are taxed and labour income is not). In our case, this requires that  $D$  is taxed more heavily than  $C$ . Since a proportional labour income tax is equivalent to having uniform tax rates on both commodities, the positive labour income tax must be supplemented with a positive tax rate on commodity  $D$ .

This general insight is exemplified in our particular case<sup>16</sup>, where the Ramsey optimum is  $t^R = (0.27, 0.19)$ . Starting at  $t^0$ , environmental quality, employment, and welfare exclusive of environmental quality all increase as we move upwards and to the left along  $R^0$ , since the contour  $R^0$  cuts into the shaded area, i.e.,  $H^+ \cap L^+ \cap E^+$ . The tax vector  $t^\#$  in Figure 3 is defined by  $\{H(t^\#) = H(t^0), R(t^\#) = R(t^0)\}$ . In other words, for all tax rates  $t$  along  $R^0$  between  $t^0$  and  $t^\#$ ,  $H(t)$  is higher than  $H(t^0)$ . Hence, if the preferences over environmental quality are such that the Sandmo optimum  $t^S$  is somewhere between  $t^R$  and  $t^\#$ , a discrete jump from the initial point  $t^0$  into the optimum  $t^S$  gives strictly higher level of employment, environmental quality,

<sup>16</sup> Returning to (18b), we now have that the optimal tax on commodity  $D$  at the Sandmo optimum is higher than  $t_{DP}$  (since  $t_D^R > 0$ ), whereby (18b) does not reduce into (18c).

welfare defined exclusive of environmental quality, and – needless to say – welfare inclusive of environmental quality. This example<sup>17</sup> therefore sharply contrasts the results obtained with the preference structure assumed by BdM.

For the sake of symmetry and completeness, let us finally illustrate the case where we switch the above elasticities, i.e., we stipulate that  $\sigma^F = 0.75$  and  $\sigma^H = 1.25$ .

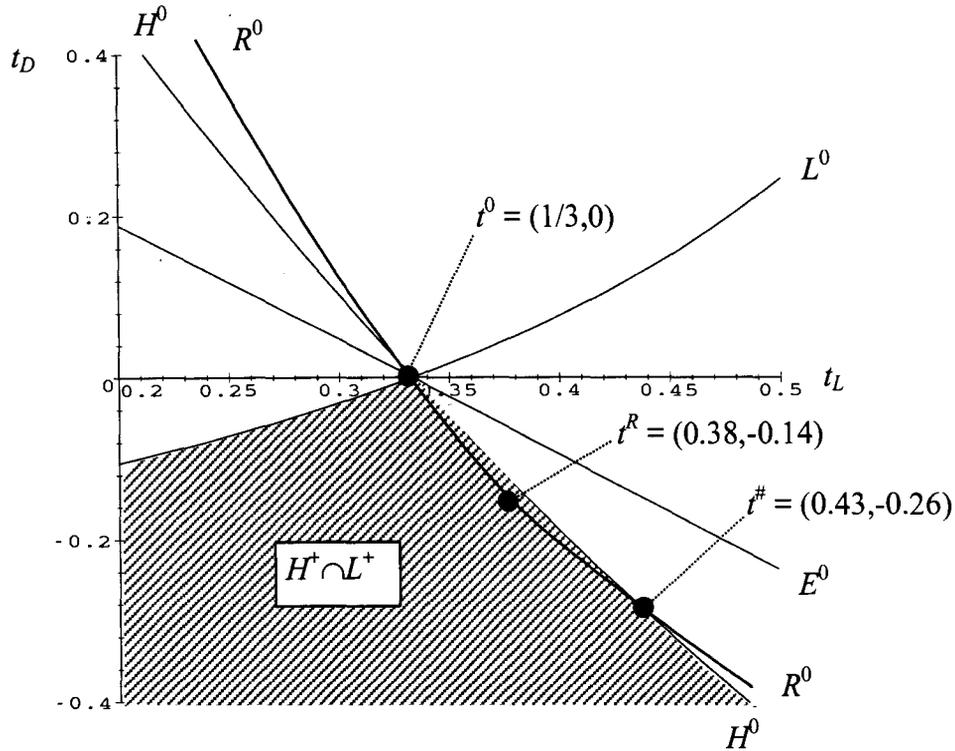


Figure 4. Higher degree of substitutability between  $V$  and  $D$  than between  $V$  and  $C$ .

With this preference structure,  $H^+$  is below and to the left of  $H^0$ ,  $E^+$  above and to the right of  $E^0$ , while  $L^+$  is below and to the right of  $L^0$ . Thus, the contour  $R^0$  cuts into the shaded area,  $H^+ \cap L^+$ , at points along  $R^0$  below and to the right of  $t^0$ . Since  $E^+$  is above and to the right of  $E^0$ , however, it follows that neither WDD nor LDD is feasible. The tax vector  $t^\# = (0.43, -0.26)$  is defined analogously to  $t^\#$  in the previous Figure, i.e.,  $\{t^\# = t : H(t^\#) = H(t^0), R(t^\#) = R(t^0)\}$ .

Without specifying the top level utility function  $U(E, H(\cdot))$  and the environmental damage function  $e(D)$  explicitly, the only thing we know about the Sandmo optimum  $t^S$  is that it is located somewhere along  $R^0$  above and to the left of  $t^R$ . If the marginal utility of environmental quality is relatively low compared to the marginal utility of the  $H$ -aggregate, it

<sup>17</sup> Parry (1995) points out the same basic insight, viz. that the degree of complementarity between the various consumption goods and leisure is an important issue for the existence of a double dividend. Due to differences in methodology, his results are somewhat hard to compare directly with those concerning WDD and LDD, however. Most importantly, Parry seems to adopt a different notion of «double dividend» than those studied in this paper, cf. his footnote 2.

is therefore possible that  $t^S$  is below and to the right of  $t^0$ . Thus, it could turn out that reforming the tax system into the Sandmo optimum  $t^S$  implies lowering the tax on commodity  $D$  while increasing the income tax! By so doing, we would obtain higher employment, higher welfare exclusive of environmental quality, and higher top level utility, but reduced environmental quality.

Before closing this section, we generalise the main insight provided by Figures 3 and 4. Given the assumed preferences,  $H(V,C,D) = H(D,F(V,C))$ , the crucial factor for whether or not we obtain a WDD from the starting point  $t^0$  is the *relative* magnitudes of  $\sigma^H$  and  $\sigma^F$ . In general, all combinations of  $\sigma^F$  and  $\sigma^H$  such that  $\sigma^F > \sigma^H$  implies that  $t_L^R, t_D^R > 0$  (cf. Figure 3), while  $\sigma^F = \sigma^H > 0$  implies<sup>18</sup>  $t_L^R > 0, t_D^R = 0$  (cf. Figures 1 and 2), and  $\sigma^F < \sigma^H$  implies  $t_L^R > 0, t_D^R < 0$  (cf. Figure 4). In other words, it suffices for  $\sigma^F$  to be marginally higher than  $\sigma^H$  (whereby  $D$  has a marginally higher degree of complementarity with leisure than  $C$ ) for a marginal tax reform  $dt_D > 0, dt_L < 0$  from  $t^0$  to produce a WDD.

## 9. Concluding comments

Bovenberg and de Mooij (1994) has served as an influential benchmark for the literature on green taxes and double dividends. Several recent contributions have extended BdM's results, e.g. by including intermediate inputs to production and by allowing for imperfections in the labour market. Rather than investigating the consequences of extending the model, this paper retains BdM's assumptions, and focuses on how the results obtained within this model framework hinges on alternative assumptions regarding the representative consumer's preferences.

BdM's results have recently been reinterpreted by Fullerton (1997), who focuses on the role played by different price normalisations. Fullerton's comment is an important contribution for the discussion of whether or not the second best optimal tax rate on the polluting commodity is above or below the first best optimal level, viz. the marginal environmental damage. However, while different price normalisations change the *absolute* tax rates and consumer prices in the model, all relative prices and real variables are unaffected. The approach in this paper has therefore been to investigate the consequences of revenue neutral tax reforms for the employment level, environmental quality, and welfare exclusive and inclusive of environmental quality.

---

<sup>18</sup> The requirement that both  $\sigma^F$  and  $\sigma^H$  are strictly positive is included since  $\sigma^F = \sigma^H = 0$  implies that  $t^R$  is no longer unique. If  $\sigma^F = \sigma^H = 0$  there are no tax distortions (Leontief preferences), and all tax rates  $\{t : R(t) = G\}$  give the same value for  $H(t)$ .

A main conclusion from the previous section is that quite a number of possibilities arise once we depart with the assumption that leisure is weakly separable from the private consumption goods. This suggests that the results derived by BdM are too specific to support the apparently general conclusions cited above in section 7.1.

First, there is hardly a «typical» effect from revenue neutral green tax reforms. Rather, the effects on two of the crucial variables for the double dividend issue, employment and welfare exclusive of environmental quality, are highly sensitive to the exact specification of preferences and initial tax structure. Unless one believes that BdM's assumptions are the empirically «typical» ones, their results therefore need to be supplemented with some counterfactual scenarios.

Second, it is clearly not a general fact that a «narrow based tax» like the tax on commodity  $D$  in our model is less efficient from a fiscal point of view than the «broad based» labour income tax. Of course, the tax on commodity  $D$  changes the relative price between the two private consumption goods, while the labour income tax does not. Since there are three commodities in the model, however, there are also three relative consumer prices. There exists no general result in the theory of optimal commodity taxation which commands us to minimise the number of relative price changes. Neither does there exist a result saying that the relative efficiency of the available tax instruments may be ranked according to the sizes of the respective tax bases. Without detailed empirical insight into the complete system of own and cross price effects, few general statements can therefore be made about the relative efficiency of the set of available tax instruments.

## References:

- Bovenberg, A.L. and R. de Mooij (1994): Environmental levies and distortionary taxation, *American Economic Review* 84, 1085-1089.
- Bovenberg, A.L. and L.H. Goulder (1996): Optimal environmental taxation in the presence of other taxes: general-equilibrium analysis, *American Economic Review* 86, 985-1000.
- Bovenberg, A.L. and F. van der Ploeg (1994): Environmental policy, public finance and the labour market in a second best world, *Journal of Public Economics* 55, 349-390.
- Bovenberg, A.L. and F. van der Ploeg (1996): Optimal taxation, public goods and environmental policy with involuntary unemployment, *Journal of Public Economics* 62, 59-83.
- Bye, B. (1997): Taxation, unemployment and growth: Dynamic welfare effects of "green" policies, Discussion Paper No. 183, Statistics Norway.
- Corlett, W.J. and D.C. Hague (1953-54): Complementarity and the excess burden of taxation, *Review of Economic Studies* 21, 21-30.
- Fullerton, D. (1997): Environmental levies and distortionary taxation: comment, *American Economic Review* 87, 245-51.
- Goulder, L. H. (1995): Environmental taxation and the double dividend: A reader's guide, *International Tax and Public Finance* 2, 157-183.
- Håkonsen, L. (1996): Fiscal and environmental efficiency: A trade off rather than a double dividend, Discussion Paper 24/96, Dep. of Economics, Norwegian School of Economics and Business Administration, Bergen, Norway.
- Lee, D.R. and W.S. Misiolek (1986): Substituting pollution taxation for general taxation: Some implications for efficiency in pollutions taxation, *Journal of Environmental Economics and Management* 13, 338-347.
- Myles, G. D. (1995): *Public Economics*, Cambridge University Press.
- Oates, W.E. (1995): Green taxes: Can we protect the environment and improve the tax system at the same time?, *Southern Economic Journal* 61, 915-22.
- Parry, I.W.H. (1995): Pollution taxes and revenue recycling, *Journal of Environmental Economics and Management* 29, S-64 - S-77.
- Pearce, D.W. (1991): The role of carbon taxes in adjusting to global warming, *Economic Journal* 101, 938-948.
- Sandmo, A. (1975): Optimal taxation in the presence of externalities, *Swedish Journal of Economics* 77, 86-98.
- Sandmo, A. (1987): A reinterpretation of elasticity formulae in optimum tax theory, *Economica* 54, 89-96.
- Schöb, R. (1996): Evaluating tax reforms in the presence of externalities, *Oxford Economic Papers* 48, 537-555.
- Schöb, R. (1997): Do green taxes on factors and consumption alleviate unemployment?, paper presented at Research Forum on Taxation, June 2-4, Oslo, Norway.
- Terkla, D. (1984): The efficiency value of effluent tax revenues, *Journal of Environmental Economics and Management* 11, 107-123.

## Chapter 4. Negative externalities, dead weight losses, and the cost of public funds\*

### 1. Introduction

Taxes are often divided into the following three broad categories: i) efficiency improving taxes, ii) neutral taxes, and iii) distortionary taxes. The most prominent example of taxes in category i) is the Pigouvian tax, i.e., a tax on a negative externality. In category two we could think of lump sum taxes and profits taxes<sup>1</sup>, and the last category contains most other taxes like income taxes and various indirect consumption taxes. As instruments for raising tax revenues, these taxes have fundamentally different characteristics. An optimal Pigouvian tax both raises tax revenues and corrects for the distortion created by the negative externality. A neutral tax raises tax revenues without either creating new or correcting existing distortions. Finally, the distortionary taxes raises tax revenues but have the unfortunate effect that they create dead weight losses.

In this paper we focus on how the existence of negative externalities influence the marginal cost of public funds (MCF). Is it a general fact that MCF is reduced if there are negative externalities present in the economy? Is MCF less than one if raising tax revenues by means of (first or second best) optimal Pigouvian taxes? Could MCF drop below zero? These are the fundamental questions to be analysed in this paper.

Previous studies based on computable general equilibrium models (Ballard and Medema (1993) and Brendemoen and Vennemo (1996)) have reported MCF values less than one when there exist negative externalities. This paper argues that such results mainly are due to the economies in question being outside the optimum, however.

In an economy where a first best optimum (for any level of tax revenues) is realised by means of an optimal Pigouvian tax and a lump sum tax, it is argued that MCF is one. In other words, first best optimal taxation implies that MCF is one both with and without externalities present in the economy.

---

\* This paper has been presented at the Norwegian School of Economics and Business Adm. (Nov. 97) and at the Nordic Tax Seminar in Oslo (Dec. 97). Vidar Christiansen, Lars Mathiesen, Agnar Sandmo, and Bjørn Sandvik have provided valuable comments and suggestions.

<sup>1</sup> Provided that taxable profits equals factual profits.

If lump sum taxation is not available, second best optimal taxation is defined in terms of taxes on externality producing goods alongside with taxes on other (non externality producing) goods, cf. Sandmo (1975) and Bovenberg and van der Ploeg (1994a). When solving the optimal tax problem numerically, we find that the presence of externalities fundamentally changes the optimal tax rates compared to the no-externality case. However, the change in MCF is relatively small except for relatively low levels of collected tax revenues. In fact, there is a striking similarity between the behaviour of MCF in cases with and without externalities in the numerical examples studies in this paper.

The paper adopts the same model framework as in Bovenberg and van der Ploeg (1994a). This model is described in section 2. Section 3 analyses first best optimal taxation, and derives MCF in this setting. Section 4 derives the optimal tax solution for the second best case where the lump sum tax may not be used. In section 5 we develop maximum value functions from which the total, average and marginal dead weight losses are derived, while section 6 describes alternative cost of funds measures. In section 7 a numerical simulation model is used in order to compute the previously derived measures, while section 8 concludes the paper.

## 2. Model outline

Several recent articles on environmental taxation, e.g. Bovenberg and de Mooij (1994) and Bovenberg and van der Ploeg (1994a), have been based on more or less the same model. This model serves well for studying the topics to be considered also here. Parts of the analysis lends heavily from Bovenberg and van der Ploeg (1994a), hereafter referred to as BP.

Let  $U(E, x_0, x_1, x_2)$  express a representative consumer's preferences over environmental quality, leisure, and two private consumption goods, respectively. Let further  $T$  denote the consumer's time endowment and  $P = (P_0, P_1, P_2)$  the vector of consumer prices for  $x_0, x_1$ , and  $x_2$ . The consumer's budget constraint is

$$P_1 x_1 + P_2 x_2 = P_0 (T - x_0) - a, \quad (1)$$

where  $a$  is a lump-sum tax. Maximisation of  $U(\cdot)$  subject to (1) yields the demand functions for  $x_0, x_1$ , and  $x_2$ , respectively, i.e.,  $x_i(P, a, E)$ , for  $i = 0, 1, 2$ , and the indirect utility function  $V(P, a, E)$ . The negative externality is modelled as a feedback from the consumption of  $x_2$  on  $E$ ,  $E = e(x_2)$ ,  $e' < 0$ . We assume that this relationship is not internalised in the consumer's

optimisation problem, however. Rather, it is assumed that  $E$  is regarded as an exogenous parameter from the consumer's point of view, i.e., there is a market failure which calls for a price correction.

Labour is the only factor of production, and the production technology, (2), exhibits constant returns with fixed productivity coefficients. For simplicity, all productivity coefficients are normalised.  $G$  denotes a commodity consumed and financed by the public sector.

$$x_0 + x_1 + x_2 + G = T \quad (2)$$

It is assumed that the public sector has no other source of income than taxes, so that the expenditures on  $G$  must be met by a corresponding amount of tax income. Given (2) and assuming perfect competition, the producer prices of final consumption goods,  $p = (p_1, p_2, p_G)$ , equal the producer price of labour,  $w$ , which is chosen as the numeraire.

The available tax instruments are  $(t_L, t_1, t_2, a)$ , where  $t_L$  is a labour income tax,  $t_1$  and  $t_2$  are indirect consumption taxes on commodities 1 and 2, and  $a$  is a lump sum tax. We do not impose non-negativity for the tax rates. Since there are no pure profits in this economy, one of the tax rates may without loss of generality be set to zero, cf. Munk (1978). In the following, we choose  $t_1 = 0$ , whereby  $P_1 = p_1 = 1$ . Given the above tax system and the fact that all producer prices equal one, the consumer's budget constraint may be restated as

$$x_1 + (1 + t_2)x_2 = (1 - t_L)(T - x_0) - a, \quad (1')$$

whereby we may express the demand functions  $x_i(\cdot)$ , and the indirect utility function  $V(\cdot)$  as functions of the tax rates  $(t_L, t_2, a)$  and the level of environmental quality.

### 3. First best optimum

#### 3.1. Command economy

Consider a command economy where a central planner may choose the level of each  $x_i$  subject to the material balance condition (2) and a pre-specified level of public consumption,  $G = G^0$ . The condition for Pareto-optimality (see also BP, equation (2)) is that<sup>2</sup>

---

<sup>2</sup> We use the notation  $\partial U / \partial x_i = U_i$ .

$$U_0 = U_1 = U_2 + U_E e'. \quad (3)$$

From (3) we have that the marginal rate of substitution between the externality producing good and the numeraire is  $U_2/U_1 = 1 - U_E e'/U_1$ , where  $-U_E e'/U_1 > 0$  is the term accounting for the negative externality created by the consumption of  $x_2$ . For later reference, we define the marginal environmental damage (*MED*) by

$$MED = -\frac{U_E e'}{U_1}. \quad (4)$$

In a competitive equilibrium, (3) is realised by setting  $t_L = 0$  and  $t_2^P = MED$ , where  $t_2^P$  denotes the Pigouvian tax on  $x_2$ . Since the resulting tax revenue may not satisfy the material balance condition with the pre-specified level of public consumption  $G^0$ , the lump-sum tax is scaled so as to realise an equilibrium which satisfies (2) and (3).

### 3.2. Dual analysis

A thorough comparison of first- and second best tax optima is greatly facilitated by restating the conditions for having a first best optimum (i.e., an allocation satisfying (2) and (3)) in terms of a dual maximisation problem. In the first-best situation where all tax instruments are available, the public sector budget constraint is<sup>3</sup>

$$R(t_L, t_2, a) = t_L l + t_2 x_2 + a \geq G^0, \quad (5)$$

where  $l \equiv (T - x_0)$  denotes labour supply and  $R(\cdot)$  the tax revenue function. A maximum for  $V(t_L, t_2, a)$  with respect to  $(t_L, t_2, a)$  subject to (5) is a point of tangency<sup>4</sup> between a contour for  $V(t_L, t_2, a)$  and the border of the constraint set  $\{(t_L, t_2, a) : R(t_L, t_2, a) \geq G^0\}$ . Such tangency points satisfy the first-order conditions (7)-(10), derived<sup>5</sup> from the Lagrange function (6). Let  $\mu$  be the multiplier associated with the tax revenue constraint (5), and let  $\lambda$  be the multiplier associated with the budget constraint (1) in the consumer's utility maximisation problem, i.e.,

<sup>3</sup> Since the demand functions are derived from the consumer's utility maximisation problem, the consumer's budget constraint (2) is automatically fulfilled. Hence, by Walras' law, an allocation satisfying (5) also satisfies the material balance constraint (2).

<sup>4</sup> Since we do not impose non-negativity constraints on the tax instruments, we do not need to consider corner solutions.

the marginal utility of lump-sum income or, equivalently, the marginal utility of the numeraire,  $\lambda = U_1$ .

$$L(t_L, t_2, a, \mu) = V(t_L, t_2, a) + \mu [R(t_L, t_2, a) - G^0] \quad (6)$$

$$\frac{\partial L}{\partial a} = -\lambda + \tilde{U}_E e^{\frac{\partial x_2}{\partial a}} + \mu \left[ 1 + t_L \frac{\partial l}{\partial a} + t_2 \frac{\partial x_2}{\partial a} \right] = 0 \quad (7)$$

$$\frac{\partial L}{\partial t_L} = -\lambda l + \tilde{U}_E e^{\frac{\partial x_2}{\partial t_L}} + \mu \left[ l + t_L \frac{\partial l}{\partial t_L} + t_2 \frac{\partial x_2}{\partial t_L} \right] = 0 \quad (8)$$

$$\frac{\partial L}{\partial t_2} = -\lambda x_2 + \tilde{U}_E e^{\frac{\partial x_2}{\partial t_2}} + \mu \left[ x_2 + t_L \frac{\partial l}{\partial t_2} + t_2 \frac{\partial x_2}{\partial t_2} \right] = 0 \quad (9)$$

$$\frac{\partial L}{\partial \mu} = a + t_L l + t_2 x_2 - G^0 \geq 0, \quad \mu \geq 0, \quad \mu \frac{\partial L}{\partial \mu} = 0 \quad (10)$$

BP denote the term  $\tilde{U}_E$  in (7)-(9) as *the marginal social utility of the environment*, defined as

$$\tilde{U}_E \equiv \left( \frac{U_E + \mu \left[ t_L \frac{\partial l}{\partial E} + t_2 \frac{\partial x_2}{\partial E} \right]}{1 - e^{\frac{\partial x_2}{\partial E}}} \right) \quad (11)$$

While  $U_E$  is the direct impact of a marginal increase in environmental quality on utility,  $\tilde{U}_E$  in addition includes the indirect effects of a marginal increase in environmental quality on the taxed commodities. Observe that  $\tilde{U}_E$  and  $U_E$  coincide if  $E$  is weakly separable from leisure and the private consumption goods.

In the first best case where the lump sum tax is available,  $t_L$  is redundant in the sense that maximising  $V(t_L, t_2, a)$  subject to (6) always yields the solution  $t_L = 0$ . In other words, in the three-dimensional space spanned by  $t_L, t_2, a$ , all first best solutions occur in the two dimensional plane spanned by  $t_2$  and  $a$ , where  $t_L = 0$  (see Appendix 2). Furthermore, it is

---

<sup>5</sup> See Appendix 1 for the derivation of equations (7)-(9).

shown in Appendix 2 that the solution to (7)-(10) have the following properties (with optimal values denoted by superscripts  $FB$ ):

$$\begin{aligned}
t_2^{FB} &= t_2^P = -\frac{U_E e'}{\lambda} = MED, \\
a^{FB} &= G^0 - t_2^{FB} x_2, \\
t_L^{FB} &= 0, \\
\mu^{FB} &= \lambda^{FB}, \\
\tilde{U}_E &= U_E.
\end{aligned} \tag{12}$$

From (12), we see that for all levels of the exogenous tax revenue requirement  $G^0$  the tax on  $x_2$  fully internalises the marginal environmental damage, and that the lump sum tax balances the tax revenue requirement. Further, since lump sum taxation is available, the marginal cost of public funds,  $(\mu/\lambda)^{FB}$  equals one, and the social marginal utility of the environment equals the direct impact on private utility,  $\tilde{U}_E = U_E$ . Recollect that  $(\mu/\lambda)^{FB}$  equals one also in the case where first best taxation is available in models without negative externalities. The intuition is that in both situations (with and without externalities), any level of tax revenues may be raised without creating dead weight losses. Thus, raising more tax revenues changes the distribution of income, but leaves efficiency unaffected.

#### 4. Second best optimum

In the second best situation, the lump-sum tax  $a$  is no longer available. In technical terms this amounts to pre-specifying  $a = 0$  in the model relations<sup>6</sup>. Two endogenous variables remains;  $t_L$  and  $t_2$ . The second best tax optimisation problem is therefore  $\max V(t_L, t_2)$  with respect to  $(t_L, t_2)$  subject to  $R(t_L, t_2) = t_L l + t_2 x_2 \geq G^0$ . The relevant first order conditions are then (8)-(10). These conditions are explored extensively in BP, which may be consulted for further details. In geometrical terms, a point  $(t_L, t_2)$  satisfying (8)-(10) is a point of tangency between a contour for  $V(t_L, t_2)$  and the border of the constraint set  $\{(t_L, t_2) : R(t_L, t_2) \geq G^0\}$ , see figure 4b.

Solving (8) and (9) for  $t_L$  and  $t_2$ , we find that

---

<sup>6</sup> We could alternatively require that  $a \leq 0$ , i.e., that lump sum transfers are allowed while lump sum taxation is not possible. Of course, the model adopted here (with one representative consumer) does not itself explain why lump sum taxation may not be used. Rather, the infeasibility of lump sum taxation is external to the model framework.

$$t_L^{SB} = \frac{\left(1 - \frac{\lambda}{\mu}\right)}{A} \left( x_2 \frac{\partial x_2}{\partial t_L} - l \frac{\partial x_2}{\partial t_2} \right), \text{ and} \quad (13)$$

$$t_2^{SB} = \frac{\left(1 - \frac{\lambda}{\mu}\right)}{A} \left( l \frac{\partial l}{\partial t_2} - x_2 \frac{\partial l}{\partial t_L} \right) + \frac{\lambda - \tilde{U}_E e'}{\mu U_1}, \quad (14)$$

where  $A = \frac{\frac{\partial l}{\partial t_L} \frac{\partial x_2}{\partial t_L}}{\frac{\partial l}{\partial t_2} \frac{\partial x_2}{\partial t_2}}$ . (13) and (14) correspond<sup>7</sup> to (22) and (23) in Sandmo (1975). The

expressions for  $t_L$  and the first term in the expression for  $t_2$  are identical to the Ramsey tax scheme, i.e., the second best optimum without the negative externality. In (14), we see the separability property described by Sandmo (1975); the second best optimal tax on  $x_2$  is a weighted sum of a Ramsey-term defined in terms of the derivatives of the demand system, and a term containing the marginal social utility of the environment, with  $(1-\lambda/\mu)$  and  $\lambda/\mu$  being the weights respectively.

In order to rewrite the first order conditions in an illuminating way, let us define (cf. BP eq. (14)) the price corrective term in (14) by

$$-\frac{\tilde{U}_E e' \lambda}{U_1 \mu} \equiv t_2^{PC}. \quad (15)$$

This makes clear that there are two reasons why the price corrective part of a second best optimum,  $t_2^{PC}$ , may differ from MED: i) It is defined in terms of the marginal social utility of the environment,  $\tilde{U}_E$ , while MED is defined in terms of the direct marginal utility,  $U_E$ , and ii)  $\mu$  and  $\lambda$  do not in general coincide in a second best situation.

Using (15), we may reformulate (8) and (9) as follows:

$$t_L \frac{\partial l}{\partial t_L} + (t_2 - t_2^{PC}) \frac{\partial x_2}{\partial t_L} = \frac{\lambda - \mu}{\mu} l \quad (8')$$

<sup>7</sup> Note that  $\tilde{U}_E$  in (15) is replaced by  $U_E$  in Sandmo's eq. (23). This difference is due to an implicit assumption made by Sandmo, viz. that a marginal change in  $E$  does not affect the demands for the taxed commodities.

$$t_L \frac{\partial l}{\partial t_2} + (t_2 - t_2^{PC}) \frac{\partial x_2}{\partial t_2} = \frac{\lambda - \mu}{\mu} x_2 \quad (9')$$

From the previous section it is clear that a first best optimal allocation must be characterised by  $\{t_L = 0, t_2 = \text{MED} = -\frac{U_E e'}{U_1}\}$ . In the second-best situation, suppose that the price corrective element,  $t_2^{PC}$ , exactly fulfils the tax revenue requirement, such that  $\{t_L = 0, t_2 = t_2^{PC}\}$ . Then, from (8') and (9'),  $\mu = \lambda$ . Furthermore, using  $\{t_L = 0, t_2 = t_2^{PC}, \mu = \lambda\}$  and (11) in (15), we find that  $\tilde{U}_E = U_E$ , whereby  $t_2^{PC}$  reduces in  $t_2^P = \text{MED}$ , which is the first best optimal Pigouvian tax. In other words, a Pareto optimum is realised although the lump sum tax is not available in the special case where public expenditures equal the tax revenue generated by the optimal Pigouvian tax alone,  $R(t_L, t_2) = t_2^P x_2(0, t_2^P) \equiv G^P$ . Any other level for  $G$ , requires a positive or negative labour income tax rate in addition to the tax on  $x_2$ . Hence, in the case where the lump sum tax may not be used, a Pareto optimum may only be sustained for the public expenditure level  $G^P$ , whereby dead-weight losses accrue for all  $G \neq G^P$ .

## 5. Maximum utility and tax revenue functions

In this section we define four maximum value functions and their slopes. With these tools, we may develop interesting measures for the total and marginal dead weight loss, and the total and marginal cost of public funds.

### 5.1. Maximum utility functions

The solution to the first best maximisation problem,  $\max V(t_L, t_2, a)$  s.t.  $R(t_L, t_2, a) \geq G$  is found from (7)-(10), yielding  $\{t_L^{FB}(G), t_2^{FB}(G), a^{FB}(G), \mu^{FB}(G)\}$ . Let us denote the corresponding maximum value function by the *first best maximum utility function*,  $V^{FB}(G) \equiv V(a^{FB}(G), t_L^{FB}(G), t_2^{FB}(G))$ . From the envelope theorem, we have that  $\mu^{FB} = -dV^{FB}(G)/dG$ . Therefore,  $\mu^{FB}$  is the negative of the slope of  $V^{FB}(G)$ , measuring the negative of the change in the maximum utility level following from a marginal increase in  $G$ .

Correspondingly, the solution to the second best optimisation problem  $\max V(t_L, t_2)$  s.t.  $R(t_L, t_2) \geq G$  is  $\{t_L^{SB}(G), t_2^{SB}(G), \mu^{SB}(G)\}$  characterised by (8)-(10) (with  $a \equiv 0$ ). This yields the *second best maximum utility function*,  $V^{SB}(G) \equiv V(t_L^{SB}(G), t_2^{SB}(G))$ . Again, the interpretation of  $\mu^{SB}$  is that  $\mu^{SB} = -dV^{SB}/dG$ . In other words,  $\mu^{SB}$  measures the negative of the

change in the maximum utility level from a marginal increase in  $G$  when the increase in  $G$  must be financed by means of distortionary taxes.

By definition, we must have that  $V^{FB}(G) \geq V^{SB}(G)$ . Furthermore, since the solutions to the first- and second best problems coincide at  $G = G^P$ , we have that  $V^{FB}(G^P) = V^{SB}(G^P)$ , while  $V^{FB}(G) > V^{SB}(G)$  for all levels of  $G$  where second best taxation produces dead weight losses. Finally,  $V^{FB}(G)$  and  $V^{SB}(G)$  must obviously be declining functions of  $G$ . Based on these facts we may sketch the following figure<sup>8</sup>.

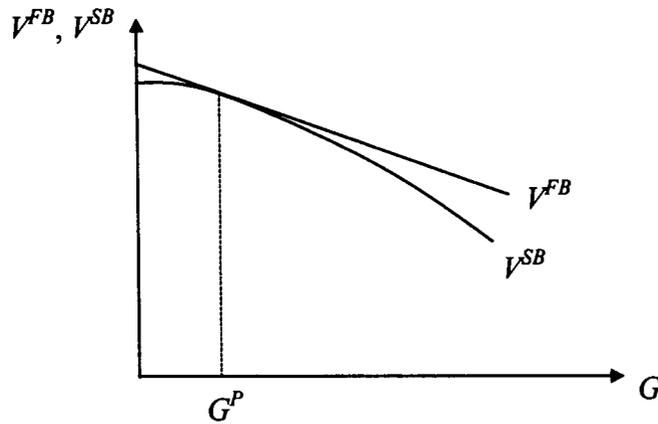


Figure 1. First and second best maximum utility functions

This figure immediately reminds us about the relationship between long and short run costs, where the long run cost function envelopes the short run cost function. The central point is that both the functions and their derivatives coincide at  $G^P$ , i.e., we have that  $V^{FB} = V^{SB}$  and  $\mu^{FB} = \mu^{SB}$  at  $G^P$ . The vertical distance between  $V^{FB}$  and  $V^{SB}$  represents the total dead weight loss measured in units of utility (the amount by which utility would increase while holding the tax revenue constant if going from second best to first best taxation). Since the utility-difference ( $V^{FB} - V^{SB}$ ) is not invariant to arbitrary positive monotone transformations of the utility index, we develop an alternative measure of the dead weight loss in the next section.

## 5.2. First and second best maximum tax revenue functions

Consider the following alternative formulation of the optimal tax problem,

$$\max R(t_L, t_2, a) \text{ s.t. } V(t_L, t_2, a) \geq U, \quad (16)$$

<sup>8</sup> We have only established that the slope of  $V^{FB}(G)$  is negative. For simplicity,  $V^{FB}(G)$  is shown as a linear function of  $G$  in Figure 1.

and the corresponding Lagrangean,  $\Lambda = R(t_L, t_2, a) + \beta[V(t_L, t_2, a) - U]$ , where  $U$  is some pre-specified utility level. Of course, the structure of the first order conditions shown in (7) - (10) does not change, except for the obvious fact that the multiplier  $\mu$  in front of the partials of  $R(\cdot)$  disappear and that the multiplier  $\beta$  enters in front of the partials of  $V(\cdot)$ . It is therefore of little interest for our purposes to restate the first order conditions for this alternative formulation of the optimisation problem (see Appendix 4, however). The crucial thing is that the dual representation of the maximisation problem enables us to define measures of the total, average and marginal dead weight losses which are invariant with respect to positive monotone transformations of the utility index.

Let  $\{t_L^{FB}(U), t_2^{FB}(U), a^{FB}(U), \beta^{FB}(U)\}$  be the first best solution<sup>9</sup> to (16). We define the corresponding maximum value function as the *first best maximum tax revenue function*,  $R^{FB}(U) \equiv R(t_L^{FB}(U), t_2^{FB}(U), a^{FB}(U))$ . It again follows from the envelope theorem that  $\beta^{FB} = -dR^{FB}/dU$  measures the increase in maximum tax revenues following from a marginal decrease in the utility requirement.

In the second best case, let the solution to (16) be  $\{t_L^{SB}(U), t_2^{SB}(U), \beta^{SB}(U)\}$ , and let the corresponding maximum value function be  $R^{SB}(U) \equiv R(t_L^{SB}(U), t_2^{SB}(U))$ , with slope  $\beta^{SB} = -dR^{SB}/dU$ . It follows that  $\beta^{SB}$  measures the increase in maximum second best financed tax revenue following from a marginal decrease in the utility requirement.

Defining  $U$  in (16) as  $U \equiv V^{SB}(G)$ , we obtain the composite maximum value functions  $\rho^{FB}(G) \equiv R^{FB}(V^{SB}(G))$  and  $\rho^{SB}(G) \equiv R^{SB}(V^{SB}(G))$ . By construction, we have that  $\rho^{SB}(G) = G$ . The relationship between  $\rho^{SB}(G)$  and  $\rho^{FB}(G)$  becomes a mirror image of that of  $V^{FB}(G)$  and  $V^{SB}(G)$  shown in Figure 1. We by definition have that  $\rho^{FB}(G) \geq \rho^{SB}(G)$  for all  $G$  and that  $\rho^{FB}(G^P) = \rho^{SB}(G^P)$ . These results are illustrated in Figure 2.

---

<sup>9</sup> If  $U$  in (16) equals  $V^{FB}(t_L^{FB}(G), t_2^{FB}(G), a^{FB}(G))$ , we by definition have that  $\{t_L^{FB}(U), t_2^{FB}(U), a^{FB}(U)\}$  equals  $\{t_L^{FB}(G), t_2^{FB}(G), a^{FB}(G)\}$ .

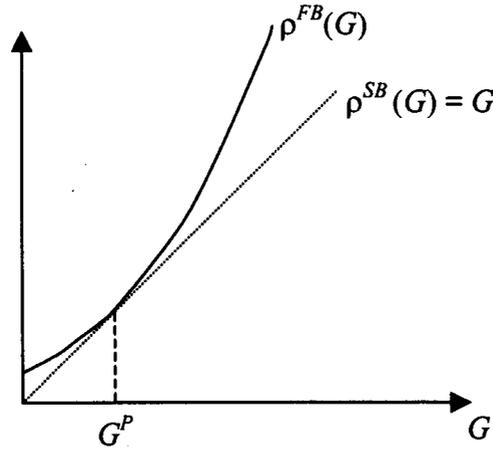


Figure 2. First and second best maximum tax revenue functions

Based on Figure 2, we may define the total, average and marginal dead weight losses generated by distortionary taxation. The total dead weight loss (measured in units of tax revenue) is found as the vertical distance between  $\rho^{FB}$  and  $\rho^{SB}$ ,

$$TDWL(G) \equiv \rho^{FB}(G) - G, \quad (17)$$

measuring the extra tax revenues which could have been collected by switching from second to first best taxation while keeping the consumer at the utility level  $U = V^{SB}(G)$ . The average dead weight loss is of course  $TDWL(G)/G$ ,

$$ADWL(G) = \frac{\rho^{FB}(G)}{G} - 1. \quad (18)$$

Equally obvious, the marginal dead weight loss is  $TDWL'(G)$ , which (somewhat less obvious<sup>10</sup>) becomes

$$MDWL(G) = \frac{\beta^{FB}(G)}{\beta^{SB}(G)} - 1. \quad (19)$$

<sup>10</sup> We have that  $dTDWL(G)/dG = d\rho^{FB}/dG - 1 = (dR^{FB}/dU)(dV^{SB}/dG) - 1 = \beta^{FB}\mu^{SB} - 1$ . By construction of  $\rho^{SB}(G)$ , we have that  $\beta^{SB} = 1/\mu^{SB}$ , obtaining (19). The notation  $\beta^i(G)$  is shorthand for  $\beta^i(V^{SB}(G))$ ,  $i = FB, SB$ .

## 6. Defining alternative cost of funds measures

In Section 3, we derived the result that  $(\mu/\lambda)^{FB}$  equals one. This is the familiar result that MCF is one in the first best case. In e.g. Bovenberg and van der Ploeg (1994a), (1994b), and Goulder (1995), the same kind of measure,  $(\mu/\lambda)$ , is referred to as MCF also in the second best case. Solving (13) for  $(\mu/\lambda)^{SB}$  we obtain

$$\text{MCF} = (\mu / \lambda)^{SB} = \frac{1}{1 + \frac{\partial l}{\partial t_L} \frac{t_L}{l} + \frac{\partial x_2}{\partial t_L} \frac{(t_2 - t_2^{PC})}{l}} \quad (20)$$

Other writers, e.g. Ballard and Medema (1993) and Brendemoen and Vennemo (1996), have defined MCF as the negative of the ratio between "change in consumer welfare" and "change in tax revenues" following from a tax increase, where "change in consumer welfare" is measured as the equivalent variation, i.e.,

$$\text{MCF} = \frac{-EV}{dR} \quad (21)$$

In the absence of negative externalities, it would be the case that  $-EV/dR = (\mu/\lambda)^{SB}$ , cf. Håkonsen (1997). With negative externalities, however, the perceived marginal utility of income,  $\lambda$ , is greater than the true marginal utility of income,  $\lambda - U_{Ee}'\partial x_2/\partial a$ . It follows that  $-EV/dR$  is greater than  $(\mu/\lambda)^{SB}$ , i.e. (for details, see Appendix 3),

$$-\frac{EV}{dR} = \frac{\mu}{\lambda - U_{Ee}'\frac{\partial x_2}{\partial a}} \quad (22)$$

Still another MCF-candidate may be derived from the maximum value functions developed in the previous Section. Since second best taxation is Pareto dominated by first best taxation, it seems reasonable to infer that the total cost of financing tax revenues is greater with second best taxation. This extra cost is due to the dead weight loss generated by the tax distortions. The total cost of public funds (TCF) thus consists of i) the actual tax revenue collected, and ii) the total dead weight loss,

$$\text{TCF}(G) = G + \text{TDWL}(G). \quad (23)$$

Since  $\text{TDWL}(G)$  equals  $\rho^{FB}(G) - G$ , we may alternatively write<sup>11</sup>  $\text{TCF}(G) = \rho^{FB}(G)$ . The corresponding average cost measure (ACF) is

$$\text{ACF}(G) = \frac{\text{TCF}(G)}{G} = \frac{R^{FB}(G)}{R^{SB}(G)}, \quad (24)$$

which equals  $(1 + \text{ADWL}(G))$ , cf. (18). Finally, the marginal cost measure corresponding to  $\text{TCF}(G)$  becomes<sup>12</sup>

$$\text{MCF}(G) = \text{TCF}'(G) = \frac{d\rho^{FB}(G)}{dG} = \frac{dR^{FB}(U)}{dU} \frac{dV^{SB}(G)}{dG} = \frac{\beta^{FB}(G)}{\beta^{SB}(G)}, \quad (25)$$

which from (19) equals  $(1 + \text{MDWL}(G))$ . This measure has quite intuitive properties. With no tax distortions,  $\text{MDWL}(G)$  is zero, and MCF is one. If tax distortions are generated, the marginal dead weight loss is added to the extra dollar of tax income in order to get the full marginal cost. Observe that  $\beta^{FB}/\beta^{SB}$  is independent of the choice of numeraire, while  $-EV/dR$  and  $(\mu/\lambda)^{SB}$  are not. Changing numeraire from  $P_1$  to  $W$  would imply a relatively large increase in  $-EV/dR$  and  $(\mu/\lambda)^{SB}$ , Håkonsen (1997).

By reference to Figure 2 we may sketch the following graphs of  $\text{ACF}(G)$  and  $\text{MCF}(G)$ :

---

<sup>11</sup> Let  $e(P, U)$  be the representative consumer's expenditure function. In the case *without* negative externalities,  $\text{TCF}(G)$  in (23) equals  $-EV = e(P^0, U^1) - e(P^1, U^1)$ , where  $U^1$  is the utility level  $V^{SB}(G)$ ,  $P^0$  are the pre-tax prices, and  $P^1$  are the post tax prices. The interpretation of  $-EV$  is that  $-EV$  measures the amount of lump-sum income which could have been taken away from a consumer facing the pre-tax prices  $P^0$  in order for the consumer to realise the utility level  $U^1 = V^{SB}(G)$ . Thus,  $-EV$  equals  $\rho^{FB}(G) = a^{FB}$ , where  $a^{FB}$  denotes the solution to  $\max R(a, t_L, t_2)$  s.t.  $V(a, t_L, t_2) = V^{SB}(G)$ , where the solutions for  $t_L$  and  $t_2$  are zero. With negative externalities present in the model, however, it is no longer the case that  $-EV = R^{FB}(G)$ , since the first-best optimal tax system no longer consists of only the lump sum tax. For further details concerning the equality between  $-EV$  and  $\rho^{FB}(G)$  in the no-externality case, see Håkonsen (1997).

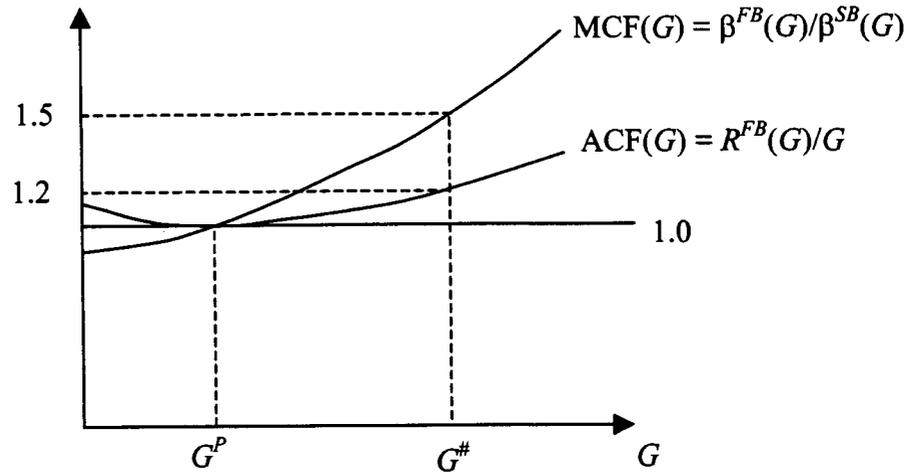


Figure 3.  $\beta^{FB}(G)/\beta^{SB}(G)$  and  $R^{FB}(G)/G$

These curves immediately remind us of  $AC(y)$  and  $MC(y)$  for a firm which minimises the cost of producing output  $y$ . In the Figure there is a stipulated level of public expenditures of  $G^\#$ , where  $ACF(G^\#) = 1.2$  and  $MCF(G^\#) = 1.5$ . In other words, the total dead weight loss (in units of tax revenue) is  $0.2G^\#$ , the average dead weight loss is 0.2 (20%) and the marginal dead weight loss is 0.5 (50%).

We have already established that  $-EV/dR > (\mu/\lambda)^{SB}$  in (22). As shown in Appendix 4, we further find that

$$\frac{\beta^{FB}}{\beta^{SB}} = \frac{1}{1 + \varepsilon + \frac{\partial x_2}{\partial t_L} \frac{(t_2 - t_2^{PC})}{l}} \frac{\lambda^{SB}}{\lambda^{FB}} = \left( \frac{\mu}{\lambda} \right)^{SB} \frac{\lambda^{SB}}{\lambda^{FB}}, \quad (26)$$

i.e.,  $(\mu/\lambda)^{SB}$  scaled by the factor  $(\lambda^{SB}/\lambda^{FB})$  yields  $\beta^{FB}/\beta^{SB}$ . We could alternatively cancel the term  $\lambda^{SB}$  and write  $\beta^{FB}/\beta^{SB}$  as  $\mu^{SB}/\lambda^{FB}$ , i.e., the multiplier  $\mu^{SB}$  in the second best solution to  $\max V(\cdot)$  s.t.  $R(\cdot) \geq G$  divided by the marginal utility of income at the consumer prices generated by the first best solution to  $\max R(\cdot)$  s.t.  $V(\cdot) \geq V^{SB}(G)$ . Since the first and second best optimal consumer prices are identical at  $G^P$ , we have that  $\lambda^{SB}(G^P) = \lambda^{FB}(G^P)$ , such that  $\beta^{FB}/\beta^{SB}$  and  $(\mu/\lambda)^{SB}$  coincide in the special case where  $G = G^P$ .

<sup>12</sup> The last step again uses the fact that  $\mu^{SB} = 1/\beta^{SB}$ .

## 7. Some illustrative examples

In this Section, we adopt a numerical version of the above model. Supplementing the theoretical constructs with explicit solutions and graphical illustrations will hopefully contribute to the understanding and interpretation of the results derived so far. We shall also illustrate some of the concepts *without* having negative externalities in the model. This will make it easier to interpret the results obtained *with* externalities in the model, and to see more clearly how the presence of externalities influence the tax optima, dead weight losses and the cost of public funds.

### 7.1. The numerical model

There are of course a number of parameters which need to be specified when developing a numerical version of the model. Since the main purpose of this section is to illustrate some of the above results, we will not go much into the consequences of alternative modelling assumptions.

We assume the following utility function,

$$U(E, F(x_0, H(x_1, x_2))), \quad (27)$$

where  $U(\cdot)$ ,  $F(\cdot)$ , and  $H(\cdot)$  are CES functions, with  $\sigma^U$ ,  $\sigma^F$ , and  $\sigma^H$  being the elasticities of substitution. With this structure, we know quite a lot about the optimal tax solutions before we actually compute them. First of all, since  $E$  is weakly separable<sup>13</sup> from  $F(\cdot)$ ,  $\partial x_i / \partial E = 0$  for  $i = 0, 1, 2$ . This has the implication that  $\tilde{U}_E$  is reduced into  $U_E$  in all above expressions involving  $\tilde{U}_E$ . Second, since  $x_0$  is assumed to be weakly separable from the homothetic aggregate  $G(x_1, x_2)$ , uniform commodity taxes would be optimal if negative externalities were not present in the model. With our choice of numeraire, uniform commodity taxation is implemented with a proportional tax on labour income alone. Thus, the assumed preference structure makes it especially easy to see how the optimal tax solutions are affected by the existence of negative externalities<sup>14</sup>.

<sup>13</sup> It would be interesting to study the implications of having  $E$  as an argument in the demands for  $x_0, x_1$ , and  $x_2$  also in the numerical examples. Due to recursive definitions of the model variables, however, ( $E$  is a function of  $x_2$ , which is a function of  $E$ , etc.) I have not succeeded in implementing this in the numerical simulations.

<sup>14</sup> The assumed structure implies that the first term on the right hand side of (14) is zero. The optimal tax rate on  $t_2$  thus equals the second term on the right hand side, i.e.,  $t_2 = \text{MED}(\lambda/\mu)$ , since  $\frac{-\tilde{U}_E e'}{U_1}$  reduces into  $\frac{-U_E e'}{U_1} = \text{MED}$ .

The remaining parameters to be specified are the time endowment  $T$ , the environmental damage function  $E = e(D)$ , and the sizes of the substitution parameters  $\sigma^U$ ,  $\sigma^F$ , and  $\sigma^H$  and the share parameters in the utility function. We choose  $T = 100$ , whereby 100 is the full endowment income if there are no taxes in the model. We further specify that  $\sigma^U$  and  $\sigma^H$  equal one (i.e., that  $U(\cdot)$  and  $H(\cdot)$  are Cobb-Douglas functions) and that  $\sigma^F$  equals 1.5. We shall also consider a case where  $\sigma^F < 1$ , and briefly comment on the consequences of alternative values for  $\sigma^H$ . The model is calibrated at the no-tax equilibrium such that  $x_0 = x_1 = x_2 = T/3 = 100/3$ . In other words, the consumer spends his full endowment income of 100 by consuming equal amounts of leisure and the two private consumption goods. Since  $\sigma^F$  is 1.5, the labour supply curve is upward sloping, with a wage elasticity of labour supply of 0.17 at the calibration point. Finally, we choose a linear environmental damage function,  $E = 40 - D$ , whereby the level of environmental quality at the calibration point is 7.77. The level  $E = 40$  may be thought of as the "pre-industrial level" which would be obtained in the absence of polluting activities. The top level aggregate is  $U = E^{0.05}F^{0.95}$ , which together with the choice of  $e(D)$  gives a reasonable balance between environmental concerns and the other arguments in the utility function.

## 7.2. First and second best tax optima.

We first illustrate the solutions in the first- and second best cases for four levels of the public expenditures,  $G = \{0, 8.2 = G^P, 20, 35\}$ . In the first best case the government may collect the full endowment income of  $T = 100$ . In the second best case, however, there is a global maximum of  $R(\cdot)$ ,  $R^{MAX} = 37.8$  at  $(t_L, t_2) = (0.79, 0)$ .

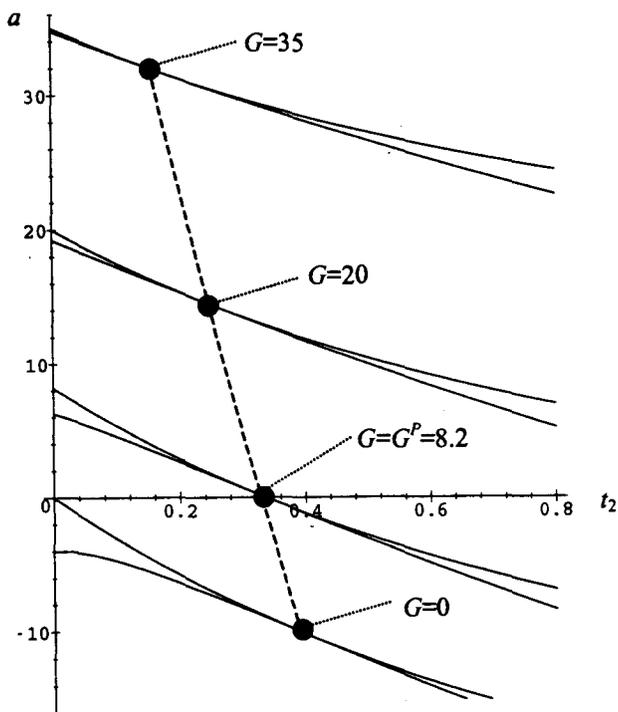


Figure 4a. First best optima

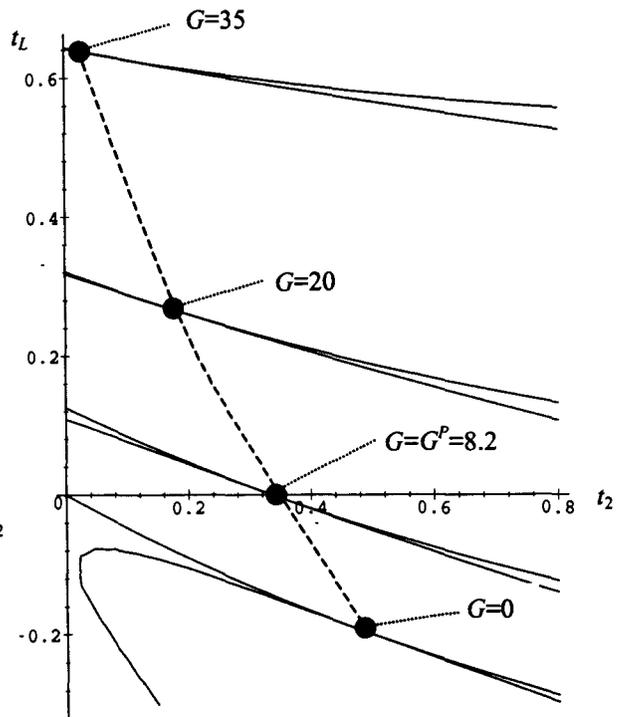


Figure 4b. Second best optima

As noted in sections 3 and 4, the first and second best solutions coincide at the tax revenue level raised by the optimal Pigouvian tax alone, which in our example is  $G^P = 8.2$ . The heavy dotted lines in the two figures are drawn through the family of tangency points between the contours of  $R(\cdot)$  and  $V(\cdot)$ , and resemble the output expansion path in a firm's cost minimisation problem. A suggestion for a name for such curves might be "tax revenue expansion paths".

By definition, all tax rates along the "first best tax revenue expansion path" are Pareto optimal. This fact serves well as a starting point for explaining why MCF is one in the first best case. In a first best solution in cases *without externalities*, the tax revenue is raised by means of the lump sum tax alone. The "first best tax revenue expansion path" would then simply be the  $a$ -axis in figure 4a. For all levels of  $G$ , the optimal tax is  $a = G$ , and all such allocations are Pareto optimal. Different levels of  $G$  imply different income distributions between the government and the consumer, but have no consequences for efficiency.

In the case *with externalities*, there is one important difference: Points along the  $a$ -axis are no longer Pareto optimal. For a market economy to support an allocation satisfying (3), we need the price correction  $t_2 = \text{MED}$ . As long as  $t_2 = \text{MED}$  for all levels of  $G$  (i.e., as long as we stay on the tax revenue expansion path), more or less tax revenue may be

raised without affecting efficiency. For exactly this reason, MCF equals one ( $\mu$  equals  $\lambda$ ) in the first best case both with and without negative externalities present in the model. We notice that the optimal Pigouvian tax rate is declining in  $G$ . This is due to the fact that as  $a$  increases, the resulting negative income effect on  $x_2$  contributes to improved environmental quality, making environmental quality less scarce compared to private consumption, such that MED and  $t_2$  are reduced.

In figure 4b, the distortionary labour income tax replaces the lump sum tax. Therefore, all points along the "second best tax revenue expansion path" are Pareto dominated by first best taxation except at the tax revenue level  $G^P$ . As we increase  $G$  in Figure 4b, the tax optima move upwards and to the left, i.e.,  $t_L$  increases and  $t_2$  decreases. When we approach the global maximum for the tax revenue function,  $t_2$  converges towards zero, since the labour income tax represents the most efficient revenue raising instrument<sup>15</sup>, dominating the price corrective tax.

**7.3. Maximum value functions, dead weight losses, and cost of public funds**

We start by computing the maximum tax revenue functions  $\rho^{FB}(G)$  and  $\rho^{SB}(G)$ . Since there is a global maximum for second best financed tax revenues,  $R^{MAX}$ , the functions  $\rho^{FB}(G)$  and  $\rho^{SB}(G)$  are only defined for  $G \leq R^{MAX} = 37.8$ .

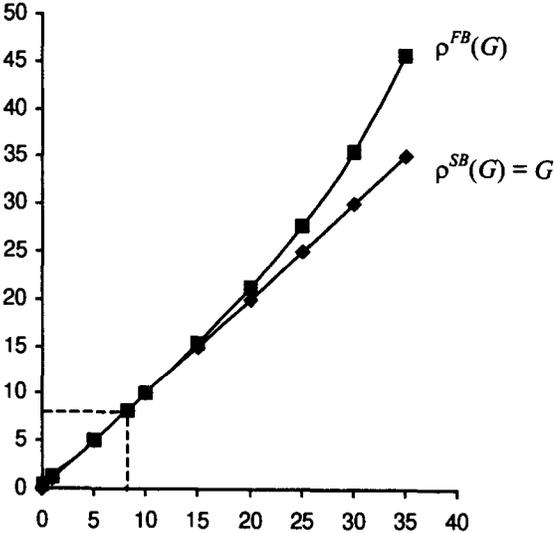


Figure 5.  $\rho^{SB}(G)$  and  $\rho^{FB}(G)$

<sup>15</sup> If there were no externalities in the model, the assumed preferences implies that the second best tax revenue expansion path would be the  $t_L$ -axis (cf. section 7.5).

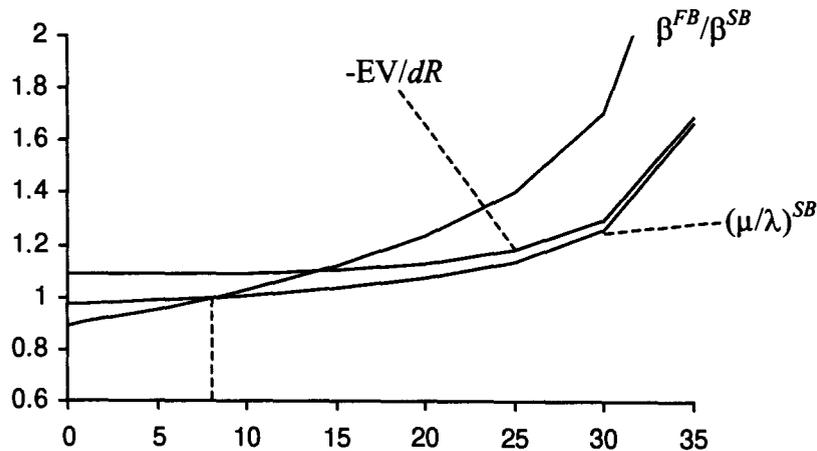
The point of tangency between  $\rho^{FB}$  and  $\rho^{SB}$  (indicated by the dotted lines) is at  $G^P = 8.2$ . The total, average and marginal dead weight losses in this example are seen in Table 1.

**Table 1. Total, average and marginal dead weight losses**

G	0	5	$G^P=8.2$	10	15	20	25	30	35
$TDWL(G)=R^{FB}(G)-R^{SB}(G)$	0.46	0.07	0.00	0.03	0.38	1.25	2.81	5.48	10.62
$ADWL(G)=TDWL(G)/G$	$+\infty$	0.01	0.00	0.00	0.03	0.06	0.11	0.18	0.30
$MDWL(G)=TDWL(G)$	-0.11	-0.05	0.00	0.03	0.12	0.23	0.40	0.70	1.58

At  $G = 35$ , we are relatively close to the maximum tax income  $R^{MAX}$ , whereby the marginal dead weight loss is as high as 1.58. The marginal dead weight loss rises towards infinity as we approach  $R^{MAX}$  from below.

Finally we plot the three alternative MCF definitions  $(\mu/\lambda)^{SB}$ ,  $-EV/dR$ , and  $\beta^{FB}/\beta^{SB}$ , where we recall that the latter equals one plus  $MDWL(G)$ .



**Figure 6.  $(\mu/\lambda)^{SB}$  and  $\beta^{FB}/\beta^{SB}$**

$(\mu/\lambda)^{SB}$  and  $\beta^{FB}/\beta^{SB}$  are both less than one for  $G < G^P$ , and higher than one for  $G > G^P$ . As noted previously, the conversion factor (cf. eq. (26)) is the ratio of marginal utilities of income at the first and second best optimal prices respectively, i.e.,  $(\beta^{FB}/\beta^{SB}) = (\mu/\lambda)^{SB}(\lambda^{SB}/\lambda^{FB})$ . For  $G > (<) G^P$  the income tax is positive (negative), making  $\lambda^{SB} > (<) \lambda^{FB}$ , explaining the pattern shown in Figure 6.

The difference between  $-EV/dR$  and  $(\mu/\lambda)^{SB}$  is reduced as  $G$  increases. This is due to the fact that the term  $U_E$  (cf. eq. (22)) declines as more and more taxes are collected, since the environmental quality rises as the consumption of the polluting good  $x_2$  diminishes.

While the three measures take distinctly different numerical values, all of them are strictly increasing in  $G$ , which seems like a natural property for MCF measures. As will be shown in the next section, however, this development is not a general one;  $-EV/dR$  and  $(\mu/\lambda)^{SB}$  may decrease while  $\beta^{FB}/\beta^{SB}$  increases.

#### 7.4. A case with negative labour supply elasticity

So far, we have assumed that  $\sigma^F$  (the elasticity of substitution between leisure and the private consumption goods) is 1.5, which implies an upward sloping labour supply curve. Since the tax wedge between the consumer price of leisure and private consumption goods is a fundamental source for dead weight losses in our model, it is important to have some idea about the effects of alternative choices of  $\sigma^F$ . Let us therefore reduce  $\sigma^F$  to 0.75, which gives a wage elasticity of labour supply of -0.08 at the calibration point, i.e., a downward sloping labour supply schedule.

There is no point in repeating all previous illustrations for this scenario. Consider first the total and marginal dead weight losses in Table 2.

**Table 2. Total and marginal dead weight losses in the case where  $\sigma^F = 0.75$ .**

$G$	0	$G^P=8.8$	20	30	40	50	60	70	80
$TDWL(G)=R^{FB}(G)-R^{SB}(G)$	0.25	0.00	0.40	1.42	3.05	5.3	7.96	10.82	12.56
$MDWL(G)=TDWL'(G)$	-0.06	0.00	0.07	0.13	0.19	0.25	0.29	0.27	0.01

As one should expect, the dead weight loss is reduced when the elasticity of substitution between leisure and consumption goods becomes smaller. For example, while the total dead weight loss at  $G = 30$  was nearly 5.5 with  $\sigma^F = 1.5$ , it is now only 1.42. We also observe that the first best Pigou tax generates a higher tax revenue;  $G^P$  rises from 8.2 to 8.8. This is due to the tax base becoming less elastic when the degree of substitutability towards the untaxed good (leisure) is reduced. However, the most significant change from the previous scenario is that there no longer is an upper bound on  $G$ ; when  $\sigma^F < 1$ , the full endowment income of  $T = 100$  may be collected also in the second best case<sup>16</sup>. Consequently, the total dead weight loss,  $TDWL(G)$ , is no longer a monotonically growing and convex function of  $G$  for  $G > G^P$ .

<sup>16</sup> As  $G$  increases, the contours of  $R(\cdot)$  and  $V(\cdot)$  become flatter and flatter. The numerical search for tangency points between contours of  $R(\cdot)$  and  $V(\cdot)$  (cf. figure 4) thus becomes increasingly difficult. Although there exist tax rates such that  $R(\cdot)$  is greater than 80, the computer program has not been able to solve the tax optima for  $G > 80$ . See the next section, however, where optimal tax solutions are computed up to  $G = 99$  in the case where the negative externality no longer exists.

Rather,  $TDWL(G)$  follows a convex growth pattern up to  $G \approx 60$ , where it has an inflection point, cf. Figure 7.

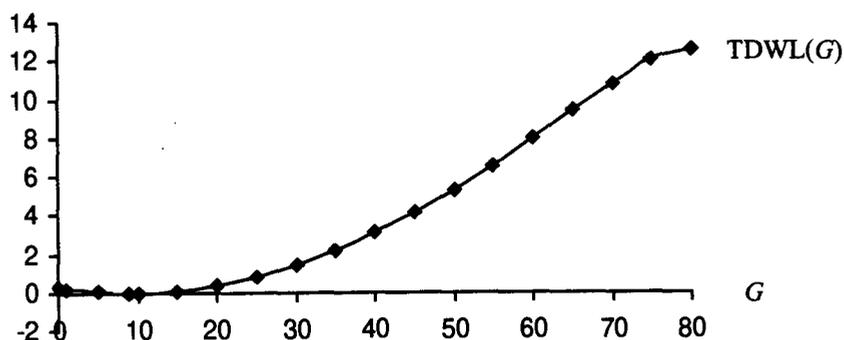


Figure 7. The total dead weight loss when  $\sigma^F = 0.75$ .

Since the marginal dead weight loss,  $MDWL(G)$ , is the derivative of  $TDWL(G)$ , the maximum for  $MDWL(G)$  is at the inflection point of  $TDWL(G)$  at  $G \approx 60$ . At  $G = 80$ ,  $TDWL$  is nearly flat, whereby  $MDWL$  is close to zero.<sup>17</sup>

Let us turn our attention to the alternative MCF concepts  $(\mu/\lambda)^{SB}$ ,  $-EV/dR$ , and  $\beta^{FB}/\beta^{SB}$ .

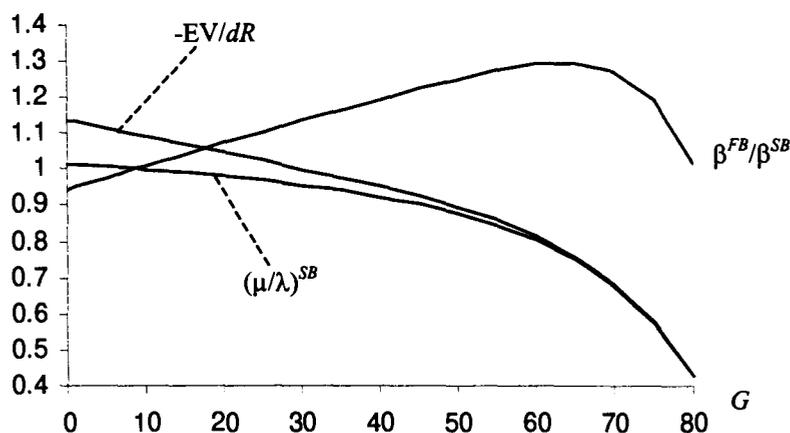


Figure 8.  $(\mu/\lambda)^{SB}$  and  $\beta^{FB}/\beta^{SB}$  when  $\sigma^F = 0.75$ .

The graph of  $\beta^{FB}/\beta^{SB}$  is in principle already known, since it only adds one to the previously shown marginal dead weight loss in Table 2. As in the previously seen case in Figure 6, the difference between  $(\mu/\lambda)^{SB}$  and  $-EV/dR$  becomes smaller as the tax revenue and the level of environmental quality increase, whereby  $U^E$  in (22) decreases.

<sup>17</sup> If we had been able to compute the optimal tax solutions also for the region  $G \in (80,100)$ ,  $TDWL(G)$  would be declining and  $MDWL(G)$  would be negative (see the next subsection for such a case).

What makes this example particularly interesting is the fact that  $(\mu/\lambda)^{SB}$  and  $-EV/dR$  decrease while  $\beta^{FB}/\beta^{SB}$  increases (for  $G < 60$ ). If we choose to interpret  $(\mu/\lambda)^{SB}$  or  $-EV/dR$  as MCF, it would thus be the case that MCF and the marginal dead weight loss are inversely related. The reason for this somewhat paradoxical result is seen by solving (26) for  $(\mu/\lambda)^{SB}$ , i.e.,  $(\mu/\lambda)^{SB} = (\beta^{FB}/\beta^{SB})(\lambda^{FB}/\lambda^{SB})$ , and investigating the two elements  $(\beta^{FB}/\beta^{SB})$  and  $(\lambda^{FB}/\lambda^{SB})$  separately. We have already seen that  $(\beta^{FB}/\beta^{SB})$  is rising in  $G$  up to  $G \approx 60$ . On the other hand, the second best optimal tax rates  $\{t_L^{SB}(G), t_2^{SB}(G)\}$  are characterised by  $t_L^{SB}(G)$  becoming higher and  $t_2^{SB}(G)$  becoming smaller as  $G$  increases.<sup>18</sup> Ceteris paribus, this makes both leisure consumption and commodity 2 cheaper and the marginal utility of income,  $\lambda^{SB}$ , higher as  $G$  rises. The ratio  $(\lambda^{FB}/\lambda^{SB})$  thus declines as  $G$  rises, with the per cent decline in  $(\lambda^{FB}/\lambda^{SB})$  being higher than the per cent growth in  $(\beta^{FB}/\beta^{SB})$ .

### 7.5. Externalities vs. no externalities.

The main question in this paper has been how the presence of negative externalities affects the dead weight loss and marginal cost of public funds. It is therefore interesting to compare some of the above results also for the case *without* negative externalities present in the model. To perform such a comparison, let us assume that a costless technological breakthrough produces a new abatement technology, such that the consumption of commodity 2 no longer deteriorates the environment. The no-externality case is analysed by specifying that the level of environmental quality is fixed at the no emission level. In the numerical model this amounts to having  $E = 40$ . All other parameters are retained from the previous sections. In this subsection we focus on the measure  $\beta^{FB}/\beta^{SB}$ , referred to as MCF, leaving  $(\mu/\lambda)^{SB}$  and  $-EV/dR$  aside.

Consider first the case where  $\sigma = 1.5$ . Figure 9 shows the development of  $\rho^{FB}(G)$  and  $\rho^{SB}(G)$  for the case with and without negative externalities, respectively.

---

<sup>18</sup> See figure 4b for the case where  $\sigma = 1.5$ . The development of  $t_L$  and  $t_2$  is qualitatively similar when  $\sigma = 0.75$ .

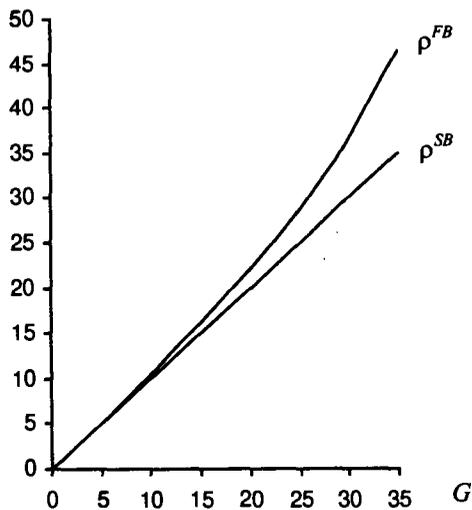


Figure 9a.  $\rho^{FB}$  and  $\rho^{SB}$  without externalities.

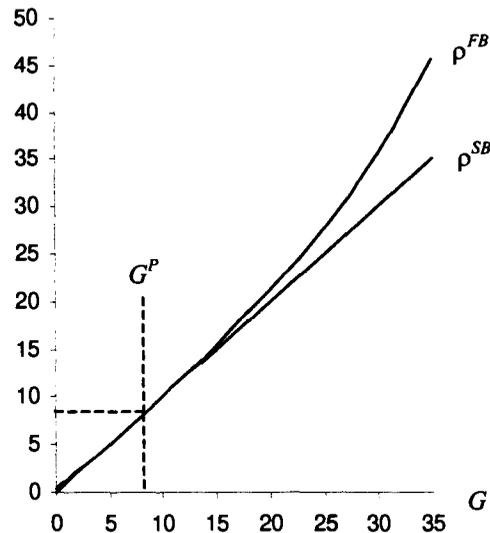


Figure 9a.  $\rho^{FB}$  and  $\rho^{SB}$  with externalities.

Without externalities, the only point in common for  $\rho^{FB}$  and  $\rho^{SB}$  is  $G = 0$ . Consequently, dead weight losses are generated from the first unit of tax revenue collected (although very small for low levels of tax revenue). The existence of negative externalities implies that the tax revenue level  $G^P$  (8.2 in this example) may be collected without a dead weight loss. For low levels of tax revenues this significantly reduces the dead weight loss. For example, at  $G = 20$ , the total dead weight losses are 2.1 and 1.2 without and with externalities, respectively. As we approach the maximum available tax revenue, however, the dead weight losses converge for the two cases.

These patterns for  $\rho^{FB}$  and  $\rho^{SB}$  translate into the following graphs for MCF as defined by  $\beta^{FB}/\beta^{SB}$ , where the upper curve shows the no-externality case.

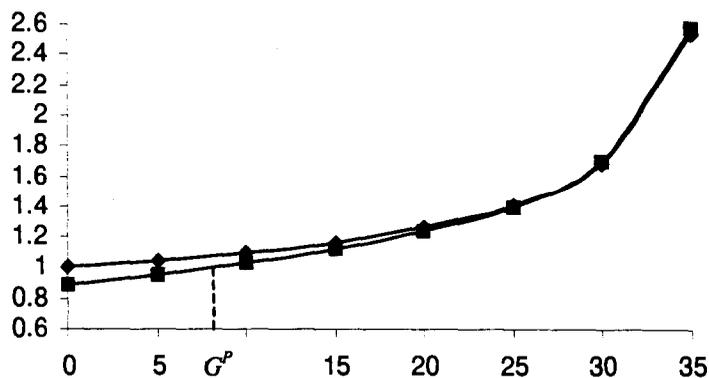
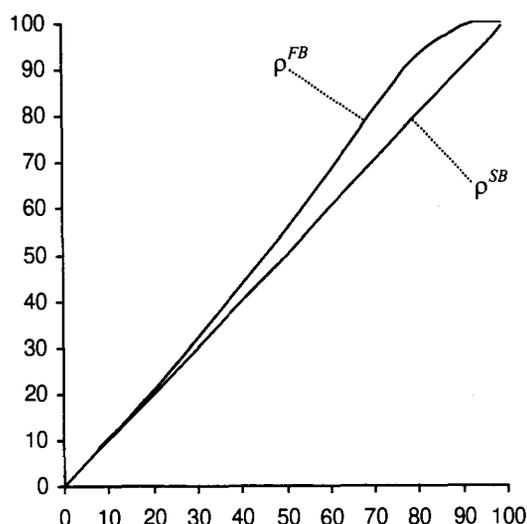


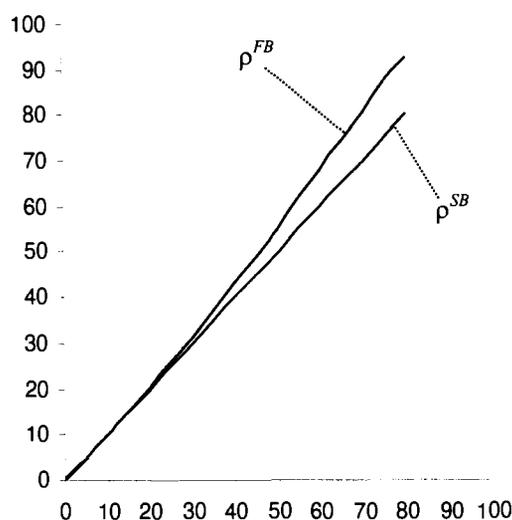
Figure 10.  $\beta^{FB}/\beta^{SB}$  with and without negative externalities.

In the case without the negative externality,  $\beta^{FB}/\beta^{SB}$  starts at 1.0. When there are negative externalities,  $\beta^{FB}/\beta^{SB}$  starts below one and passes this level at  $G^P$ . As  $G$  increases,  $\beta^{FB}/\beta^{SB}$  is less and less influenced by the negative externality.

Finally, we show the equivalents of Figures 9 and 10 for the case where the elasticity of substitution between leisure and private consumption goods,  $\sigma^F$ , is 0.75.



**Figure 11a.**  $\rho^{FB}$  and  $\rho^{SB}$  without externalities when  $\sigma^F = 0.75$ .



**Figure 11b.**  $\rho^{FB}$  and  $\rho^{SB}$  with externalities when  $\sigma^F = 0.75$ .

The principles seen in Figures 9a and b are retained also in this case. The central case here is figure 11a, where optimal tax solutions have been computed also for the interval<sup>19</sup>  $G \in (80,99]$ . Quite interestingly, the dead weight loss converges towards zero at both ends of the interval  $G \in [0, 100]$ . This simply reflects that  $G = 100$  is the maximum theoretical level of tax revenues irrespective of whether there are tax distortions or not. The corresponding graphs for MCF are shown in figure 12, where the no-externality case again is the upper curve.

<sup>19</sup> In the case without negative externalities, the tax optima are computable also for these extreme values. Recall that the time endowment  $T$  is 100, whereby the values for  $G$  may be read as the public sector's resource use in per cent of the economy's full endowment income.

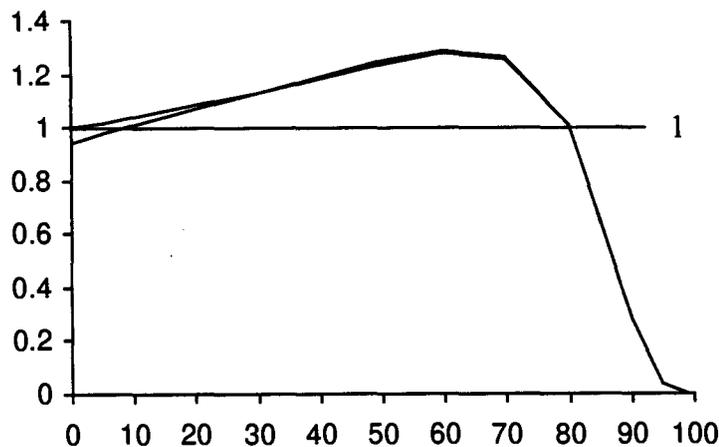


Figure 12.  $\beta^{FB}/\beta^{SB}$  with and without negative externalities when  $\sigma^F = 0.75$ .

Again we see that the negative externality significantly only affects  $\beta^{FB}/\beta^{SB}$  at relatively low levels of tax revenues.  $\beta^{FB}/\beta^{SB}$  stay above one as long as the total dead weight loss is an increasing function of  $G$ , and converges towards zero<sup>20</sup> as  $G$  converges towards the resource limit  $T = 100$ .

An obvious objection to the above figures is that what happens to the optimal tax solutions and the dead weight losses at extremely high levels of public sector resource use is of little interest for real-life tax policy analysis. Despite this, there are interesting phenomena which occur at relatively modest tax revenue levels. Without having seen how the dead weight loss develops over the full range of theoretically possible values of  $G$ , the pattern shown in Figure 12 could be hard to explain. For example, one would perhaps imagine that the marginal dead weight loss is a convex and monotonically increasing function of  $G$ .

In the example in Figure 12, the point where the marginal dead weight loss is at the highest is at the inflection point for the total dead weight loss, i.e., at  $G \approx 60$  (cf. Figure 7). While this level of public sector resource usage still is unrealistically high, the point where the marginal dead weight loss is no longer a convex function of  $G$  (the inflection point for  $MDWL(G)$ ) is at  $G \approx 40$ . This level might be sufficiently low to be of some practical interest. If one started out measuring the development of the marginal dead weight loss, and found a point where the rate of growth in the MDWL started to decline, it would perhaps be hard to explain such a pattern if one did not know the principles derived by studying the development at extreme levels.

<sup>20</sup> The values for  $\beta^{FB}/\beta^{SB}$  at  $G = 90$  and  $99$  have only been computed in the no-externality case.

### 7.6. Initial tax rates outside the optimum.

All conclusions so far have rested on the assumption that tax revenues are optimally financed. Allowing for an initial tax system outside the optimum certainly brings other results. For example, suppose that the government initially consumes twenty units of commodity  $G$ , financed by a labour income tax only. In the case where  $\sigma^F = 1.5$ , the necessary income tax rate would be  $t_L = 0.32$ . The effect of being outside the optimum must by definition be that one could i) make the consumer better off while maintaining the tax revenue of 20, ii) increase the tax revenues while maintaining the utility level for the consumer, or iii) increase both the utility level and the tax revenues. Figure 13 illustrates these possibilities.

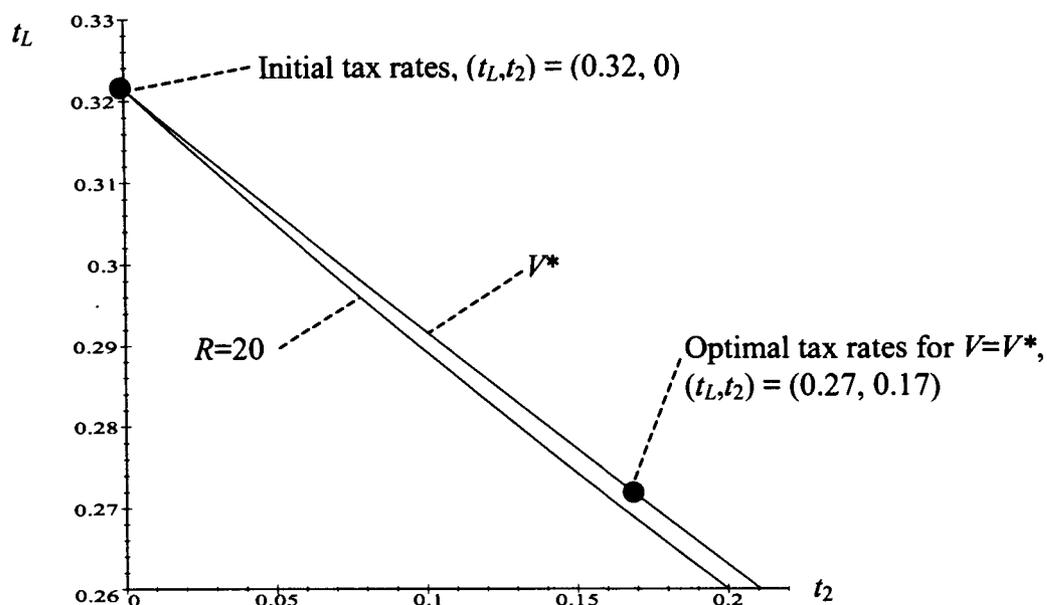


Figure 13. Initial tax rates outside the optimum

Following the contour  $V(t_L, t_2) = V^*$  downwards and to the left increases  $R(t_L, t_2)$  until the point  $(t_L, t_2) = (0.27, 0.17)$  is reached, where  $R = 20.16$ . In other words, the cost associated with generating the 0.16 extra units of tax revenues is zero. Moving somewhere in between  $R=20$  and  $V^*$  would both give higher tax revenues and higher utility level, whereby the cost of funds (however defined) would be negative. Such possibilities are not dependent of the existence of externalities, however. The same points could be made if starting outside the optimum in a case without externalities.

In order to provide a closer comparison to the results of Ballard and Medema (1993) and Brendemoen and Vennemo (1996) (in which the ratio  $-EV/dR$  is computed), consider Figure 14, in which there are four alternative combinations of tax rates. Point a) represents the benchmark, where  $R = 20$  and the tax rates are  $t_L = 0.321$  and  $t_2 = 0$ . We study three alternative tax reforms from point a), which all increase the tax revenues to  $R = 20.1$ . Point b) is obtained by maintaining  $t_2 = 0$ , while increasing  $t_L$  to 0.323, point c) by maintaining  $t_L = 0.321$  while increasing  $t_2$  to 0.005, and d) by increasing  $t_2$  to 0.05 and reducing  $t_L$  to 0.31.

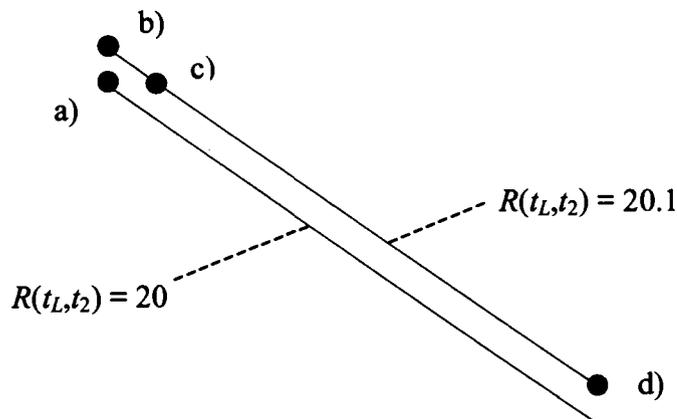


Figure 14. Three alternative tax reforms.

The corresponding values for the ratio  $-EV/dR$  are seen in Table 3.

Table 3.  $-EV/dR$  for three alternative tax reforms.

From point a) to:	b)	c)	d)
$-EV/dR$	1.04	0.93	0.17

These results are qualitatively similar to the balanced budget experiments in Ballard and Medema<sup>21</sup>: the ratio  $-EV/dR$  is above one when increasing the tax rates on non-externality producing goods and below one when increasing the Pigouvian tax. In our model, we find that  $-EV/dR$  is close to zero when going to point d) in Figure 14. If going further to the right and downwards along  $R = 20.1$ , we enter into the region in Figure 13 where  $V(t_L, t_2) > V^*$ , whereby  $-EV/dR$  becomes negative. A jump from point a) to the optimum along  $R = 20.1$ ,  $(t_L, t_2) = (0.27, 0.17)$ , yields  $-EV/dR = -0.65$ , i.e., the government collects 0.1 extra units of tax revenue at the same time as the equivalent variation for the household is positive,  $EV = 0.065$ .

<sup>21</sup> Brendemoen and Vennemo find negative values for  $-EV/dR$  when increasing the tax rates on gasoline, oil, and CO<sub>2</sub>. This is mainly due to very large environmental benefits, which are subtracted from the "traditional" MCF computed without the environmental effects.

## **8. Concluding comments**

In a first best economy, the presence of negative externalities does affect the dead weight losses and the marginal cost of public funds. Since any amount of tax revenues may be raised without creating dead weight losses, alternative levels of tax revenues only mean different distributions of income between the private and public sector. The marginal cost of public funds in a first best economy therefore equals one both with and without negative externalities.

In a second best economy, negative externalities imply that a positive amount of tax revenue may be raised without generating a dead weight loss even though a lump sum tax is not available. This amount is the tax revenue generated by a optimal Pigouvian tax alone. At relatively modest levels of tax revenues, MCF is thus considerably lower than in the no-externality case. As the amount of tax revenue increases, however, the share of tax revenues raised by the corrective tax on polluting goods becomes smaller and smaller relative to the revenues raised by other tax instruments. The MCF therefore converges towards the level in the no-externality case as the amount of tax revenues increases.

These conclusions apply to the case where the tax revenues are collected optimally. If the tax system starts outside the second best optimum, however, it is by definition possible to achieve Pareto improvements. One such improvement is to keep the utility of the consumer constant while increasing the tax revenues until a second best optimum is reached. The cost associated with the extra tax revenues from such an operation would be zero. To the extent that the tax authorities have under-taxed polluting activities, there could be cases where the cost of raising extra tax revenues by increasing tax rates on polluting commodities is substantially smaller than for other tax instruments. It is important to bear in mind that such possibilities are not due to the existence of externalities as such. Rather, they are only reflections of the initial tax system being outside the optimum. Indeed, Ballard and Medema's (1993) and Brendemoen and Vennemo's (1996) results obtained in CGE models for the US and Norwegian economies, respectively – that including negative externalities in the analysis brings the marginal cost of funds below unity – indicate that polluting commodities have been under-taxed in these economies.

## Appendix 1. Derivation of the first-order conditions (7)-(9).

We first write the partial of  $L$  w.r.t.  $a$  as

$$\frac{\partial L}{\partial a} = -\lambda + U_E \frac{dE}{da} + \mu \left[ 1 + t_L \frac{dl}{da} + t_2 \frac{dx_2}{da} \right] = 0. \quad (\text{a1.1})$$

Since  $E = e(x_2(t_L, t_2, E(\cdot)))$ , we have that

$$\frac{dE}{da} = e' \left( \frac{\partial x_2}{\partial a} + \frac{\partial x_2}{\partial E} \frac{dE}{da} \right), \quad (\text{a1.2})$$

which (provided that the stability condition  $\partial x_2 / \partial E < 1$  holds) gives

$$\frac{dE}{da} = \frac{e' \frac{\partial x_2}{\partial a}}{1 - e' \frac{\partial x_2}{\partial E}}. \quad (\text{a1.3})$$

The terms  $dl/dE$  and  $dx_2/dE$  in (a1.1) are

$$\frac{dl}{da} = \frac{\partial l}{\partial a} + \frac{\partial l}{\partial E} \frac{dE}{da} \quad \text{and} \quad \frac{dx_2}{da} = \frac{\partial x_2}{\partial a} + \frac{\partial x_2}{\partial E} \frac{dE}{da}, \quad (\text{a1.4})$$

which from (a1.3) becomes

$$\frac{dl}{da} = \frac{\partial l}{\partial a} + \frac{\partial l}{\partial E} \frac{e' \frac{\partial x_2}{\partial a}}{1 - e' \frac{\partial x_2}{\partial E}} \quad \text{and} \quad \frac{dx_2}{da} = \frac{\partial x_2}{\partial a} + \frac{\partial x_2}{\partial E} \frac{e' \frac{\partial x_2}{\partial a}}{1 - e' \frac{\partial x_2}{\partial E}}. \quad (\text{a1.5})$$

Inserting (a1.3) and (a1.5) into (a1.1) yields

$$\frac{\partial L}{\partial a} = -\lambda + U_E \frac{e' \frac{\partial x_2}{\partial a}}{1 - e' \frac{\partial x_2}{\partial E}} + \mu \left[ 1 + t_L \left( \frac{\partial l}{\partial a} + \frac{\partial l}{\partial E} \frac{e' \frac{\partial x_2}{\partial a}}{1 - e' \frac{\partial x_2}{\partial E}} \right) + t_2 \left( \frac{\partial x_2}{\partial a} + \frac{\partial x_2}{\partial E} \frac{e' \frac{\partial x_2}{\partial a}}{1 - e' \frac{\partial x_2}{\partial E}} \right) \right] = 0, \quad (\text{a1.6})$$

which using (11) yields (7). The derivations of (8) and (9) follow by analogy.

## Appendix 2. Derivation of (12)

Inserting the Slutsky equation in (8) and (9) and rewriting (7)-(9) by collecting all compensated derivatives (with superscripts C) on the left hand side yields

$$0 = \lambda - \mu \left[ t_L \frac{\partial l}{\partial a} + t_2 \frac{\partial x_2}{\partial a} \right] - \tilde{U}_E e' \frac{\partial x_2}{\partial a}, \quad (\text{a2.1})$$

$$\tilde{U}_E e' \frac{\partial x_2^C}{\partial t_L} + \mu \left[ t_L \frac{\partial l^C}{\partial t_L} + t_2 \frac{\partial x_2^C}{\partial t_L} \right] = l \left( \lambda - \mu \left[ t_L \frac{\partial l}{\partial a} + t_2 \frac{\partial x_2}{\partial a} \right] - \tilde{U}_E e' \frac{\partial x_2}{\partial a} \right), \quad (\text{a2.2})$$

$$\tilde{U}_E e' \frac{\partial x_2^C}{\partial t_2} + \mu \left[ t_L \frac{\partial l^C}{\partial t_2} + t_2 \frac{\partial x_2^C}{\partial t_2} \right] = x_2 \left( \lambda - \mu \left[ t_L \frac{\partial l}{\partial a} + t_2 \frac{\partial x_2}{\partial a} \right] - \tilde{U}_E e' \frac{\partial x_2}{\partial a} \right), \quad (\text{a2.3})$$

whereby (since (a2.1) holds) the system reduces into

$$\left( \frac{\tilde{U}_E e'}{\mu} + t_2 \right) \frac{\partial x_2}{\partial a} + t_L \frac{\partial l}{\partial a} = \frac{\lambda}{\mu} - 1, \quad (\text{a2.4})$$

$$\left( \frac{\tilde{U}_E e'}{\mu} + t_2 \right) \frac{\partial x_2^C}{\partial t_L} + t_L \frac{\partial l^C}{\partial t_L} = 0, \quad (\text{a2.5})$$

$$\left( \frac{\tilde{U}_E e'}{\mu} + t_2 \right) \frac{\partial x_2^C}{\partial t_2} + t_L \frac{\partial l^C}{\partial t_2} = 0. \quad (\text{a2.6})$$

The system (a2.4)-(a2.6) is satisfied for  $t_L = 0$  and  $t_2 = -\frac{\tilde{U}_E e'}{\mu}$ , which from (a2.4) implies that

$\mu = \lambda$ , which again gives  $t_2 = -\frac{\tilde{U}_E e'}{\lambda}$ . Inserting  $t_2 = \text{MED} = -U_E e'/\lambda$  into  $\tilde{U}_E$  in (11), we find

that  $\tilde{U}_E$  reduces into  $U_E$  (i.e., the social marginal utility of the environment reduces into the direct marginal utility of the environment). Finally, using (10), we obtain the solution for  $a = G^0 - t_2 x_2$ , whereby (7)-(10) implies the pattern shown in (12).

### Appendix 3. Derivation of (22).

Consider a second best optimally financed marginal increase in  $G$ . From the public sector's budget constraint we must have

$$dG = dR = \frac{\partial R}{\partial t_L} dt_L + \frac{\partial R}{\partial t_2} dt_2. \quad (\text{a3.1})$$

The impact on utility from this increase in tax revenues is

$$dV = \frac{\partial V}{\partial t_L} dt_L + \frac{\partial V}{\partial t_2} dt_2 = \left( -\lambda I + U_E e' \frac{\partial x_2}{\partial t_L} \right) dt_L + \left( -\lambda x_2 + U_E e' \frac{\partial x_2}{\partial t_2} \right) dt_2, \quad (\text{a3.2})$$

which from the first-order conditions (8) and (9) becomes

$$dV = -\mu \left( \frac{\partial R}{\partial t_L} dt_L + \frac{\partial R}{\partial t_2} dt_2 \right) = -\mu dR. \quad (\text{a3.3})$$

The effect on utility from a marginal increase in lump-sum income,  $dI = -da$ , is

$$dV = \frac{\partial V}{\partial I} dI = \left( \lambda + U_E e' \frac{\partial x_2}{\partial I} \right) dI. \quad (\text{a3.4})$$

The equivalent variation (EV) is defined as the change in lump-sum income,  $dI = EV$ , which at the pre-tax-increase tax rates  $t_L^0, t_2^0$  gives the same change in utility as in (a3.3), i.e.,

$$\left( \lambda + U_E e' \frac{\partial x_2}{\partial I} \right) EV = -\mu dR, \quad (\text{a3.5})$$

whereby the ratio  $-EV/dR$  becomes

$$\frac{-EV}{dR} = \frac{\mu}{\lambda + U_E e' \frac{\partial x_2}{\partial I}} = \frac{\mu}{\lambda - U_E e' \frac{\partial x_2}{\partial a}}, \quad (\text{a3.6})$$

cf. eq. (22).

#### Appendix 4. Derivation of (26).

The first order conditions associated with the Lagrangean

$$\Lambda = R(t_L, t_2, a) + \beta [V(t_L, t_2, a) - U] \quad (\text{a4.1})$$

are (cf. Appendix 1)

$$\frac{\partial \Lambda}{\partial a} = 1 + t_L \frac{\partial l}{\partial a} + t_2 \frac{\partial x_2}{\partial a} - \lambda \beta + \tilde{R}_E e^{\frac{\partial x_2}{\partial a}} = 0, \quad (\text{a4.2})$$

$$\frac{\partial \Lambda}{\partial t_L} = l + t_L \frac{\partial l}{\partial t_L} + t_2 \frac{\partial x_2}{\partial t_L} - \lambda \beta l + \tilde{R}_E e^{\frac{\partial x_2}{\partial t_L}} = 0, \quad (\text{a4.3})$$

$$\frac{\partial \Lambda}{\partial t_2} = x_2 + t_L \frac{\partial l}{\partial t_2} + t_2 \frac{\partial x_2}{\partial t_2} - \lambda \beta x_2 + \tilde{R}_E e^{\frac{\partial x_2}{\partial t_2}} = 0, \text{ and} \quad (\text{a4.4})$$

$$\frac{\partial \Lambda}{\partial \beta} = V(t_L, t_2, a) - U = 0, \quad (\text{a4.5})$$

where  $\tilde{R}_E$  is the direct and indirect effects on tax revenues of a marginal increase in  $E$ , i.e.,

$$\tilde{R}_E = \frac{\beta U_E + t_L \frac{\partial l}{\partial E} + t_2 \frac{\partial x_2}{\partial E}}{1 - e^{\frac{\partial x_2}{\partial E}}}, \quad (\text{a4.6})$$

which becomes the dual counterpart to  $\tilde{U}_E$ . Following the same procedure as in Appendix 2, we substitute for the Slutsky equation and collect all compensated terms in the left hand side, obtaining

$$0 = -1 + \beta \lambda - t_L \frac{\partial l}{\partial a} - t_2 \frac{\partial x_2}{\partial a} - \tilde{R}_E e^{\frac{\partial x_2}{\partial a}}, \quad (\text{a4.7})$$

$$\tilde{R}_E e^{\frac{\partial x_2^C}{\partial t_L}} + t_L \frac{\partial l^C}{\partial t_L} + t_2 \frac{\partial x_2^C}{\partial t_L} = l \left( -1 + \beta \lambda - t_L \frac{\partial l}{\partial a} - t_2 \frac{\partial x_2}{\partial a} - \tilde{R}_E e^{\frac{\partial x_2}{\partial a}} \right), \quad (\text{a4.8})$$

$$\tilde{R}_E e^{\frac{\partial x_2^C}{\partial t_2}} + t_L \frac{\partial l^C}{\partial t_2} + t_2 \frac{\partial x_2^C}{\partial t_2} = x_2 \left( -1 + \beta \lambda - t_L \frac{\partial l}{\partial a} - t_2 \frac{\partial x_2}{\partial a} - \tilde{R}_E e^{\frac{\partial x_2}{\partial a}} \right), \quad (\text{a4.9})$$

which, using (a4.7), becomes

$$(\tilde{R}_E e' + t_2) \frac{\partial x_2}{\partial a} + t_L \frac{\partial l}{\partial a} = \beta \lambda - 1, \quad (\text{a4.10})$$

$$(\tilde{R}_E e' + t_2) \frac{\partial x_2}{\partial t_L} + t_L \frac{\partial l}{\partial t_L} = 0, \quad (\text{a4.11})$$

$$(\tilde{R}_E e' + t_2) \frac{\partial x_2}{\partial t_2} + t_L \frac{\partial l}{\partial t_2} = 0. \quad (\text{a4.12})$$

This system is consistent with  $t_L = 0$ ,  $t_2 = -\tilde{R}_E e'$  and  $\beta^{FB} = 1/\lambda^{FB}$ . Inserting  $t_2 = \text{MED} = (-U_{EE}')/\lambda$  in (a4.6) and using  $\beta^{FB} = 1/\lambda^{FB}$ , we find that  $-\tilde{R}_E e'$  reduces into  $(-U_{EE}')/\lambda = \text{MED}$ , whereby the solution of the first order conditions in the first best case imply that  $t_L = 0$ ,  $t_2 = \text{MED}$ ,  $\lambda^{FB} = 1/\lambda^{FB}$  and  $\tilde{R}_E e' = (U_{EE}')/\lambda^{FB}$ , cf. (12).

In the second best case, we define the price corrective second best tax element by  $t_2^{PC} = -\tilde{R}_E e'$  and write (a4.3) as

$$l + t_L \frac{\partial l}{\partial t_L} + (t_2 - t_2^{PC}) \frac{\partial x_2}{\partial t_L} = \beta^{SB} \lambda^{SB} l, \quad (\text{a4.13})$$

obtaining

$$\beta^{SB} = \frac{1 + \frac{\partial l}{\partial t_L} \frac{t_L}{l} + \frac{\partial x_2}{\partial t_L} \frac{(t_2 - t_2^{PC})}{l}}{\lambda^{SB}} = \frac{1 + \varepsilon + \frac{\partial x_2}{\partial t_L} \frac{(t_2 - t_2^{PC})}{l}}{\lambda^{SB}}. \quad (\text{a4.14})$$

Finally, using (20) and the fact that  $\beta^{FB} = 1/\lambda^{FB}$ , we have the result shown in (26) in the paper, i.e.,

$$\frac{\beta^{FB}}{\beta^{SB}} = \frac{1}{1 + \varepsilon + \frac{\partial x_2}{\partial t_L} \frac{(t_2 - t_2^{PC})}{l}} \frac{\lambda^{SB}}{\lambda^{FB}} = \left( \frac{\mu}{\lambda} \right)^{SB} \frac{\lambda^{SB}}{\lambda^{FB}}. \quad (\text{a4.15})$$

## References:

- Ballard, C.L. and S.G. Medema (1993): The marginal efficiency effects of taxes and subsidies in the presence of externalities, *Journal of Public Economics* 52, 199-216.
- Bovenberg, A.L. and R. de Mooij (1994): Environmental levies and distortionary taxation, *American Economic Review* 84, 1085-1089.
- Bovenberg, A.L. and F. van der Ploeg (1994a): Environmental policy, public finance and the labour market in a second best world, *Journal of Public Economics* 55, 349-390.
- Bovenberg, A.L. and F. van der Ploeg (1994b): Environmental policy, public goods and the marginal cost of public funds, *Economic Journal* 104, 444-454.
- Brendemoen, A. and H. Vennemo (1996): The marginal cost of funds in the presence of environmental externalities, *Scandinavian Journal of Economics* 98, 405-422.
- Goulder, L. H. (1995): Environmental taxation and the double dividend: A Reader's Guide, *International Tax and Public Finance* 2, 157-183.
- Håkonsen, L. (1997): An investigation into alternative representations of the average and marginal cost of public funds, Discussion Paper 11/97, Department of Economics, Norwegian School of Economics and Business Administration, Bergen, Norway.
- Munk, K.J. (1977): Optimal taxation and pure profit, *Scandinavian Journal of Economics* 79, 1-19.
- Sandmo, A. (1975): Optimal taxation in the presence of externalities, *Swedish Journal of Economics* 77, 86-98.

# Chapter 5. CO<sub>2</sub>-stabilisation may be a «no-regrets» policy: A general equilibrium analysis of the Norwegian economy<sup>\*\*</sup>

## 1. Introduction

During the last few years a number of studies on the economic consequences of reducing the emissions of greenhouse gases (GHG), in particular CO<sub>2</sub>, have been performed. The methodologies used and the conclusions reached vary considerably. Manne and Richels (1994) argue that «[t]he principal reason why studies differ is alternative views about the future character of the energy system.» Their estimates of the welfare cost of stabilising CO<sub>2</sub> emissions at the 1990-level, vary between 0.2 and 6.8 percent of GNP depending on the actual parameter values. Rather than varying parameters of the energy system as they do, we consider alternative ways of redistributing the CO<sub>2</sub> tax revenue and alternative representations of the welfare index. Like Manne and Richels we find a wide range for the cost of emission reductions.

In theoretical analyses, the welfare function is often taken to include consumption of traded commodities as well as leisure and consumption of amenities. In such models, reduced deterioration of the environment may affect welfare positively through two channels. An improved environment, e.g. reduced air and water pollution, has a direct welfare effect. In addition, it increases the efficiency of production, thereby allowing increased consumption of traded commodities. In response to such ideas, there is a growing literature on the measurement of economic welfare, e.g. Nordhaus and Tobin (1972), Cobb and Daly (1989), and Brekke and Gravningsmyhr (1994). The issue is how to adjust our conventional GNP-measure for defensive expenditures on road maintenance and injuries from traffic or pollution, for the value of leisure, household work and disamenities of urbanisation, etc. Our work is in line with this kind of reasoning.

Numerical analyses differ in a number of dimensions, e.g. with respect to time frame, whether or not there is a feedback on the economy from global warming, and in details

---

<sup>\*</sup> Published in *Environmental and Resource Economics* 9, 171-198 (co-author: Lars Mathiesen).

<sup>\*</sup> Helpful comments and suggestions from Kjetil Bjorvatn, Dag Morten Dalen, Rolf Golombek, Ulf Pedersen, Guttorm Schelderup, Øystein Thøgersen, two anonymous referees, and seminar participants at NHH, in Kristiansand S, and Voss are appreciated. The research has received support from the Department of Finance. As usual, errors and presented views remain the authors' responsibility.

of the energy system, economic activities and mechanisms featured. Typically, however, the welfare-index is based on a GNP-measure neglecting feedbacks from a damaged local environment. (A notable exception is Ballard and Medema (1993).) Furthermore, analyses based on optimisation models that do not easily accommodate *ad valorem* taxes, disregard the presence of an existing tax system, i.e., they use a first best equilibrium as their reference. In this setting, any additional constraint, like a restriction on total CO<sub>2</sub> emission, necessarily implies a welfare cost.

Simplified modelling is in line with an economic tradition to include only variables and relations that can be measured with a reasonable degree of precision. It is often suggested that one is better off in terms of reliability and credibility of policy advice by supplementing hard numbers with verbal description of softer arguments. Our point is that in the present context vital information is lost by choice of modelling framework. Carbon regulations have pervasive consequences on our economies, some of these consequences are large and some also have positive effects on welfare.

Failing to reach a cost-effective international agreement, countries agreed at the 1992 Rio Earth Summit that they would attempt to stabilise yearly emissions at the 1990-levels by the turn of the century. This goal may imply reducing emissions by some 20 percent compared to a «business-as-usual»-scenario. A number of analyses suggest that the costs of such reductions will amount to one percent or more of GNP. Seemingly, many policy-makers feel that this is too high a cost to inflict on present generations, and the implementation of necessary measures is halted<sup>1</sup>. This stand may be rationalised as follows. Firstly, the cost will be borne by the rather few who will become involved in the reallocation of resources, in particular those who will lose their present occupation. In an era of high unemployment, a political reluctance towards creating higher unemployment - even on a temporary basis - is understandable. Secondly, the benefits from such regulation, accruing in the rather distant future, are highly uncertain. Finally, inaction could be ascribed to a Prisoner's Dilemma situation where it is felt to be individually rational not to regulate. (See Hoel (1991).)

We sympathise with these points, but find the premises to be misleading. We think that a too simple representation of the economic system and the welfare-index of present generations have biased these cost-calculations towards a negative conclusion. In fact, we demonstrate that by including relations and feedbacks which there is wide agreement exist, the con-

---

<sup>1</sup> In spite of considerable effort, EU has been unable to reach an agreement among member countries on the implementation of a suggested energy and carbon tax.

clusion could be reversed. That is, welfare for present generations could be improved by restricting CO<sub>2</sub> emissions. Any benefits to future generations would only add to the profitability of such a regulation. Our result hinges on inefficiencies in the present economic system. We do *not* say that restricting CO<sub>2</sub> emissions is a first best solution to such inefficiencies, but that it is likely to be a «no-regrets» policy.

The paper builds up this argument in a stepwise fashion. Employing a fairly detailed, static CGE-model of the Norwegian economy, we analyse the welfare cost of reducing total Norwegian CO<sub>2</sub> emissions by year 2000. In the reference version, labour supply is exogenously stipulated and the revenue from the uniform CO<sub>2</sub> tax is returned lump-sum to private households. The first extension of the model is to endogenise labour supply and thus include leisure. In the second step, the CO<sub>2</sub> tax revenue is used to reduce a distortionary tax, namely employer's social security contributions. The third step is to include negative externalities. We identify two sources. One is the emission of several non-GHGs like CO, SO<sub>2</sub>, and NO<sub>x</sub>, which in various ways cause damages to health, buildings, etc. From previous analyses of the Norwegian economy we know that a by-product of reduced CO<sub>2</sub> emissions is a significant reduction of these components as well. (See Brendemoen and Vennemo, 1994 and Håkonsen and Mathiesen, 1994). The other source is transportation. Reducing traffic volumes will reduce CO<sub>2</sub> emissions and costs related to road maintenance, accidents and time loss due to congestion. The final version of the model accounts for the subjective part of such externalities, i.e., the disutility from noise, dust, and other «bads» that are not represented through markets.

These steps have been analysed in some way or another in previous studies<sup>2</sup>. Our contribution is to bring together a systematic comparison of the effects on the cost of reducing CO<sub>2</sub> emissions. According to the reference version of our model, 20 percent CO<sub>2</sub> reduction is associated with a one percent welfare *cost* compared to a «business-as-usual»-scenario without any CO<sub>2</sub> regulation. In the final version with reduced labour tax and externalities accounted for, the 20 percent regulation yields a one percent *gain*. The difference in results from different descriptions of the economic system and its welfare-index is thus significant.

---

<sup>2</sup> Goulder (1995a) and Nordhaus (1993) compare alternative ways of redistributing the revenue from CO<sub>2</sub> taxation. Our modelling of local pollution is close to that of Ballard and Medema (1993). Data for this modelling is from several Norwegian analyses in Statistics Norway, e.g. Alfsen et al. (1992) and Brendemoen et al. (1992).

The remaining part of this paper is organised as follows. Sections 2 and 3 present the principals of our model and the reference case. Further details are contained in Appendices 1-3. In sections 4 and 5 we discuss results from alternative formulations of labour supply and redistribution of tax revenue. Externalities are introduced in section 6. We evaluate our findings and conclude in section 7.

## **2. The model framework**

We employ a static CGE-model representing the Norwegian economy. Industrial and other economic activities are disaggregated so as to obtain fairly homogeneous sectors with respect to emission coefficients. Thus, transportation activities and process industries, like primary metals production, are treated in large detail, while service sectors are more lumped together<sup>3</sup>. The model considers one aggregate private household. Export or import of any commodity depends on the endogenous Norwegian price relative to a stipulated world market price of that commodity. Real investments, public consumption and the value of net exports are all kept constant at fixed prices, whereby private consumption is the only element in GNP that varies in real terms across the scenarios. The public budget is balanced by a lump-sum transfer to households.

Within our rather short time horizon of 10 years, the modelling of resource reallocation may be critical. We have conducted the analysis under different assumptions regarding the mobility of capital and labour. Although equilibrium values differ across assumptions, we arrive, however, at essentially identical conclusions. The model disregards reallocation costs. Thus, while the regulated equilibria call for reallocation between sectors thereby possibly causing structural unemployment, we assume that any unemployment has vanished by the time horizon. Hence, our analysis, like most other CGE analyses, underestimates some of the costs of CO<sub>2</sub>-reducing policies.

Total emissions are caused by burning fossil fuels or processing carbon-containing materials like limestone. Hence, emissions are related to production and consumption sectors' demand for these commodities. Assume that no sector is allowed to emit CO<sub>2</sub> unless it has obtained a permit to do so. In this interpretation, the aggregate demand for emission permits from production and consumption sectors equals total emission. Our model focuses on the

---

<sup>3</sup> Appendix 1 provides an overview of model disaggregation and activity levels. Commodities follow a similar disaggregation. Energy goods and the various types of emission components are also listed. Finally, the hierarchy of commodity transformation is illustrated.

emission permit, an economic good, rather than the emission, which is a bad. It is assumed that the government issues emission permits according to a specified goal for the total emission level,  $E$ . The permit is treated like other inputs to production or consumption. Its demand,  $e(\mathbf{p}, q)$ , is a function of the vector of factor prices,  $\mathbf{p}$ , and the price of the permit,  $q$ . Its equilibrium price  $q^*$  is determined by the complementary slackness condition

$$\begin{aligned} E - e(\mathbf{p}^*, q^*) &\geq 0, \\ q^* &\geq 0, \\ q^* [E - e(\mathbf{p}^*, q^*)] &= 0. \end{aligned}$$

That is, excess supply of the permit and its price must both be non-negative, and if either one is positive, the other has to be zero.

The interpretation of an emission permit as a commodity does not necessarily imply that a market for the permit has to be established in the economy. The price can alternatively be thought of as the required tax per unit of  $\text{CO}_2$  in order to accomplish a prespecified reduction of total emissions.<sup>4</sup> In an economy outside a second best equilibrium, the cost of obtaining a given reduction in  $\text{CO}_2$  emissions is most likely<sup>5</sup> minimised with *non-uniform*  $\text{CO}_2$  tax rates, see Håkonsen (1995). Our model does not easily allow the computation of optimal commodity taxes, however. Therefore we follow the programming tradition and use the shadow price of the emission constraint as a uniform  $\text{CO}_2$  tax rate.

### 3. The reference case

In an emission-reducing scenario, the government collects a  $\text{CO}_2$  tax revenue,  $qE$ , that in general may be balanced by reductions in other taxes, increased transfers or expansion of expenditures. Because of this multitude of ends there is no single answer to the question of how much a given reduction of  $\text{CO}_2$  will cost. In fact, there will be one answer for every possible way of redistributing the revenue. As a reference case, we compute an equilibrium where the  $\text{CO}_2$  tax revenue is returned as a lump-sum to the household. In this reference case welfare is measured by an index of private consumption goods,  $C \equiv V(C_1, \dots, C_4)$ . (See Figure A2.1 in

---

<sup>4</sup> We will use the descriptions «emission tax» and «price on emission permits» interchangeably.

<sup>5</sup> If other tax rates depart from the second best levels, differentiated  $\text{CO}_2$  tax rates might bring the economy closer to the second best optimum than equal  $\text{CO}_2$  tax rates for all emitting activities.

appendix 2.) This set-up comes close to the economic content of optimisation models. Observe that C is the only component of GNP that changes with regulation.

Figure 1 displays the price of the CO<sub>2</sub> emission permit and the resulting CO<sub>2</sub> tax revenue, while Figure 2 presents the corresponding change in the welfare index. The conclusion in this setting is that welfare is reduced along with the CO<sub>2</sub> reductions. A 20 percent reduction of CO<sub>2</sub> emissions reduces V by about 0.8 percent<sup>6</sup>. This equilibrium may be obtained by levying a tax<sup>7</sup> of US\$ 107 per metric ton of CO<sub>2</sub>. The resulting revenue is 3.8 billion US\$, which amounts to 8 percent of total public tax revenue and about 2/3 of the employers' contributions to social security. Hence a CO<sub>2</sub> tax may provide a substantial source for public revenue and allow for a considerable tax reform.

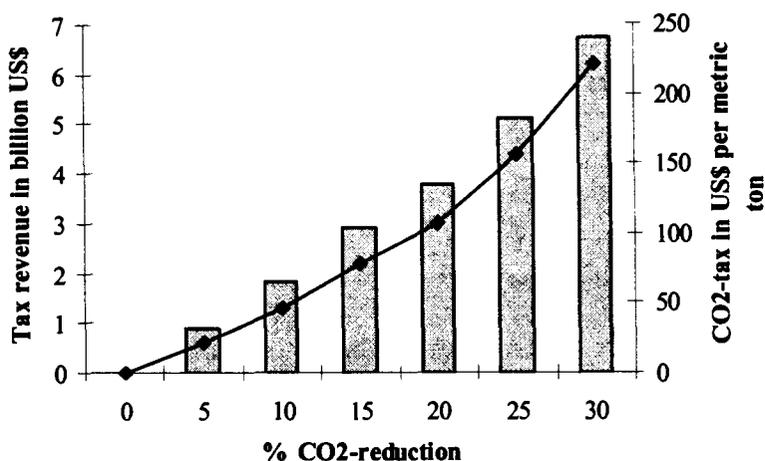


Figure 1. CO<sub>2</sub> tax rate (line) and tax revenue (bars) in the reference case

The original argument for a CO<sub>2</sub> tax is to correct for the externalities - damages from global warming - caused by CO<sub>2</sub> emissions. We ignore such externalities in our model for two reasons: Firstly, the cost from global warming will be felt decades into the future, and could not possibly be incorporated within our model's time horizon. Secondly, we analyse unilateral Norwegian reductions which will have a negligible effect on global warming within any time horizon. Thus, the CO<sub>2</sub> tax is only distortionary in our model. Negative externalities due to domestic emissions of NO<sub>x</sub> and SO<sub>2</sub> are accounted for in section 6, however.

<sup>6</sup> Nordhaus (1991) and Manne and Richels (1994) review tax and cost estimates from a number of studies. Our observed 0.8% loss for 20% CO<sub>2</sub> reduction is at the lower end of such estimates. Since 20% CO<sub>2</sub> reduction is required for stabilising the Norwegian emissions at the 1990-levels, we focus on this particular goal for emission reductions throughout the paper.

Outside a second-best optimum, one cannot in general say for certain whether a new tax improves or worsens welfare. In most cases, one would expect that raising extra revenue with a distortionary tax and redistributing the proceeds in a lump-sum manner, would increase the efficiency loss from taxation<sup>8</sup>. In a closed economy, an increased efficiency loss would translate directly into reduced welfare. In an open economy, however, there may be a possibility for improved terms-of-trade from domestic taxation. We describe export demand for Norwegian products by falling and fairly elastic demand curves for Norwegian exports. This assumed ability to transfer some part of a domestic cost increase into higher export prices, explains the slight increase in private consumption in the 5% reduction scenario in Figure 2. The general picture in the reference case, however, is that a CO<sub>2</sub> tax aimed at curbing emissions, translates into higher costs of production and thus higher consumer prices and reduced consumption.

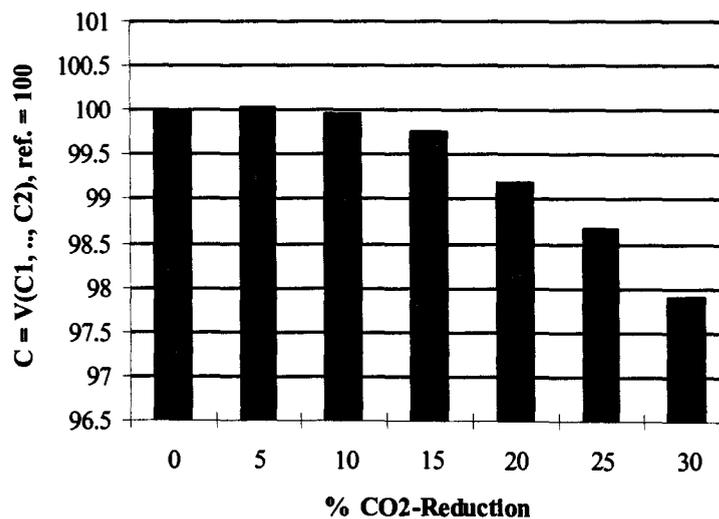


Figure 2. Reduction in private consumption

#### 4. Labour supply

With the labour-leisure choice being endogenous, an increase in consumer prices relative to the wage may cause a shift from commodity consumption towards leisure. The relevant welfare-indicator for such a trade-off is a combined index of commodity and leisure consumption, i.e., the utility-index  $U(C,L) = U(V(C_1, \dots, C_4), L)$ , for which we assume a CES-function

<sup>7</sup> Tax rates per ton *carbon* is 3.66 times the tax rates per ton CO<sub>2</sub>.

<sup>8</sup> If the tax system is at a second-best optimum at the outset, raising extra tax revenue with a distortive tax will certainly increase the efficiency loss. Being outside the second-best optimum, however, income effects matter, and they could, at least in principle, reverse this result.

$$U(C,L) = [aC^r + (1-a)L^r]^{\frac{1}{r}}$$

Parameterisation of this function involves two issues: the weighing of C and L, and the implied wage-elasticity of labour supply. Rather than stipulate the parameters of the utility-function directly, we consider derived demands, impose empirically plausible elasticities on consumption demand and labour supply, and then infer parameters of the utility function<sup>9</sup>.

We maintain the assumption that the CO<sub>2</sub> tax revenue is transferred as a lump-sum to the household. Figure 3 presents the indices  $V(C_1, \dots, C_4)$  and  $U(C,L)$ . In addition, an index of commodity-consumption only is computed for the model of endogenous labour supply. This is denoted  $U(C,0)$  in Figure 3. Hence there are 3 columns for each reduction level. The first column of each triplet corresponds to that of Figure 2.

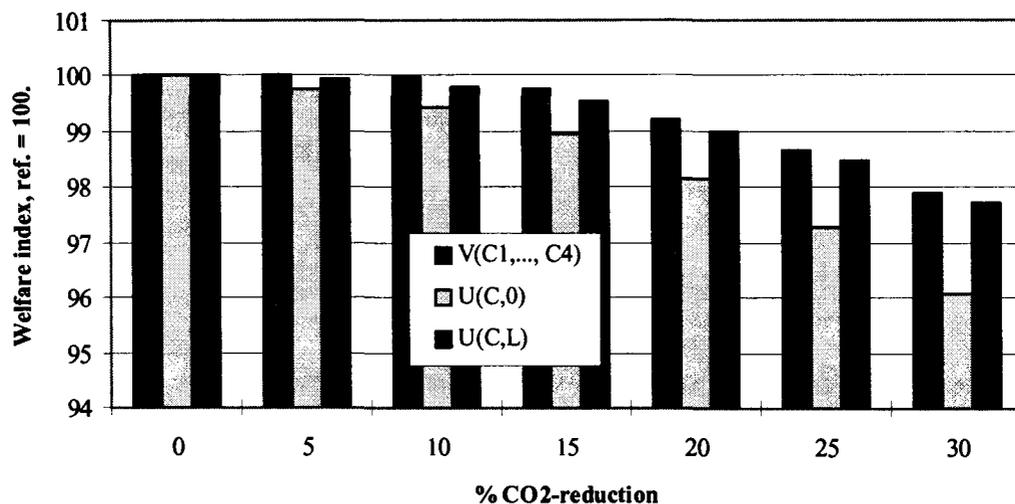


Figure 3. Endogenous vs. exogenous labour supply

There is a noticeable difference in the change of welfare from a CO<sub>2</sub> regulation according to these three measures. Choosing the index that follows directly from either of the two alternative model formulations, i.e., columns number one or three, suggest that the difference may be insignificant. This picture, however, is considerably changed when the model is further extended in the next section.

Choosing a GNP-measure like commodity consumption as welfare indicator when the model (and presumably also the economy) allows for a labour-leisure choice, may seri-

<sup>9</sup> Our calibration implies a 80-20-weighing of C and L in terms of expenditure, and a wage elasticity of labour supply of 0.3. See Appendix 2.

ously distort the impression of the welfare cost of regulation. Compare columns 2 and 3. According to  $U(C,0)$ , welfare is reduced about twice as fast as for  $U(C,L)$ . Of course, the index in the second column is inconsistent: This index exaggerates the welfare loss since the household's optimal substitution from C to L is neglected. At the 20 percent  $CO_2$  reduction, for example, the welfare loss measured by  $U(C,L)$  is a weighted average of a 1.8 percent drop in commodity consumption and a 2 percent increase in leisure.

Acknowledging the need for a shift away from traditional consumption of «dirty» commodities seems to be the essence of the debate on a sustainable development. Admittedly, our model does not represent all consequences of increased leisure and changes in the consumption and production patterns. E.g. the model has no direct link between leisure-time and the consumption of leisure-commodities like air-travel, and no accounting is made for changes in commuting patterns because of a new industrial structure.

## 5. The redistribution of the $CO_2$ tax revenue

As pointed out above, adding a distortionary tax and redistributing the revenues in a lump-sum manner will, apart from possible terms-of-trade-effects, in most cases increase the efficiency loss from taxation. A far more interesting procedure is to balance the budget by reducing some existing distortionary tax rate. In this manner, environmentally motivated policy changes are integrated into the more general field of revenue-neutral tax reform analysis.

We have chosen to reduce the employer's statutory tax on labour. This is not an arbitrary choice. Firstly, the effective tax wedge in Norway between what labour income can buy in terms of consumption goods and the cost of labour for producing such goods, is almost 70 percent for individuals in higher income brackets.<sup>10</sup> Secondly, policies implying a higher intensity of employment and lower intensity of energy and natural resource use is at the heart of the policy debate both in Norway and in the European Community.

«The current development model in the Community is leading to a sub-optimal combination of two of its main resources, i.e. labour and nature. The model is characterised by an insufficient use of labour and an excessive use of natural resources, and results in a deterioration of the quality of life.» [Commission of the European Communities, White Paper, 1993.]

---

<sup>10</sup> There are three wedges between the worker's net wage and consumer prices: i) the employer's labour tax, with the rate 14.3% (maximum), ii) the employee's income taxation with a maximum marginal rate of 49,5%, and iii) the VAT of 23%. In addition to VAT, there are also other indirect taxes on e.g. gasoline, tobacco, and alcoholic beverages, making the effective tax wedge even higher for these particular commodities.

Finally, within a similar CGE-model of the Norwegian economy, Mathiesen (1993) found that the reduction of the employer's tax on labour would give a considerably larger boost to welfare (measured in terms of equivalent variation) than ex ante equal-revenue reductions in either energy-taxes or other taxes on inputs to production.

In Figure 4 we compare utility-indexes of lump-sum and labour-tax redistribution in a model with endogenous labour supply. The scenarios are denoted LUMP-SUM<sup>11</sup> and LABOUR TAX respectively. In light of the «double dividend hypothesis»-debate, this Figure provides interesting results. Goulder (1995b) defines three alternative varieties of double dividend claims:

- a) The *weak form* says that using the revenue from an environmental tax to finance reductions in existing distortionary taxes reduces the costs of the policy compared to lump-sum redistribution to tax payers.
- b) The *intermediate form* says that it is possible to find a distortionary tax rate such that a revenue neutral substitution of an environmental tax for this particular tax involves a zero or negative gross cost<sup>12</sup>.
- c) The *strong form* says that a revenue neutral substitution of an environmental tax for *typical* or *representative* distortionary taxes involves a zero or negative gross cost.

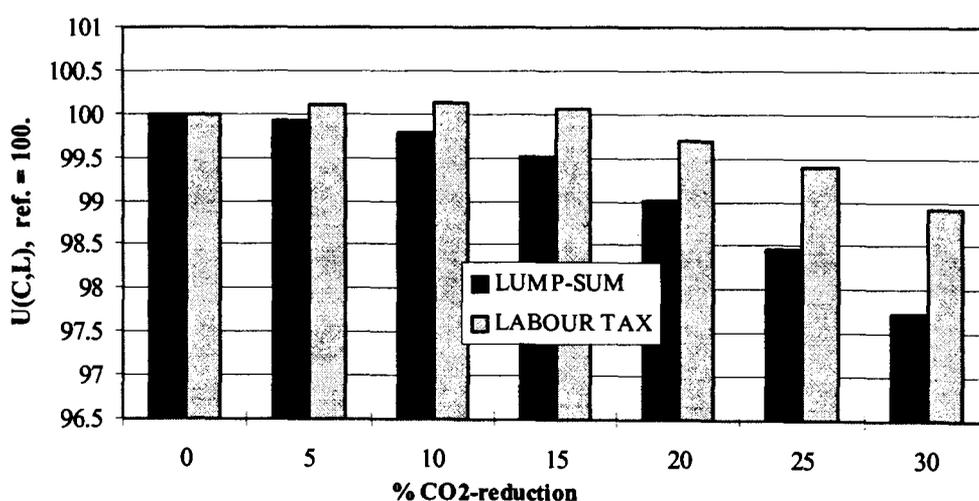


Figure 4. Alternative redistribution mechanisms with endogenous labour supply

<sup>11</sup> The LUMP-SUM-scenario corresponds to the rightmost of the three columns in Figure 3.

<sup>12</sup> Note that «gross cost» is the welfare cost when neglecting effects on the environment, i.e., exactly what we do in this section. Thus the results in this section only relate to one of the two «dividends», namely the efficiency effect of substituting one tax for another. When we proceed to include the effects of CO<sub>2</sub> taxes on environmental variables in section 6, both «dividends» will be accounted for.

By inspection of the whole range of emission reductions from 5% to 30%, our results confirm the *weak form*: Cutting employer's labour tax rates is superior in terms of efficiency relative to lump-sum redistribution. This result is not surprising and is also confirmed by Goulder (1995a). The utility-indexes develop considerably different. At 20 percent CO<sub>2</sub> reduction, the welfare loss is 0.3 percent when reducing the labour tax, as compared to 1 percent with the lump-sum transfer. Goulder suggests that the use of CO<sub>2</sub> tax revenue for reduction of some other distortionary tax might reduce the welfare loss by 36 to 53 percent<sup>13</sup> for the US economy. Nordhaus (1993) performs a similar experiment although his computation of welfare gain is different. He concludes that (p. 316): «The surprising result of this experiment is that the gain from efficient use of green taxes is quite substantial.»

In the reduction range from 5% to 15%, our computations also support the *intermediate form* of the double dividend hypothesis, namely that CO<sub>2</sub> taxation in combination with reduced labour taxation is *welfare improving*. This effect is not found by Goulder. The Figure suggests that an «optimal» policy would be to reduce CO<sub>2</sub> emissions by 10 percent. This would call for a CO<sub>2</sub> tax of about 50 US\$ per metric ton CO<sub>2</sub> and a reduction in employer's labour tax rates of almost 30% from the benchmark level.

In a theoretical model, Bovenberg and de Mooij (1994) reject the intermediate and strong versions of the double dividend hypothesis. They find that welfare (exclusive of environmental quality) is reduced when increasing an environmental tax and reducing the labour income tax. This result follows from their assumed preference structure, however, since a proportional labour income tax represents the Ramsey-optimum in their model. Cutting the labour income tax and introducing a tax on a polluting commodity therefore implies a departure from the Ramsey-optimum, and must by definition decrease welfare (exclusive of environmental quality). Hence, their analysis is invalid as a *general* rejection of the intermediate and stronger forms of the double dividend hypothesis. It brings out the insight though, that there must be some kind of inefficiency in the initial tax structure for the two stronger forms of the double dividend hypothesis to materialise. This conclusion is also reached by Goulder (1995b), section 2.3.

---

<sup>13</sup> Goulder considers four alternative tax instruments, personal tax, profits tax, payroll tax and all taxes. Compared to the lump-sum case, the welfare loss is 36% smaller with cuts in the personal tax rates, 37% smaller with cuts in the profits tax, 53% smaller with cuts in the payroll taxes, and 42% smaller when all taxes are reduced.

Since there is no reason to believe that today's tax-system represents even an approximate second-best optimum, changing one tax for another could indeed be efficiency-improving. Our case is just one example. Of course, the conclusion from such a tax experiment is country-specific, since the tax structure varies from country to country. Thus, while the efficiency gain from cutting the labour tax rates is higher than the efficiency loss from increasing CO<sub>2</sub> taxes at our model's benchmark equilibrium, the opposite may be the case for other countries. Furthermore, different results from such a tax experiment may also be due to different modelling assumptions. This is illustrated by the fact that Goulder (1995a) finds no support for the intermediate form for the US economy, while such support is found from the Jorgenson-Wilcoxon model<sup>14</sup>, cf. Shackleton et al. (1992).

Carraro et al. (1994) point out that a reduction of the labour tax instead of lump-sum-redistribution, counteracts a goal of reducing CO<sub>2</sub> because the cost of production is reduced, thereby increasing the scale of production. This point is confirmed in our analysis, but is of negligible empirical significance. The required CO<sub>2</sub> tax for achieving 20% reduction increases from 50 to 52 US\$ from the LUMP-SUM to the LABOUR TAX scenario. Alternatively stated, using the CO<sub>2</sub> tax computed from the LUMP-SUM scenario in the LABOUR TAX scenario, will reduce emissions by 19.3% and not the full 20 %.

In addition to efficiency arguments, distributional considerations will contribute to the attractiveness of the labour-tax reducing alternative. Several researchers have pointed out that a CO<sub>2</sub> tax may be regressive because low-income households have a higher budget share for energy than do high-income households. Reduced taxation of labour results in increased wages and favours those who have their income mainly from labour.

«Political» considerations could also be relevant for the acceptance and implementation of a CO<sub>2</sub> tax. Firms will be directly taxed by the proposed CO<sub>2</sub> tax. Returning this tax revenue by reducing employers' labour taxes would presumably increase the attractiveness of the CO<sub>2</sub> policy from a producer's point of view compared to a policy where the revenue is paid directly to households.

Empirical evidence on the wage elasticity of labour supply seems to indicate that this parameter is positive, but low. One should be careful, however, and not interpret a low elasticity as being approximately synonymous with a fixed labour supply. In the endogenous-

---

<sup>14</sup> In the Jorgenson-Wilcoxon model increased CO<sub>2</sub> taxation increases welfare both with labour tax and capital tax redistribution; the welfare gain being highest with cuts in the labour tax rates.

supply model and with a zero wage elasticity, the elasticity of substitution between C and L is calibrated so that the income and substitution effects from an increased real wage, *ceteris paribus*, are exactly equal. However, the real wage is certainly not the only price that changes across the studied scenarios. When CO<sub>2</sub> taxes are introduced in change for reduced labour taxation, the returns to capital on average go down. Since the household's income stems from both labour and capital, reduced returns produce a negative income effect, and hence increased labour supply. Figure 5 illustrates this point. Observe that the utility-index from the model with exogenous labour supply (denoted EXOG in Figure 5, and  $V(C_1, \dots, C_4)$  in Figures 2 and 3), differ considerably from the three alternative model versions with endogenous supply (i.e., the utility index  $U(C,L)$ ) and wage elasticities of 0, 0.3 and 0.6 respectively (denoted ENDOG in Figure 5). This finding supports a claim that the choice of «model-structure» is more important than «precision» of estimates.

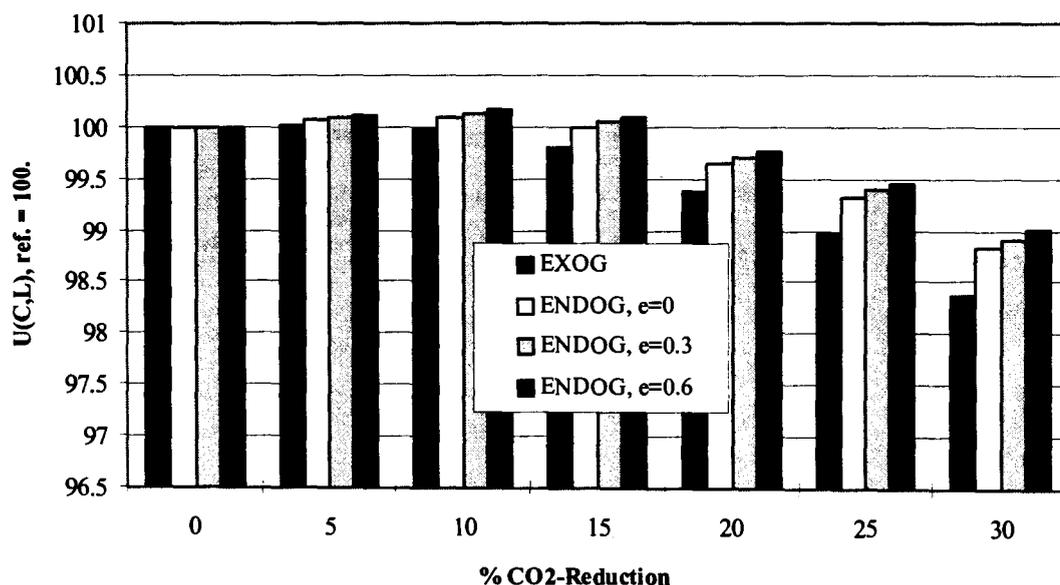


Figure 5. Labour supply: model structure vs. parameter value

## 6. Including emissions of other components and negative externalities

So far, we have only considered CO<sub>2</sub> emissions. Our model also includes a number of other emission components like SO<sub>2</sub> and NO<sub>x</sub><sup>15</sup>. While CO<sub>2</sub> is the only component under *direct regulation* in this analysis, emissions of all other components will be indirectly affected by the CO<sub>2</sub> policy. Håkonsen and Mathiesen (1994), using the same model, reported that such emis-

<sup>15</sup> For a complete list, see Appendix 1, Table A1.3.

sion levels are reduced when CO<sub>2</sub> emissions are restricted. That is, all such components are *complements* to CO<sub>2</sub>.<sup>16</sup> This point is also demonstrated by Brendemoen and Vennemo (1994). Figure 6 illustrates these side-effects from reduced CO<sub>2</sub> emissions. The graphs correspond to the equilibria presented in the LABOUR TAX case in Figure 4. It is not surprising that components are complements rather than substitutes to CO<sub>2</sub>. This follows because most emissions are generated from the use of the same energy inputs. The burning of heating oil, for example, leads to emissions of all components, although very little of some. The lack of abatement possibilities for CO<sub>2</sub> implies that reduced emissions have to be based on reduced use of fossil fuels.

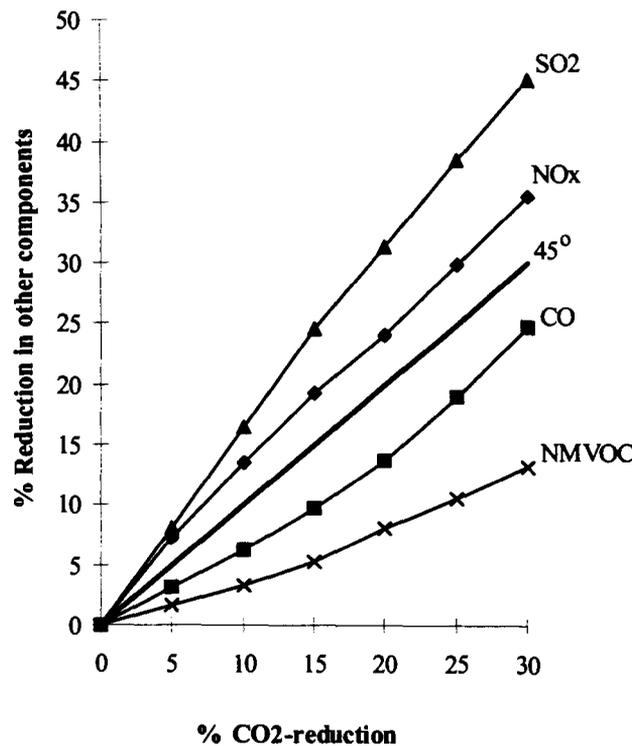


Figure 6. Effects on non-greenhouse gas emissions

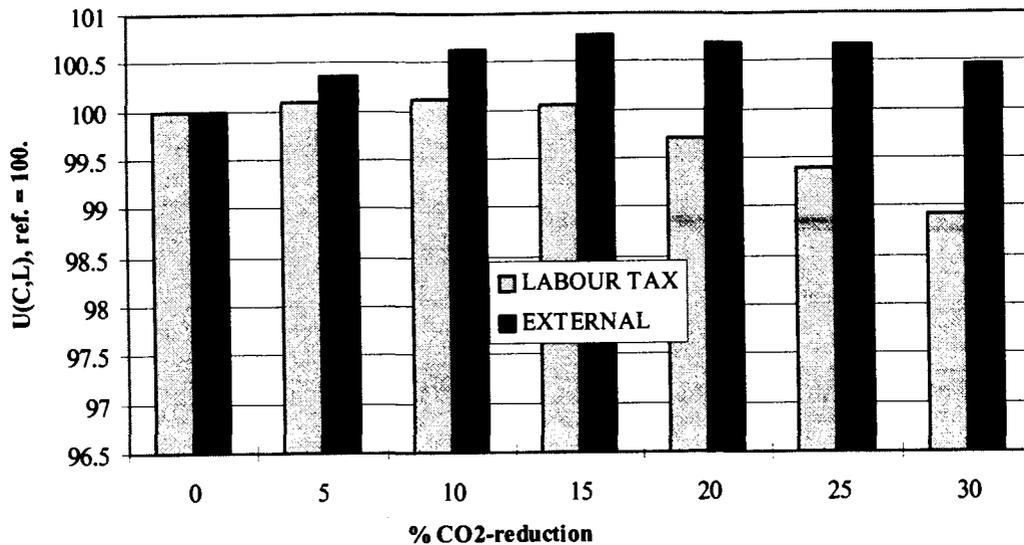
As seen from Appendix 1, all non-GHG components have negative consequences for consumers, producers, structures or vegetation. That is, they create negative externalities. Since these costs are external, the linkages between the externality generating activity and the level of externality costs are not represented in the National Account's input-output data from which

<sup>16</sup> For any component  $i$  there is a total demand  $e_i$ , defined analogously to CO<sub>2</sub> in section 2. Component  $i$  is said to be a complement to CO<sub>2</sub> if  $\frac{\partial e_i}{\partial q} < 0$ , where  $q$  denotes the price of CO<sub>2</sub> emission permits.

the model is calibrated. We have to look elsewhere to find data for such a description. Alfsen et al. (1992) and Brendemoen et al. (1992) provide estimates on the magnitude of externality costs from emissions to air and from road traffic. Road traffic is one of the most prominent sources for CO<sub>2</sub> emissions, and the CO<sub>2</sub> reducing scenarios will indeed have consequences for the overall traffic volume. For this reason, externality costs from road traffic are highly relevant for our analysis.

In the reports referred to above, these effects have been expressed in terms of costs. In a general equilibrium framework, costs and value terms result from the consumption of some scarce resource or commodity. Thus, if a marginal cost measure is to be integrated into the model, that cost figure must be translated into unit consumption of some commodity, e.g. time (labour or leisure), or public sector services. We have conducted such a translation, and some of the most important cost measures have thus been explicitly integrated into the model. Appendix 3 provides a brief description of this modelling exercise. Some of the marginal cost data provided by Alfsen et al. and Brendemoen et al. do not correspond directly to consumption of traded commodities, rather they concern subjective utility assessments. The consequences from this part of the externality cost data are considered towards the end of this section.

The linkages between emissions and road traffic on the one hand and resource costs on the other hand will necessarily increase the benefits (or reduce the costs) of the studied CO<sub>2</sub> policy. With the direction thus clarified, the remaining question is the significance of the integration of these external costs. This is seen in Figure 7, where the scenario which includes the externality costs is denoted EXTERNAL. Welfare is still measured by  $U(C,L)$  as established above.



**Figure 7. Including gains from reduced externality costs**

Without externalities accounted for, an «optimum» was found at a 10 percent CO<sub>2</sub> reduction (LABOUR TAX in Figures 4 and 7). With externalities integrated into the model, the «optimum» is extended to 15 percent reduction of CO<sub>2</sub>, see Figure 7. This reduction corresponds to a tax of 80 US\$ per metric ton of CO<sub>2</sub> accompanied by a halving of the labour tax rates. At larger CO<sub>2</sub> reductions, welfare will be somewhat lower, but even a 30 percent reduction of emissions provides a higher welfare than the reference equilibrium with its unrestricted emissions. Hence, incorporating the externalities caused by non-GHGs and road traffic, substantially alters the results and extends the scope for a «no-regrets» CO<sub>2</sub> reduction.

Integrating externality costs into the general equilibrium provides a new and interesting dimension to an analysis of the welfare effects of reduced externalities. Two different externalities, that ex ante may be evaluated the same, translate differently into the general equilibrium welfare measure. Traffic-related externalities provide a good example. Among the negative effects of road traffic are i) time consumption for households on congested roads, and ii) road damage paid for by the public sector. Suppose that the traffic volume is reduced so that time usage and resource use for road maintenance are reduced.<sup>17</sup> Suppose further that the evaluation of both of these effects are one dollar when evaluated at benchmark prices. In the model, reduced resource usage in the public sector will lower the required tax-revenue. Since taxation is distortionary, the shadow price of public funding is higher than one. Thus,

when the labour tax is reduced by one dollar, private consumption in general increases by more than one dollar. In our model private consumption increases by 1.25 dollars, which represents the marginal cost of public funds for this specific tax-instrument at the benchmark equilibrium.<sup>18</sup> Reduced time consumption, also worth one dollar at benchmark prices, translates into increased private consumption worth 1.1 dollars in the model. This shows that increasing a resource endowment worth one dollar at status quo prices in general does not give the same value when measured in terms of private consumption in the new equilibrium. While these numbers are of little interest as such, the example demonstrates the additional information from integrating externality cost measures as real variables in general equilibrium. In this respect our treatment of externalities departs from Brendemoen and Vennemo (1994) who compute welfare gains in a separate sub-model and after finding the equilibrium.

As mentioned, the marginal cost estimates reported by Alfsen et al. (1992) and Brendemoen et al. (1992) contain a mixture of market cost estimates and subjective utility assessments. While we have integrated the market cost estimates into our model, some subjective costs – traffic-related noise and health costs other than labour-losses and resource use in medical treatment – have so far been neglected. Since the subjective cost data are not represented in the equilibrium model, the remaining gains from reduced subjective costs are computed outside the model, and added to the U(C,L)-index. These subjective costs are small relative to the already included externalities. At a 15 percent reduction of CO<sub>2</sub> emissions, the inclusion of reduced subjective costs increase the welfare index as reported in Figure 7 by 0.2 percent points, increasing to 0.4 percent points at the 30 percent reduction of CO<sub>2</sub> emissions (cf. Figure 8 in the next section).

## 7. Concluding comments

We started this paper by referring to Manne and Richels (1994) and their analysis of the uncertainty in model parameters describing the energy system. Our focus has been on the welfare index of such analyses. Its description is probably more controversial than uncertain. The controversy applies to the choice of which factors to include and to the appropriate weights for aggregating different dimensions of welfare into one index.

---

<sup>17</sup> Since the aggregated household allocates its time optimally given the prevailing wage rate, the value of time is the same irrespective of whether the time is used to increase leisure consumption or labour supply.

<sup>18</sup> Such a treatment of reduced resource use in public sector has linkages to the discussion in the literature on correction of GDP for «defensive resource use», see e.g. Brekke and Gravningsmyhr (1994).

Of course, these added features are also marred with uncertainty. In our model the question of parameters like the wage elasticity of labour supply, the content and curvature of damage functions from road traffic and pollution, etc. are uncertain. With the exception of the wage elasticity of labour supply, we have not looked into the possible effects of such uncertainties. Our result on this score is interesting in itself. The computations suggest that choice of model structure – here exogenous vs. endogenous labour supply – is more critical than parameter accuracy. In fact, the differences in results between a model with exogenous labour supply and three models with endogenous supply, but different supply-elasticities, are larger than the differences among these three models.

Analogous to Manne and Richels, we find the welfare cost estimate to span a fairly wide range, and equally important, our analysis demonstrates that it may be negative. That is, a CO<sub>2</sub> regulation may be welfare-improving for the present generations. Figure 8 summarises our main results. Starting with the LUMP-SUM scenario, there is a significant increase in the welfare index by redistributing the CO<sub>2</sub> tax revenues as reduced labour taxation. Including the reduced levels of negative externalities increases the welfare index even more. Finally, the dotted line at the top of Figure 8 indicates the extra benefits from reduced levels of subjective «disutility» factors (EXTERNAL+SUBJ.). At 20% reduction, the difference in the welfare index from the LUMP-SUM scenario to the EXTERNAL+SUBJ.-scenario is 2 percentage points; a one percent loss in the LUMP-SUM scenario is turned into a one percent gain in the EXTERNAL+SUBJ.-scenario.

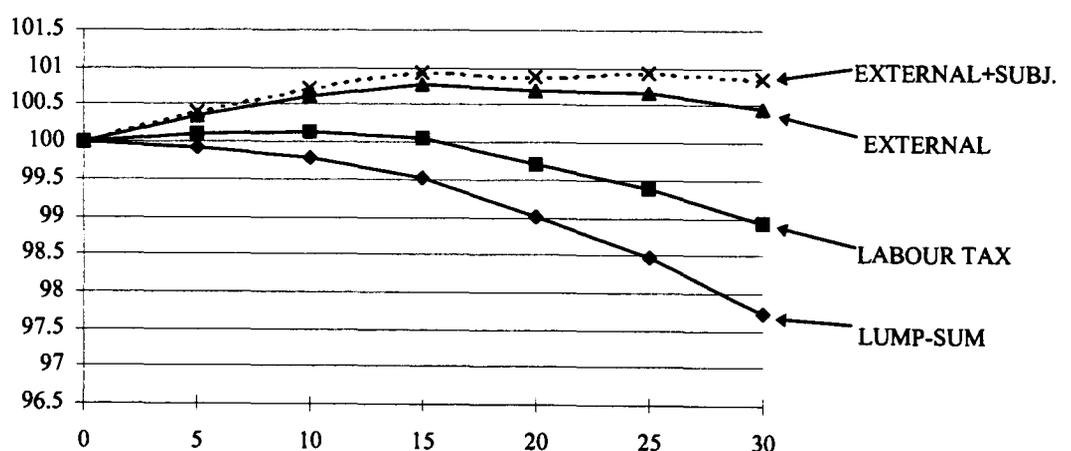


Figure 8. Welfare cost range

The possibility of a welfare gain hinges on inefficiencies in the tax-systems and in our societies' handling of externalities from local pollution. Such inefficiencies will of course be country specific, but so is the energy system, the industrial structure, the level of damages from road traffic, etc. In fact, we believe that the volumes of road traffic and local air pollution and thus the corresponding damages, are lower in Norway than in OECD-countries with larger urban areas and warmer climates.<sup>19</sup>

We do not claim that regulation of CO<sub>2</sub> emissions is a first-best solution to these inefficiencies. However, given a *ceteris paribus* assumption – that the only policy change is the analysed CO<sub>2</sub> regulation – the positive side-effects of reduced local air pollution and road traffic cannot be neglected. Of course, if more direct policies against such externalities are implemented, the positive side-effects of the CO<sub>2</sub> policy would be reduced.

Inefficiencies in our present economic system allow «no-regrets»-policies. We have singled out two areas of inefficiencies: the tax-system and negative externalities. What about other sources, e.g. imperfect competition and imperfect information? How do such market failures influence upon the welfare-effects of reduced CO<sub>2</sub> emissions? We have not looked into these questions.

We have studied a unilateral Norwegian CO<sub>2</sub> reduction. There are three important links to the outside world through which our results might be substantially changed if CO<sub>2</sub> reductions were prompted by a wide international agreement. Firstly, while domestic emissions of damage-generating pollutants like NO<sub>x</sub> and SO<sub>2</sub>, are reduced, «imported» concentrations of these substances remain unaltered. Since several of the larger damage-cost categories from air pollution in Norway mainly result from «imported» concentrations,<sup>20</sup> such costs are hardly changed at all in the studied scenarios. Secondly, the increases in energy prices in our scenarios are domestic and affect Norwegian energy-intensive exporting industries adversely. Under an international agreement, foreign competitors would presumably experience similar, if not larger price increases for their energy inputs. Such increases in production cost would translate into international price increases for products<sup>21</sup>. Thus one would think that Norway

---

<sup>19</sup> This applies both to the damages accounted for in our analysis, and to an important, but omitted pollution cost, namely the concentration of tropospheric ozone (ozone near the ground level), which is of growing concern in continental Europe.

<sup>20</sup> According to Park (1987), 92% of SO<sub>2</sub>-depositions in Norway were received from other countries.

<sup>21</sup> In such cases it is tempting to assume that cost increases are fully transferred to customers. This is reasonable in the long run, but certainly not in the short run, and furthermore it depends crucially on the international coverage of such an agreement. Manne and Mathiesen (1994) analysed the impact on the location of aluminium

would benefit even more from international agreements on reduced CO<sub>2</sub> emissions. However, a third effect pulls in the other direction. Norway consumes only about 1/10 of her total oil production. While the unilateral Norwegian reduction of CO<sub>2</sub> emissions would affect world consumption of fossil fuels and crude oil insignificantly, an international agreement or a wide adoption of such a policy, would probably cause a significant drop in the price to the producer, and hence reduce Norwegian export-revenues. This effect would hurt the Norwegian economy much more than the imposition of a unilateral CO<sub>2</sub> policy.<sup>22</sup>

Admittedly, our model is deficient in many respects. It is static and thus neglects the issue of capital formation and reallocation between sectors causing (temporary) unemployment, and it has one aggregate household and thus neglects distributional consequences. It is not at all clear in which direction such added detail would pull our results. Especially the neglect of temporary unemployment problems indicates that our results are on the optimistic side. On the other hand, we have chosen parameter values, i.e., a low wage elasticity, convex damage functions, and only a subset of subjective disutility-factors, e.g. none related to loss of amenities, that probably bias our results to underestimate the «true» welfare gain from the studied policy changes. Hence, the possibility of welfare improvements from regulation could survive a more comprehensive analysis. However, the most interesting aspect of this paper is not the discussion of the exact *level* of costs or gains from CO<sub>2</sub> regulations. Rather, it is the described *range* of such costs or gains; a range that depends on the chosen model structure, tax redistribution mechanism, and externality cost measure.

---

smelting of an OECD-carbon-tax and found that (because of relocation of smelters to non-OECD countries) only 30-40% of the tax was reflected in the world market price of aluminium.

<sup>22</sup> Mathiesen (1991) suggested that a 20 percent drop in the oil-price would cause a welfare loss about 10 times larger than a unilateral 15 percent reduction in CO<sub>2</sub> emissions.

## Appendix 1. Production sectors, emission-generating inputs, and emission components

**Table A1. 1 Production sectors and gross product in billion 1991 NOK in the reference scenario**

Sector	Gross product
Agriculture	27.8
Forestry	4.4
Fishing	5.9
Breeding of fish	5.6
Oil and gas extraction	132.4
Paper, pulp and fibre	19.4
Basic industrial chemicals except fertilizers	11.9
Fertilizers and pesticides	5.0
Paints, varnishes and laquers	1.6
Other chemical products	11.1
Petroleum refining	15.2
Other products of petroleum, coal and coke	2.1
Cement and lime	1.1
Iron and steel	3.3
Ferro-alloys	4.6
Primary Aluminum	11.2
Other non-ferrous metal	7.1
Other metal-based commodities	20.3
Other manufacturing industries	226.9
Construction, buildings, dwellings	79.5
Low-voltage electricity	21.9
High-voltage electricity	11.3
Railway transport	3.3
Scheduled bus transport	5.6
Tramway and subway	0.5
Taxi and other unscheduled passenger transport by road	1.3
Unscheduled freight transport by road	8.1
Ocean transport	52.6
Coastal and inland water transport	6.5
Air transport	13.0
Postal and telecommunication services	25.1
Wholesale and retail trade	93.7
Other private services	215.6
Sanitary and similar public services	4.1
Defence	23.8
Other governmental services	143.0

(exchange rate 6.50 NOK/US\$)

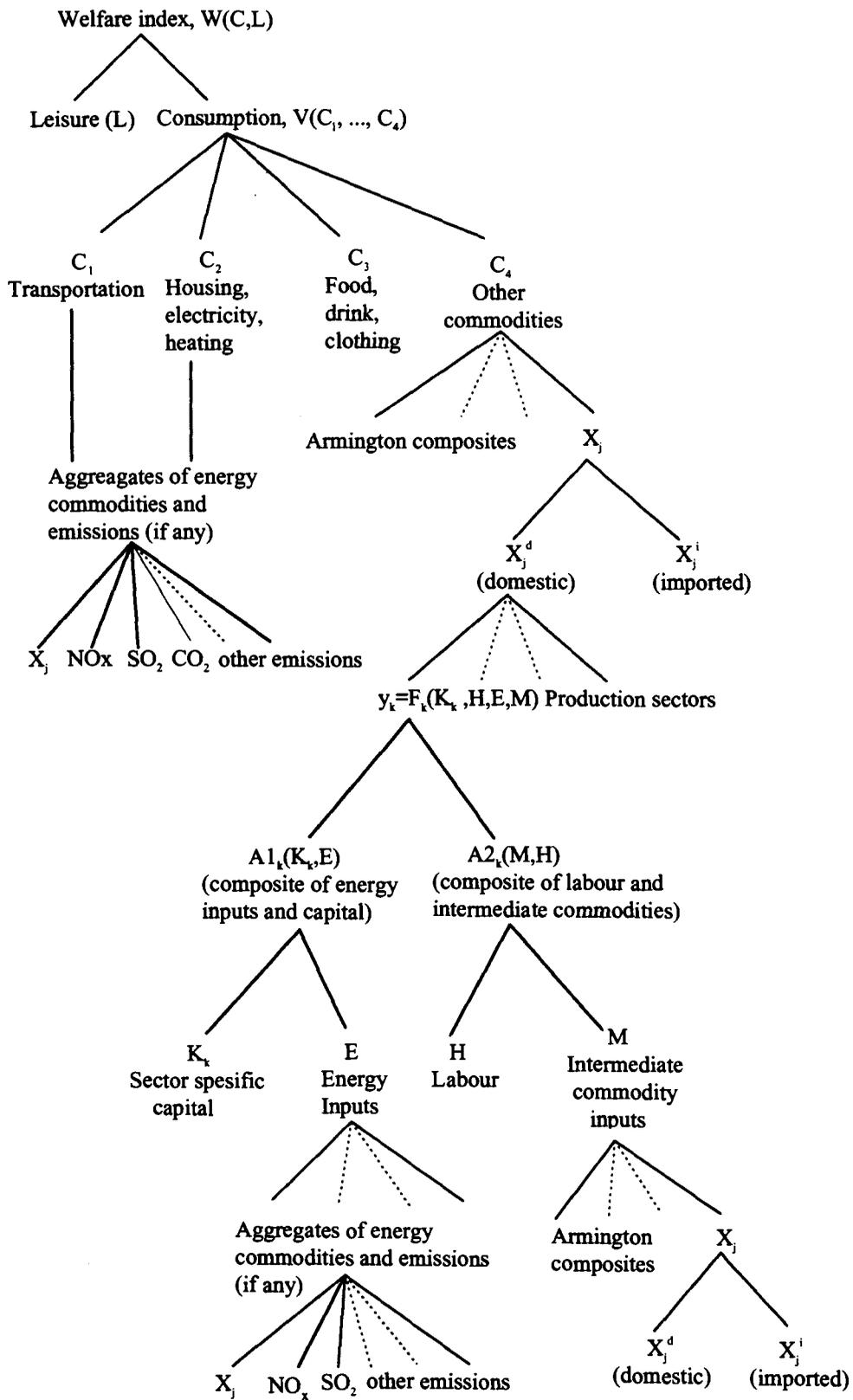
**Table A1. 2 CO<sub>2</sub> generating inputs in the model**

Gasoline
Other refined petroleum products (including diesel and various heating oils)
Coal/coke
Crude oil/gas
Other oil- and coal-based products (e.g. petrol-coke for primary aluminum smelting)
Limestone

**Table A1.3 Emission components included in the model**

<b>Component</b>	<b>Main sources</b>	<b>Negative Consequences</b>
CF <sub>4</sub> and C <sub>2</sub> F <sub>6</sub> (Perfluorized carbons)	Aluminum production	Increases the greenhouse effect
CH <sub>4</sub> (Methane)	Enteric fermentation Animal wastes Evaporation from landfills Production and transportation of crude oil, gas and coal Combustion of wood and fossile fuels	Increases the greenhouse effect Indirect effect: Contributes to increased level of trophospheric ozone (ozone near the ground level), which has negative consequences for vegetation and materials
CO (Carbon monoxide)	Esp. non-complete combustion processes - Transportation at land and sea - Wood fuel Process-related emissions in production of: Primary metals, Chemicals	Health: Binds red corpuscles, and thus reduces oxygenating capacity Indirect effect: Increased level of trophospheric ozone
CO <sub>2</sub> (Carbon dioxide)	All kinds of combustion processes Major process-related emissions in prod. of: Ferro-alloys, Aluminum, Fertilizers, Cement	Increases the greenhouse effect
NM VOC (Non-methane volatile organic compounds)	Evaporation from extraction, transportation and distribution of crude oil, gas, and gasoline. Evaporation from use and production of solvent/paint Mobile and stationary combustion (esp. wood fuel)	Health: Carrier of substances that may cause cancer, esp. PAH and benzen Indirect effect: Increased level of trophospheric ozone
NO <sub>x</sub> (Nitrogen oxides)	Combustion processes in general, esp. transportation at land and sea Process-related emissions in production of: Fertilizers, Primary metals	Health: Respiratory-passage-diseases Nature: Acidification of waters and forests Materials: Corrosion Indirect effect: Increased level of trophospheric ozone
N <sub>2</sub> O (Nitrous oxide)	Production and use of fertilizers Combustion of fossile fuels	Increases the greenhouse effect Reduces the stratospheric ozone layer
SO <sub>2</sub> (Sulphur dioxide)	Combustion of fossile fuels, esp. heavier oil distillates Process related emissions in production of: Primary metals, chemicals	Health: Respiratory-passage-diseases Nature: Acidification of waters and forests Materials: Corrosion

**Figure A1. 1 Flow of production factors and consumption goods**



## Appendix 2. Calibration of $U(C,L)$

Let  $C$  and  $L$  denote an aggregate index for commodity consumption and the volume of leisure respectively. The function  $U(C,L)$  is used for two distinct purposes in the model:

- 1) To derive aggregate households' demand for consumption goods and leisure, and thereby determine the equilibrium, and
- 2) to evaluate the (household) welfare in the resulting equilibrium.

Thus there are two issues involved when calibrating such a function:

- 1) We want the derived demands to have empirically plausible elasticities, and
- 2) we would like to obtain a reasonable weighing of the two components  $C$  and  $L$ .

Household income stems from endowments ( $E$ ) of the commodities time, capital and entitlements to transfers from the government<sup>1</sup>. Let  $p_i$  denote price net of tax for commodity  $i$ . Full (endowment) income is then

$$(1) \quad I = \sum_i p_i E_i.$$

Household demand is derived from utility maximisation

$$(2) \quad \text{maximise } U(\mathbf{X}) \text{ subject to } \mathbf{P}\mathbf{X} = I,$$

where  $\mathbf{X}$  and  $\mathbf{P}$  denote vectors of consumed commodities and their prices respectively, and where  $P_i = p_i(1+t_i)$ . The solution to (2) is a set of consumption-demand functions

$$(3) \quad X_i = F_i(\mathbf{P}, I).$$

When the household has an endowment of commodity  $i$ , its market demand function is

$$(4) \quad D_i = X_i - E_i = F_i(\mathbf{P}, I) - E_i$$

Households' preferences in our model are described by a multi-level Constant Elasticity of Substitution (CES) function. Each level (or aggregate) is described by

$$(5) \quad y = CES(\mathbf{x}; \mathbf{a}, r) = \left( \sum_i a_i x_i^r \right)^{\frac{1}{r}}$$

where  $\mathbf{x}$  is a vector of arguments,  $\mathbf{a}$  is a vector of distribution parameters and  $\sigma = 1/(1-r)$  is the elasticity of substitution. The aggregation-structure of  $U$  is shown in Figure 1. There are four (aggregate) consumer goods<sup>2</sup> that are aggregated into  $C$ , which finally is combined with leisure  $L$ . An interpretation, in terms of expenditure decisions, is that one first makes the labour-leisure choice, and then, with disposable income determined, decides on how to spend income on various commodities.

The parameters of (5), i.e.,  $\mathbf{a}$  and  $r$ , can be calibrated based on empirical facts on labour supply and market demand for various commodities.

### Labour supply

Let  $T = -D$  denote the supply of labour. The wage elasticity of labour supply, denoted  $e_w$ , can then be written

---

<sup>1</sup> This presentation is adapted from Kittelsen, S.A.C. (1992): Kalibrering av hushold og regioner i MISM0D-88, Working Paper No. 26, SNF, Bergen, 1992.

<sup>2</sup> These are «Transportation», «Housing, electricity and heating», «Food, drink & clothing», and «Other commodities». The first and second aggregate contain direct energy consumption, while the third and fourth aggregate only indirectly do so, cf. Figure A1.1.

$$(6) \quad e_w = -\frac{X}{T} [-1 + (1 - \sigma)(1 - \theta) + \Gamma],$$

where  $\sigma$  is the elasticity of substitution between C and L in the top-aggregate,  $\theta$  is leisure's expenditure share, X and T is leisure consumption and labour supply (respectively), and  $\Gamma$  is time's share in full (endowment) income. Given an estimate of  $e_w$ , we can freely stipulate either  $\sigma$  or the endowment of time,  $E=T+X$ . Table 1 provides resulting values for the elasticity of substitution,  $\sigma$ , for some alternative values of  $e_w$  and E.

Although estimates of the wage elasticity range from negative numbers to more than +1, there seems to be a consensus that it is positive, but probably fairly low. In a recent Norwegian analysis, Aaberge, Dagsvik and Strøm (1990)<sup>3</sup>, report wage elasticities of 0.18 and 1.0 for males and females respectively.

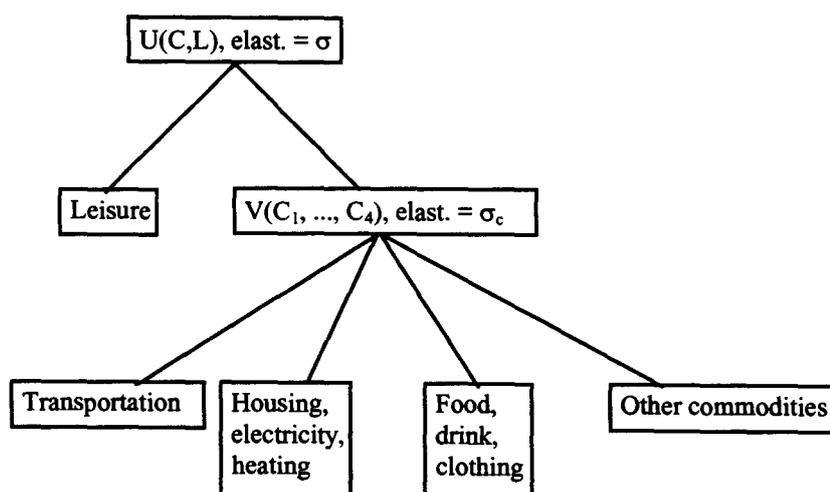
The higher the wage elasticity, the more willing the worker is to shift consumption into leisure. Because our argument is that leisure should be included in the analysis, we do not want to assume a too high wage-elasticity and thereby overdo the argument. Hence, for the base case we assume the wage elasticity to be 0.3. We do provide, however, results based on the alternative values of 0.0 and 0.6, see figure 5 in the paper.

Similar considerations lead us to stipulate a low time endowment,  $E = 500$ , so that leisure consumption (by (4)) becomes low, i.e., leisure amounts to about 20 percent of total expenditure on C and L. Putting more emphasis on leisure consumption by increasing the time endowment to  $E = 700$ , implies a 60-40 weighing of C and L in terms of expenditure shares. This only produces minor changes in the development of the welfare index  $U(C,L)$ , however.

**Table A2. 1** Combinations of time-endowment, elasticity of substitution and supply-elasticity

		Wage elasticity of labour		
		$e_w = 0$	$e_w = 0.3$	$e_w = 0.6$
Endowment of time	500	$\sigma = 0.9$	$\sigma = 1.6$	$\sigma = 2.35$
	700	$\sigma = 0.88$	$\sigma = 1.3$	$\sigma = 1.73$

**Figure A2. 1** Illustration of the CES preference structure



<sup>3</sup> Aaberge, R, J.K. Dragsvik, and S. Strøm: Labour supply, income distribution and excess burden of personal taxation in Norway, Discussion Paper No. 54, Statistics Norway, 1990.

## Appendix 3. Calibration of externality costs

### 1. Introduction

The main data sources for our calibrations are two analyses of externality costs from air pollution and road traffic performed in the research department in Statistics Norway, Alfsen et al. (1992) and Brendemoen et al. (1992)<sup>1</sup>. In these two papers, the externality cost data are largely identical, but the analyses performed are different. The total and marginal cost data are based on several publications from Statistics Norway, SFT and TØI<sup>2</sup>, where traffic- and pollution-related costs have been estimated. For costs from traffic-accidents, we have used updated cost estimates from TØI (1993).

In Alfsen and in Brendemoen, the cost data are given as total and marginal cost numbers in NOK at 1990-prices. These cost data are not integrated into the model analyses (using the models MODAG and MSG-5 respectively) of emission-reductions. Rather, the analyses of emission reductions are performed without incorporating any externality effects. In the next step, the activity levels in the computed equilibria are used to calculate changes in the levels of externality costs in a separate sub-model. Finally, the macroeconomic costs from emission reductions calculated in the traditional models are compared to the benefits from reduced level of externality costs.

Our use of the data differ from this procedure in several respects. Firstly, we have «translated» the value terms into consumption of commodities like e.g. labour, where this seems plausible. Secondly, we have used a non-linear rather than a linearised cost function over the range of output volumes. Thirdly, some of the minor cost components in Alfsen's and Brendemoen's work are omitted in our calibrations. The cost data and the assumptions underlying our «translations» of the data are briefly described below.

### 2. The original data in 1990 NOK

#### 2.1. Traffic-related costs.

Traffic-related costs are grouped into four different categories: i) Noise, ii) Road damage, iii) Congestion, and iv) Accidents. These marginal cost categories are given as linear functions  $b_i \Delta V$ , where  $i$  denotes the category and  $\Delta V$  denotes changes in the volume of road traffic.  $V$  is computed from total fuel-consumption (gasoline and diesel) in the different scenarios.

Ad iv), accidents: The total yearly cost from traffic accidents contains four separate components: a) Medical treatment, b) Loss of production due to injured and dead persons, c) Material costs (damages on vehicles), and d) Administrative costs (e.g. insurance). Further, there is a description of which group that pays for each of these cost-components: Directly involved persons, relatives to those directly involved, other private agents, and public sector. See TØI (1988) and (1993)<sup>3</sup>.

#### 2.2. Health costs from air pollution.

Health costs from air pollution are linked to the emission volumes of the following four components (indexed by  $j$ );  $\text{NO}_x$ ,  $\text{SO}_2$ ,  $\text{CO}$ , and Particulates. Changes in health damages from emissions of component  $j$  is computed by  $b_j (a_j^M \Delta M_j + a_j^S \Delta S_j)$ , where  $b_j$  is the cost of increases in the number of people exposed to pollution of component  $j$  above the ambient standard,  $a_j^M$  and  $a_j^S$  are the shares of emissions of component  $j$  from mobile and stationary sources (respectively) that causes health damages (in % of total emissions of component  $j$ ), and  $\Delta M_j$  and  $\Delta S_j$  are changes in emission levels of component  $j$  from mobile and stationary sources respectively (in metric tons).

#### 2.3. Corrosion costs

These costs are defined as a linear function of the emission volume of  $\text{SO}_2$ , namely  $b_{\text{SO}_2} \Delta \text{SO}_2$ , where  $b_{\text{SO}_2}$  is measured in thousand 1990-NOK per metric ton  $\text{SO}_2$ .

#### 2.4. Acidification of water and forests

Let  $w$  denote water and  $f$  denote forests. Acidification costs are defined as linear functions  $b_w \Delta(\text{SO}_2 + \text{NO}_x)$  and  $b_f \Delta(\text{SO}_2 + \text{NO}_x)$ , where  $b_w$  and  $b_f$  are measured in thousand 1990-NOK per metric ton  $\text{SO}_2$  and  $\text{NO}_x$ .

---

<sup>1</sup> For short referred to as Alfsen and Brendemoen respectively in the rest of the appendix.

<sup>2</sup> SFT; Norwegian Pollution Control Authority,  
TØI; Norwegian Institute of Transport Economics.

<sup>3</sup> TØI-Working Paper no. 877/1988 and TØI-Report no. 193/1993.

### 3. Our use of the data in the model

Firstly, the costs from corrosion of materials and acidification of water and forests are less than 1.5% of the total externality costs. We have omitted these two cost categories from our calculations. (This does not mean that corrosion and acidification as such are unimportant. The small cost numbers are due to the fact that Norwegian emissions of  $\text{NO}_x$  and  $\text{SO}_2$  are very small relative to «imported» concentrations from other European countries.) Secondly, the cost numbers presented by Alfsen and Brendemoen are estimates associated with a high degree of uncertainty regarding the «true» values. Therefore, the authors present three values for each parameter: Low, medium (mean value), and high, and report results from the whole range of these parameter values using Monte Carlo simulations. We use only the mean values in our calculations.

#### 3.1 Traffic related costs

Since only commercial road-traffic are direct outputs in the model (i.e. produced and purchased model commodities), we need a proxy-variable for the *total*<sup>4</sup> volume of road-traffic, denoted  $V$ . There are two fuels used for road vehicles: gasoline and diesel. However, diesel is not a specific commodity in the model. Rather it is contained in an aggregate of diesel and other (heavier) petroleum fuels, e.g. maritime fuels and fuels used for stationary combustion. Therefore, we assume that changes in gasoline consumption alone is a better proxy for changes in road-traffic than changes in gasoline and «other petroleum fuels».

##### 3.1.1 Noise

This cost component is treated as a pure «disutility»-cost. Hence, we compute the disutility from noise using the same linear function  $b_1 \Delta V$  as described in section 2.1. (With the difference that  $V$  is the chosen proxy for traffic-volume in our model.)

##### 3.1.2 Road damage

In Norway, yearly repair costs due to road damage are financed by the public sector. In the basic version of the model, total public consumption,  $G$ , is exogenously stipulated to 150 billions in fixed 1991-NOK prices ( $G^*$ ). Note that public consumption is not represented in the utility- or welfare measure for the household. In the «externality»-scenarios, we establish a link in the model between traffic volume  $V$  and the level of public consumption;  $G = G^* + b_2 p_G \Delta V$ , where  $b_2$  is defined in section 2.1. above, and  $p_G$  is the benchmark-price for public consumption in the model.

##### 3.1.3 Congestion

Traffic volumes at the benchmark activity levels imply time usage for consumers/workers. In the model, «time» is the consumer's endowment of leisure/labour. Queue theory models and studies of road congestion conclude that time consumption is a convex function of the traffic volume due to congestion. Therefore, we use a convex time-loss function:  $\Delta T = a (\Delta V)^b$ . The parameter  $b$  is arbitrarily set at 3, whereafter  $a$  is chosen so that our specification gives the same marginal loss from traffic congestion as the original marginal cost  $b_3 \Delta V$  in section 2.1 (at the benchmark level of traffic volume). The difference, of course, is that our convex specification implies that the marginal cost decreases as the traffic volume falls below the benchmark level.

##### 3.1.4 Traffic accidents

In the model, expenditures on e.g. insurance and car repairs are endogenously determined. Hence, when the demand for gasoline (i.e. road traffic) go down, these administrative and material expenditures also diminishes. Although the value of this endogenous response is less than the costs referred to in section 2.1 above, we do not calculate any «extra» benefits from reduced administrative and material costs. What we do include in our calculations of externality costs from accidents, is the part of the traffic-related costs that are i) reduced participation in the labour force due to damages and deaths, and ii) public expenditures on medical treatment. For both these cost components, we assume linearity in the traffic volume,  $V$ , and calibrate loss-functions,  $\Delta T = \alpha \Delta V$ , and  $\Delta G = \beta \Delta V$  respectively, in the same manner as described in 3.1.2 and 3.1.3 above.

#### 3.2 Health costs

First of all, the component «Particulates» is not incorporated in our model. Its contribution to health costs is small relative to  $\text{NO}_x$  and  $\text{SO}_2$  (ca. 4.5%), so its omission is of minor importance for the analysis. CO is included in the model, but also this component contributes so little to health costs that it is omitted in our analysis.

---

<sup>4</sup> In addition to buying outputs of commercial road transport activities, both the household and the production sectors use their own road vehicles.

As seen from section 2.1, the parameters  $a_j^M$  and  $a_j^S$  are crucial for the implied marginal cost numbers. These parameters give the shares of total emissions from mobile and stationary combustion sources respectively that contribute to pollution concentrations above the threshold for health-damages. Such high concentrations occur mainly in the most densely populated (and trafficated) areas. The mean values presented in Alfsen and in Brendemoen are calibrated based on the assumption that such high concentrations only take place in the five largest Norwegian cities, accounting for 25% of the Norwegian population.

We do not use parameters corresponding to  $a_j^M$  and  $a_j^S$  to establish a link between total emissions and health-damaging concentrations. Rather, for each model activity, we split emissions of  $\text{NO}_x$  and  $\text{SO}_2$  into emissions that contribute and do not contribute to health-damaging concentrations directly. This should increase the empirical validity of the performed calculations. For example, we know that emissions from ocean fisheries and aluminium smelting do not contribute to high pollution concentrations in the largest Norwegian cities. Although the model is not explicitly regionalised, we have from other sources information about where the various sector's production activities are located. For the highly aggregated sectors, we have used the share of the population that lives in the five largest cities (25%).

Also for health costs, we assume a convex damage function,  $d = a_1 (\text{HNO}_x + a_2 \text{HSO}_2)^b$ , where  $b$  is arbitrarily set at 2 and  $a_1$  and  $a_2$  are chosen so that marginal cost at the benchmark emission volumes equals Alfsen's and Brendemoen's constant marginal cost. In the case of health costs, the data itself tells little about how the aggregated «health cost»-estimate is calculated. Rather, the cost estimate contains a mixture of real costs and «willingness to pay»-estimates. In principle, health damages have at least the three following consequences: i) reduced labour participation (both because workers become ill themselves and if their children become ill, ii) medical treatment costs, and iii) the «disutility» from being ill. Lacking data for how the cost numbers are composed, we assume that the aggregated health cost splits equally into these three components. Hence, Alfsen's and Brendemoen's health cost is integrated in the model with one third of the cost affecting the time endowment for the consumer, one third affecting public resource use (public medical treatment), and the last third as subjective utility loss.

## References:

- Alfsen, K., A. Brendemoen, and S. Glomsrød (1992): Benefits of climate policies: Some tentative calculations, Discussion Paper No. 69, Statistics Norway.
- Ballard, C.L. and S.G. Medema (1993): Efficiency effects of taxes and transfers, *Journal of Public Economics* 52, 199-216.
- Bovenberg, A. L. and R. A.de Mooij (1994): Environmental levies and distortionary taxation, *American Economic Review* 94, 1085-1089.
- Brekke, K. A. and H. A. Gravningsmyhr (1994): Adjusting NNP for instrumental or defensive expenditures. An analytical approach, Discussion Paper No. 134, Statistics Norway.
- Brendemoen, A., S. Glomsrød, and M. Aaserud (1992): Miljøkostnader i makroperspektiv, Report No. 17, Statistics Norway.
- Brendemoen, A. and H. Vennemo (1994): A climate treaty and the norwegian economy: A CGE assessment, *The Energy Journal*, Vol. 15, No. 1, 77-93.
- Carraro, C., M. Galeotti, and M. Gallo (1994): Environmental taxation and unemployment: Some evidence on the double dividend hypothesis in Europe, Paper prepared for the NBER-University of Turin- FEEM-Conference on «Market Failures and Public Policy», Turin, May 1994.
- Cobb, J.B. and H.E. Daly (1989): *For the Common Good: Redirecting the Economy toward Community, the Environment, and a Sustainable Future*, Beacon Press, Boston,
- Goulder, L.H. (1995a): Effects of carbon taxes in an economy with prior tax distortions: An intertemporal general equilibrium analysis, *Journal of Environmental Economics and Management* 29, 271-297.
- Goulder, L.H. (1995b): Environmental taxation and the double dividend: A reader's guide, *International Tax and Public Finance* 2, 157-183.
- Hoel, M. (1991): Efficient international agreements for reducing emissions of CO<sub>2</sub>, *The Energy Journal* 12, 93-107.
- Håkonsen, L. (1995): Optimal commodity taxation with a binding CO<sub>2</sub>-restriction, Discussion Paper No. 5/95, Department of Economics, Norwegian School of Economics and Business Administration, Bergen.
- Håkonsen, L. og L. Mathiesen (1994): Integrasjon av skadevirkninger i generell likevekt, Working Paper No. 76, SNF, Bergen.
- Manne, A. and L. Mathiesen (1994): The impact of unilateral OECD carbon taxes upon the location of aluminium smelting, *International Journal of Global Energy Issues*, 6, 52-61.
- Manne, A. and R. Richels (1994): The cost of stabilizing global CO<sub>2</sub> emissions: A probabilistic analysis based on expert judgments, *The Energy Journal*, 15, 31-56.
- Mathiesen, L. (1991): Analyse av energibruk og CO<sub>2</sub>-utslipp i norsk økonomi i år 2000, Report No. 54, SNF, Bergen.
- Mathiesen, L. (1993): Analyse av endringer i norsk faktoravgifter, In Anderesen, C., L. Mathiesen, and J.G. Sannarnes: Indirekte skatter i Norge. En analyse med fokus på næringsvirksomhet, Report No. 52, SNF, Bergen.
- Nordhaus, W. D (1991): The cost of slowing climate change: a survey, *The Energy Journal*, 12, 37-65.

Nordhaus, W. D. (1993): Optimal greenhouse-gas reductions and tax policy in the «DICE» model, *American Economic Review* (Papers and Proceedings) 83, 313-317.

Nordhaus, W. D. and J. Tobin (1972): Is growth obsolete?, National Bureau of Economic Research General Series 96.

Park, C.C. (1987): *Acid Rain: Rhetoric and Reality*, London: Methuen.

Shackleton, R., M. Shelby, A. Cristofaro, R. Brinner, J. Yanchar, L. Goulder, D. Jorgenson, P. Wilcoxon, and P. Pauly (1992): The efficiency value of carbon tax revenues, unpublished manuscript, U.S. Environmental Protection Agency.