labor&welfareSTATE

Optimal redistributive taxation in a multi-externality model

by

Paul ECKERSTORFER

Working Paper No. 1107

July 2011

Supported by the Austrian Science Funds

FWF

The Austrian Center for Labor Economics and the Analysis of the Welfare State

JKU Linz
Department of Economics
Altenberger Strasse 69
4040 Linz, Austria
www.labornrn.at

paul.eckerstorfer@jku.at
phone +43 (0)70 2468 - 8246, - 9821 (fax)

Optimal redistributive taxation in a multi-externality model

by

Paul Eckerstorfer*

July 26, 2011

Abstract

This paper extends the previous literature on optimal redistributive taxation in the presence of externalities to a multi-externality setting. While taxes on income and on 'clean' commodities are still unaffected by the externalities, which confirms previous results, I find that the existence of more than one externality-generating commodity has important implications for the optimal Pigouvian tax rates. In general the Pigouvian parts of taxation depend also on the externalities induced by the consumption of the other commodities, implying that the interdependence of the externality-generating commodities is relevant for tax policy.

Keywords: Optimal Taxation, Externalities

JEL Classification: D82, H21, H23, H24

*Department of Economics, Altenbergerstr. 69, 4040 Linz, Austria.
E-mail: paul.eckerstorfer@jku.at

1 Introduction

The main contribution of this note is to extend the literature on optimal redistributive taxation in the presence of externalities (e.g. Pirttilä and Tuomala 1997, Cremer et al. 1998, Kopczuk 2003, Micheletto 2008) to a multi-externality setting. One important result from this literature is that the 'additivity property', first discovered by Sandmo (1975), also holds in a more general model of the Mirleesian type.[1] The additivity property consists of two components. First, the presence of an externality only affects the tax rate on that particular good (which generates the externality). Second, the internalizing part of taxation enters the tax formula additively. This result is quite remarkable and has important policy implications as it states that an externality is best addressed by taxing directly that particular good, while the rest of the tax system should remain unaffected by the externality.

However, one important shortcoming of the previous literature is the assumption that there is only one externality-generating commodity. But in reality the consumption of many commodities causes externalities. Thus, the aim of this note is to generalize the optimal tax problem (income and commodity taxes) to a multi-externality framework and to clarify the optimal tax structure in such a context. I show that the 'additivity property' remains valid with respect to the income tax and also with respect to the taxes on the non-externality generating goods, as they still remain unaffected by the externalities. However, I find that in general a tax on an externality-generating commodity also depends on the externalities induced by the consumption of the other commodities, which violates the 'additivity property'. If the level of an externality increases (decreases) the demand for another externality-generating commodity this is an additional argument to increase (decrease) the tax rate on that commodity. Thus, such interdependences between the externality-generating commodities should be taken into account when designing optimal Pigouvian tax rates. Applications where this interdependence of the externalities might be of particular relevance include alcohol and cigarette consumption or car and gasoline consumption.

The paper is organized as follows. Section 2 introduces the model and section 3 presents the main results on the optimal tax structure. Finally, section 4 concludes. The derivations of the main results are provided in the Appendix.

[1] However, Micheletto (2008) has shown that for this result to be valid it is essential that different types are equally effective as externality generating units, which is the case if the externality is of the 'atmospheric' type, meaning that the externality is created by the total consumption of a good.

2 The model

Consider an economy consisting of two types of individuals $i = L, H$, who differ in earning ability $\omega_L < \omega_H$. The size of the population is normalized to one and π_i represents the fraction of individuals of type i. Individuals provide labor supply l_i and earn gross income $z_i = \omega_i l_i$. As is common in the optimal taxation literature (e.g. Mirrlees, 1971) the tax administration can only observe gross income z_i, while earning abilities and working time of an individual i are assumed to be private information. Gross income is subject to a nonlinear income tax, and consumers allocate their net income x_i over $n+m$ consumption goods. Commodities are produced by a linear technology with labor as the only input to production. Quantities are chosen such that all producer prices are equal to one.

Let the vector of the first n commodities consumed by an individual i be denoted by $c_i = (c_{1i}, c_{2i}, ..., c_{ni})$ and the vector of the other m commodities by $d_i = (d_{1i}, d_{2i}, ..., d_{mi})$. The consumption of each of the m commodities creates a negative externality, whereas the consumption of the first n commodities does not. More precisely an externality is created by the total consumption of a commodity m, i.e. the level of an externality is given by $E_k = \sum_i \pi_i d_{ki}$, $k = 1, ..., m$.[2] That is, there are m different externalities in the model represented by the vector $E = (E_1, E_2, ..., E_m)$. Individuals have identical preferences described by the strictly concave utility function $u(c_i, d_i, l_i, E)$ with first partial derivative being positive with respect to c_{ji} and d_{ki} and negative with respect to l_i and E_k. I restrict the analysis to cases where $\partial u / \partial E_k < 0$ only for expositional reasons but allowing for positive externalities would not cause any complication. In addition, following previous work (e.g. Sandmo, 1975), I assume that individuals behave atomistically, i.e. they do not take into consideration the influence of their own consumption on the level of the externalities.

The individuals' maximization problem is analyzed in two steps. In a first step, individuals allocate a fixed amount of net income x_i over the consumption goods. Let consumer prices be denoted by $p_j = 1 + \tau_j$ for the first n commodities $j = 1, ..., n$ and by $q_k = 1 + t_k$ for the other m commodities $k = 1, ..., m$. Hence, the government can impose proportional commodity taxes τ_j and t_k, respectively, on each of the $n+m$ commodities. It is well-known that in such a tax system one tax is redundant, thus without loss of generality τ_1 is set equal to zero, i.e. $p_1 = 1$. The first stage of the maximization problem gives conditional indirect utility

$$v_i(x_i, z_i, p, q, E) \equiv \max_{c_{1i},..,c_{ni},d_{1i},..,d_{mi}} \left\{ u(c_i, d_i, z_i/\omega_i, E) \left| \sum_{j=1}^{n} p_j c_{ji} + \sum_{k=1}^{m} q_k d_{ki} \leq x_i \right. \right\}, \tag{1}$$

[2] Meade (1952) has termed this type 'atmospheric' externalities.

and conditional demand functions

$$c_{ji} = c_{ji}(x_i, z_i, p, q, E), \qquad j = 1, ..., n, \tag{2}$$

$$d_{ki} = d_{ki}(x_i, z_i, p, q, E), \qquad k = 1, ..., m. \tag{3}$$

Observe that in general the demand for all commodities depends on the level of the externalities E_k, $k = 1, ..., m$.

In a second step individuals choose their optimal labor supply by maximizing conditional indirect utility subject to the budget equation $x_i = z_i - T(z_i)$, where $T(z_i)$ denotes the nonlinear income tax function. This yields the well-known expression for the implicit marginal income tax rate $T'(z_i)$, which is given by

$$T'(z_i) = 1 + \frac{\partial v_i/\partial z_i}{\partial v_i/\partial x_i}. \tag{4}$$

The objective of the government is to design a tax system, consisting of a general income tax and proportional commodity taxes, which maximizes a utilitarian social welfare function given the informational structure of the model and an exogenous revenue requirement. The problem of finding the optimal income tax schedule can equivalently be stated by determining the optimal gross and net income bundles x_i, z_i for each type. Thus, the optimal income tax for the two types of individuals is determined implicitly as the difference $z_i - x_i$, $i = L, H$. Note that the available tax instruments are completely determined by the information structure of the model. Since earning abilities are not observable to the government (only the distribution of types is known), type specific first-best lump-sum taxes are not feasible. Therefore the government has to use a general income tax as a second-best instrument. In addition, consumption is assumed to be observable only in the aggregate, while individual consumption levels are private information. Thus, type specific nonlinear commodity taxes are not feasible either.

The utilitarian social welfare function reads

$$\max_{x_i, z_i, \tau_2, .., \tau_n, t_1, .., t_m, E_1, .., E_m} f_L v_L(x_L, z_L, p, q, E) + f_H v_H(x_H, z_H, p, q, E), \tag{5}$$

where f_L and f_H, with $f_L \geq f_H \geq 0$, represent the weights of the two types of individuals including the fractions π_L and π_H. The agent monotonicity condition is assumed to hold, meaning that $MRS^L_{zx} > MRS^H_{zx}$ at any vector (x, z), where MRS^i_{zx} is defined as $MRS^i_{zx} \equiv -(\partial v_i/\partial z_i)/(\partial v_i/\partial x_i)$. This implies that for any income tax function the high-able individual does not choose to earn less income than the low-able.

The resource constraint is given by

$$\pi_L(z_L - x_L) + \pi_H(z_H - x_H) + \sum_{j=2}^{n} \tau_j(\pi_L c_{jL} + \pi_H c_{jH}) + \sum_{k=1}^{m} t_k(\pi_L d_{kL} + \pi_H d_{kH}) \geq g, \tag{6}$$

i.e. tax revenues have to be raised to finance exogenous public spending g. In addition the government's choice of optimal taxes is restricted by a self-selection constraint, which reads

$$v_H(x_H, z_H, p, q, E) \geq v_H(x_L, z_L, p, q, E). \tag{7}$$

It assures that the allocation implemented by the government is such that the H type has no incentive to mimic or imitate the L type (by working less). The constraint that the L type does not mimic the H type can be neglected because it is not binding in the optimum, as the analysis is restricted to cases, where the government wants to redistribute from high- to low-ability persons. To abbreviate notation indirect utility of the mimicker is denoted by $v_H[L]$ and consumption of the mimicker by $c_{jH}[L]$ and $d_{kH}[L]$. In addition, the levels of the externalities are taken into consideration as separate (equality) constraints,

$$E_k = \sum_i \pi_i d_{ki}, \qquad k = 1, ..., m. \tag{8}$$

The Lagrange multipliers for the resource and the self-selection constraint are denoted by λ and μ, respectively, and for the constraints concerning the externality levels by γ_k. The first-order conditions for the maximization problem are provided in the Appendix.

3 Optimal tax structure

The optimal tax structure can be derived from the first-order conditions of the social planner's maximization problem. Before I present the results for the optimal commodity and income tax rates, I discuss in more detail the shadow prices of the externalities measured in terms of tax revenues γ_k/λ, $k = 1, ..., m$, as they will be highly relevant for the analysis of the optimal tax structure later on. This discussion is closely related to the one in Pirttilä and Tuomala (1997), which I generalize for the purposes of this study.

Let the marginal rate of substitution between E_k and x_i be defined by

$$MWP_{ki} \equiv -\frac{\partial v_i/\partial E_k}{\partial v_i/\partial x_i}. \tag{9}$$

It can be interpreted as the marginal willingness to pay of an individual i to reduce E_k by one unit. Note that MWP_{ki} is positive as we assumed E_k to be a negative externality. From the first-order conditions one can then derive an expression for γ_k/λ which is displayed in Lemma 1.

Lemma 1: In the social optimum the shadow prices of the externalities measured in terms of the government's tax revenues are given by

$$\begin{aligned}\frac{\gamma_k}{\lambda} &= \frac{1}{1-\sum_i \pi_i \frac{\partial d_{ki}^{com}}{\partial E_k}}\Big(\sum_i \pi_i MWP_{ki} - \frac{\mu}{\lambda}\frac{\partial v_H[L]}{\partial x_L}(MWP_{kH}[L] - MWP_{kL}) \\ &\quad -\sum_{j=2}^{n}\sum_i \pi_i \tau_j \frac{\partial c_{ji}^{com}}{\partial E_k} - \sum_{s=1}^{m}\sum_i \pi_i t_s \frac{\partial d_{si}^{com}}{\partial E_k} \\ &\quad +\sum_{s=1}^{k-1}\sum_i \frac{\gamma_s}{\lambda}\pi_i \frac{\partial d_{si}^{com}}{\partial E_k} + \sum_{s=k+1}^{m}\sum_i \frac{\gamma_s}{\lambda}\pi_i \frac{\partial d_{si}^{com}}{\partial E_k}\Big),\end{aligned} \tag{10}$$

for $k = 1, ..., m$, and where compensated demand for the commodities of an individual i is denoted by c_{ji}^{com} and d_{ki}^{com}. If $\frac{\partial d_{si}^{com}}{\partial E_k} \neq 0$ for $s \neq k$, then γ_k/λ depends on γ_s/λ.

Proof: The derivation of equation (10) is provided in the Appendix.

The shadow prices γ_k/λ can be interpreted as the social harm or gain of a specific externality measured in terms of tax revenues. Lemma 1 states that in general the shadow prices depend on each other, i.e. the social harm or gain induced by the consumption of commodity k also depends on the shadow prices of the externalities induced by the consumption of the other commodities s ($s \neq k$), at least if E_k affects compensated demand for these commodities. This can be seen from the last two terms on the RHS of (10).

Let me also briefly discuss the other parts of the RHS of (10) in order. The first term $1/(1 - \sum_i \pi_i \frac{\partial c_{ki}^{com}}{\partial E_k})$ captures the impact of the level of the externality on compensated demand for commodity k. If compensated demand for good k increases (decreases) with E_k, this term is larger (smaller) than 1. Hence, the shadow price is larger if an increase in the level of the externality has a positive effect on the demand for the good which generates that externality. The first-term within parenthesis is the marginal willingness to pay of all individuals to avoid the externality. It can be considered as the direct negative effect of the externality and as its sign is always positive, it increases the value of the shadow price. Then there is also an effect related to the self-selection constraint, represented by the second term within parenthesis. The sign of this effect is ambiguous. It depends on whether the mimicker or the L type has a higher marginal willingness to pay to avoid the externality. As the only difference between them is labor supply provided the sign depends on $\partial MWP_{ki}/\partial l_i \gtreqless 0$. Finally, the two terms in the second line of (10), describe the impact of the externality on government's tax revenues. The sign of these effects is again ambiguous, as the reaction of compensated demand due to a change of E_k can have either sign. Altogether γ_k/λ can be positive or negative, although E_k is a negative externality. That is, an increase in the level of the externality could also generate a social gain. However, the case that an increase of the externality is socially harmful ($\gamma_k/\lambda > 0$) appears more plausible because of the direct negative effect.

Now I present the main results on the optimal tax structure. As the focus is on the Pigouvian role of commodity taxation, preferences are assumed to be weakly separable in labor supply and consumption. It is well-known from the literature that in the absence of externalities commodity taxation is not needed in the presence of a nonlinear income tax if preferences are weakly separable (Atkinson and Stiglitz, 1976). Hence, the potential role of commodity taxes in the model is exclusively to correct for the externalities. It can be shown that the essence of my results does not depend on the separability of the utility function.[3]

In the Appendix I show that in the model optimal commodity tax rates have to satisfy

$$A \begin{pmatrix} \tau_2 \\ \vdots \\ \tau_n \\ t_1 \\ \vdots \\ t_m \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^m \sum_i \pi_i \frac{\gamma_k}{\lambda} \frac{\partial d_{ki}^{com}}{\partial p_2} \\ \vdots \\ \sum_{k=1}^m \sum_i \pi_i \frac{\gamma_k}{\lambda} \frac{\partial d_{ki}^{com}}{\partial p_n} \\ \sum_{k=1}^m \sum_i \pi_i \frac{\gamma_k}{\lambda} \frac{\partial d_{ki}^{com}}{\partial q_1} \\ \vdots \\ \sum_{k=1}^m \sum_i \pi_i \frac{\gamma_k}{\lambda} \frac{\partial d_{ki}^{com}}{\partial q_m} \end{pmatrix}, \tag{11}$$

[3]Of course optimal commodity tax rates look different if commodities are either complements or substitutes with leisure, as then they contain an additional effect on the self-selection constraint. However, the Pigouvian parts of taxation remain unaffected by these effects.

where

$$A = \begin{pmatrix} \sum_i \pi_i \frac{\partial c_{2i}^{com}}{\partial p_2} & \cdots & \sum_i \pi_i \frac{\partial c_{ni}^{com}}{\partial p_2} & \sum_i \pi_i \frac{\partial d_{1i}^{com}}{\partial p_2} & \cdots & \sum_i \pi_i \frac{\partial d_{mi}^{com}}{\partial p_2} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sum_i \pi_i \frac{\partial c_{2i}^{com}}{\partial p_n} & \cdots & \sum_i \pi_i \frac{\partial c_{ni}^{com}}{\partial p_n} & \sum_i \pi_i \frac{\partial d_{1i}^{com}}{\partial p_n} & \cdots & \sum_i \pi_i \frac{\partial d_{mi}^{com}}{\partial p_n} \\ \sum_i \pi_i \frac{\partial c_{2i}^{com}}{\partial q_1} & \cdots & \sum_i \pi_i \frac{\partial c_{ni}^{com}}{\partial q_1} & \sum_i \pi_i \frac{\partial d_{1i}^{com}}{\partial q_1} & \cdots & \sum_i \pi_i \frac{\partial d_{mi}^{com}}{\partial q_1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sum_i \pi_i \frac{\partial c_{2i}^{com}}{\partial q_m} & \cdots & \sum_i \pi_i \frac{\partial c_{ni}^{com}}{\partial q_m} & \sum_i \pi_i \frac{\partial d_{1i}^{com}}{\partial q_m} & \cdots & \sum_i \pi_i \frac{\partial d_{mi}^{com}}{\partial q_m} \end{pmatrix} \tag{12}$$

The (implicit) solution to this system of equations is given by

$$\tau_j = 0, \qquad j = 2, ..., n, \tag{13}$$

$$t_k = \frac{\gamma_k}{\lambda}, \qquad k = 1, ..., m. \tag{14}$$

It is unique if A is assumed to be non-singular. The optimal tax rates on commodities $j = 2, ..., n$ remain unaffected by the externalities, which confirms the validity of the 'additivity property' with respect to the tax rates τ_j. Hence, given weak separability the optimal tax rates on these 'clean' commodities are zero. On the other hand the optimal tax rates on the externality-generating commodities contain a Pigouvian element which is positive if the externality is socially harmful ($\gamma_k/\lambda > 0$). Looking at (14) one can see that for each $k = 1, ..., m$ the optimal t_k is equal to the shadow price of commodity k measured in terms of tax revenues. Thus, the first impression is that the 'additivity property', stating that the presence of an externality only alters the tax rate on that particular good, is still valid. But from Lemma 1 we know that the marginal social damages of the m externalities depend on each other. Hence, when taxing commodity k one has to take into account the effects of E_k on the demand for the other externality-generating commodities $s (s \neq k)$ and if these effects are nonzero the marginal social damages induced by the consumption of these commodities have an impact on the optimal t_k. For example, if the demand for commodity s increases with E_k this is an additional argument to increase t_k (provided $\gamma_s/\lambda > 0$), since reducing the level of the externality E_k also reduces the consumption of commodity s, and hence E_s.[4] Thus, if there is more than one commodity that generates an externality, the interdependence between these commodities should be taken into account.

Finally, I present the structure of the optimal income tax schedule. Combining (4) and the FOCs

[4]Essential for this argument is that demand for commodity k decreases if the price q_k increases, because then by increasing t_k one reduces E_k.

for x_i, z_i the optimal marginal income tax rates can be derived. They are given by

$$T'(z_H) = 0 \tag{15}$$

for the H type and by

$$T'(z_L) = -\frac{\mu}{\lambda \pi_L} \frac{\partial v_H[L]}{\partial x_L} (MRS_{zx}^H[L] - MRS_{zx}^L) \tag{16}$$

for the L type. One observes immediately that the marginal income tax rates do not depend on the externalities, i.e. the additivity property remains valid with respect to the income tax as well. Since the externalities do not affect the income tax schedule the optimal marginal income tax rates in our model are the same as those in the conventional mixed tax case when weak-separability is assumed. Thus, for a closer interpretation of these formulas the author refers to e.g. Edwards et al. (1994). Proposition 1 summarizes the results we obtained on the optimal tax structure.

Proposition 1: If $\frac{\partial c_{si}^{com}}{\partial E_k} \neq 0$ with $s \neq k$, the optimal tax rates t_k on the externality-generating commodities depend on all shadow prices γ_s/λ, $s = 1, ..., m$, i.e. on the marginal social harm or gain induced by each of the externality-generating commodities. This violates the 'additivity property' since the condition that an externality only affects the tax rate on that particular good no longer holds. The optimal tax rates on commodities $j = 2, ..., n$ and the income tax are unaffected by the externalities.

4 Conclusion

In this paper I study the optimal income and commodity tax structure in the presence of many externality-generating commodities, i.e. I drop the assumption made in previous contributions that there is only one externality-generating commodity. This allows me to study the interdependence between the externalities and to analyze possible implications for the optimal tax structure. I find that the income tax and the taxes on the 'clean' commodities are unaffected by the externalities, confirming previous results. On the other hand, the tax rates on the externality-generating commodities in general depend on the marginal social damage induced by the consumption of all commodities. Thus, I have shown that the optimal Pigouvian tax rates should not be considered in isolation separately for each externality, as the interdependence between them is important. This extension of the optimal tax design problem in the presence of externalities adds to further realism of the model, as in reality there are many goods whose consumption causes an externaliy. To get an idea about the magnitude of the elaborated effect more empirical evidence on the influence of externality-levels (e.g. of an increase in pollution) on the demand for other externality-generating commodities is required.

Appendix

First order conditions of the government's maximization problem

The first-order condition of the government's maximization problem with respect to the optimal income bundles (x_i, z_i), $i = L, H$ are given by

$$f_L \frac{\partial v_L}{\partial x_L} = \lambda \pi_L - \lambda \pi_L \sum_{j=2}^{n} \tau_j \frac{\partial c_{jL}}{\partial x_L} - \lambda \pi_L \sum_{k=1}^{m} t_k \frac{\partial d_{kL}}{\partial x_L} + \mu \frac{\partial v_H[L]}{\partial x_L} + \pi_L \sum_{k=1}^{m} \gamma_k \frac{\partial d_{kL}}{\partial x_L}, \tag{A1}$$

$$f_L \frac{\partial v_L}{\partial z_L} = -\lambda \pi_L - \lambda \pi_L \sum_{j=2}^{n} \tau_j \frac{\partial c_{jL}}{\partial z_L} - \lambda \pi_L \sum_{k=1}^{m} t_k \frac{\partial d_{kL}}{\partial z_L} + \mu \frac{\partial v_H[L]}{\partial z_L} + \pi_L \sum_{k=1}^{m} \gamma_k \frac{\partial d_{kL}}{\partial z_L}, \tag{A2}$$

$$f_H \frac{\partial v_H}{\partial x_H} = \lambda \pi_H - \lambda \pi_H \sum_{j=2}^{n} \tau_j \frac{\partial c_{jH}}{\partial x_H} - \lambda \pi_H \sum_{k=1}^{m} t_k \frac{\partial d_{kH}}{\partial x_H} - \mu \frac{\partial v_H}{\partial x_H} + \pi_H \sum_{k=1}^{m} \gamma_k \frac{\partial d_{kH}}{\partial x_H}, \tag{A3}$$

$$f_H \frac{\partial v_H}{\partial z_H} = -\lambda \pi_H - \lambda \pi_H \sum_{j=2}^{n} \tau_j \frac{\partial c_{jH}}{\partial z_H} - \lambda \pi_H \sum_{k=1}^{m} t_k \frac{\partial d_{kH}}{\partial z_H} - \mu \frac{\partial v_H}{\partial z_H} + \pi_H \sum_{k=1}^{m} \gamma_k \frac{\partial d_{kH}}{\partial z_H}, \tag{A4}$$

The first-order condition with respect to the commodity tax rates τ_j, $j = 2, ..., n$, and t_k, $k = 1, ..., m$, read

$$\begin{aligned} &\sum_i f_i \frac{\partial v_i}{\partial \tau_j} + \lambda \sum_i \pi_i c_{ji} + \lambda \sum_{s=2}^{n} \sum_i \tau_s \pi_i \frac{\partial c_{si}}{\partial \tau_j} + \lambda \sum_{k=1}^{m} \sum_i t_k \pi_i \frac{\partial d_{ki}}{\partial \tau_j} + \mu \frac{\partial v_H}{\partial \tau_j} - \mu \frac{\partial v_H[L]}{\partial \tau_j} \\ &- \sum_{k=1}^{m} \sum_i \gamma_k \pi_i \frac{\partial d_{ki}}{\partial \tau_j} = 0, \end{aligned} \tag{A5}$$

$$\sum_i f_i \frac{\partial v_i}{\partial t_k} + \lambda \sum_{j=2}^{n} \sum_i \tau_j \pi_i \frac{\partial c_{ji}}{\partial t_k} + \lambda \sum_i \pi_i d_{ki} + \lambda \sum_{s=1}^{m} \sum_i t_s \pi_i \frac{\partial d_{si}}{\partial t_k} + \mu \frac{\partial v_H}{\partial t_k} - \mu \frac{\partial v_H[L]}{\partial t_k}$$
$$- \sum_{s=1}^{m} \sum_i \gamma_s \pi_i \frac{\partial d_{si}}{\partial t_k} = 0. \tag{A6}$$

Finally, the first-order condition with respect to E_k, $k = 1, ..., m$ is given by

$$\sum_i f_i \frac{\partial v_i}{\partial E_k} + \lambda \sum_{j=2}^{n} \sum_i \tau_j \pi_i \frac{\partial c_{ji}}{\partial E_k} + \lambda \sum_{s=1}^{m} \sum_i t_s \pi_i \frac{\partial d_{si}}{\partial E_k} + \mu \frac{\partial v_H}{\partial E_k} - \mu \frac{\partial v_H[L]}{\partial E_k}$$
$$+ \gamma_k - \sum_{s=1}^{m} \sum_i \gamma_s \pi_i \frac{\partial d_{si}}{\partial E_K} = 0. \tag{A7}$$

Proof of Lemma 1

Take A6 and add and substract $\mu \frac{\partial v_H[L]}{\partial x_L} \frac{\partial v_L / \partial E_k}{\partial v_L / \partial x_L}$. A7 can then be transformed to

$$(f_L \frac{\partial v_L}{\partial x_L} - \mu \frac{\partial v_H[L]}{\partial x_L}) \frac{\partial v_L/\partial E_k}{\partial v_L/\partial x_L} + (f_H \frac{\partial v_H}{\partial x_H} + \mu \frac{\partial v_H}{\partial x_H}) \frac{\partial v_H/\partial E_k}{\partial v_H/\partial x_H}$$
$$- \mu \frac{\partial v_H[L]}{\partial x_L} (\frac{\partial v_H[L]/\partial E_k}{\partial v_H[L]/\partial x_L} - \frac{\partial v_L/\partial E_k}{\partial v_L/\partial x_L}) + \lambda \sum_{j=2}^{n} \sum_i \tau_j \pi_i \frac{\partial c_{ji}}{\partial E_k}$$
$$+ \lambda \sum_{s=1}^{m} \sum_i t_s \pi_i \frac{\partial d_{si}}{\partial E_k} + \gamma_k - \sum_{s=1}^{m} \sum_i \gamma_s \pi_i \frac{\partial d_{si}}{\partial E_k} = 0, \tag{A8}$$

$k = 1, .., m$. Make use of the definition for MWP_{ki} (equation (9)) and substitute for $(f_L \frac{\partial v_L}{\partial x_L} - \mu \frac{\partial v_H[L]}{\partial x_L})$and $(f_H \frac{\partial v_H}{\partial x_H} + \mu \frac{\partial v_H}{\partial x_H})$ from A1 and A3. Further, use the Slutsky decompositions $\frac{\partial c_{ji}}{\partial E_k} = \frac{\partial c_{ji}^{com}}{\partial E_k} - MWP_{ki} \frac{\partial c_{ji}}{\partial x_i}$ and $\frac{\partial d_{si}}{\partial E_k} = \frac{\partial d_{si}^{com}}{\partial E_k} - MWP_{ki} \frac{\partial d_{si}}{\partial x_i}$ for $j = 2, .., n$ and $s = 1, .., m$. Then A8 can be transformed to equation (10) from the text.

Derivation of the optimal commodity tax rates

Take A5 and plug in for $\frac{\partial v_i}{\partial \tau_j} = -c_{ji} \frac{\partial v_i}{\partial x_i}$, $\frac{\partial v_H[L]}{\partial \tau_j} = -c_{jH}[L] \frac{\partial v_H[L]}{\partial x_L}$ and for the Slutsky-equations $\frac{\partial c_{si}}{\partial \tau_j} = \frac{\partial c_{si}^{com}}{\partial p_j} - c_{ji} \frac{\partial c_{si}}{\partial x_i}$ and $\frac{\partial d_{ki}}{\partial \tau_j} = \frac{\partial d_{ki}^{com}}{\partial p_j} - c_{ji} \frac{\partial d_{ki}}{\partial x_i}$ for $s = 2, .., n$ and $k = 1, .., m$. Then A5 can be written as

$$- c_{jL} f_L \frac{\partial v_L}{\partial x_L} - c_{jH} f_H \frac{\partial v_H}{\partial x_H} + \lambda \sum_i \pi_i c_{ji} + \lambda \sum_{s=2}^{n} \sum_i \tau_s \pi_i (\frac{\partial c_{si}^{com}}{\partial p_j} - c_{ji} \frac{\partial c_{si}}{\partial x_i})$$
$$+ \lambda \sum_{k=1}^{m} \sum_i t_k \pi_i (\frac{\partial d_{ki}^{com}}{\partial p_j} - c_{ji} \frac{\partial d_{ki}}{\partial x_i}) - \mu (\frac{\partial v_H}{\partial x_H} c_{jH} - \frac{\partial v_H[L]}{\partial x_L} c_{jH}[L])$$
$$- \sum_{k=1}^{m} \sum_i \gamma_k \pi_i (\frac{\partial d_{ki}^{com}}{\partial p_j} - c_{ji} \frac{\partial d_{ki}}{\partial x_i}) = 0. \tag{A9}$$

In a next step substitute for $f_L \frac{\partial v_L}{\partial x_L}$ and $f_H \frac{\partial v_H}{\partial x_H}$ from A1 and A3. Observe that due to the assumption of weakly separable preferences we have $c_{jH}[L] = c_{jL}$ and $d_{kH}[L] = d_{kL}$. Then A9 reduces to

$$\sum_{s=2}^{n} \sum_i \pi_i \tau_s \frac{\partial c_{si}^{com}}{\partial p_j} + \sum_{k=1}^{m} \sum_i \pi_i t_k \frac{\partial d_{ki}^{com}}{\partial p_j} = \sum_{k=1}^{m} \sum_i \pi_i \frac{\gamma_k}{\lambda} \frac{\partial d_{ki}^{com}}{\partial p_j} \tag{A10}$$

$j = 2, .., n$. Applying the same steps to the first-order conditions for t_k given by A6 one obtains

$$\sum_{j=2}^{n} \sum_i \pi_i \tau_j \frac{\partial c_{ji}^{com}}{\partial q_k} + \sum_{s=1}^{m} \sum_i \pi_i t_s \frac{\partial d_{si}^{com}}{\partial q_k} = \sum_{s=1}^{m} \sum_i \pi_i \frac{\gamma_s}{\lambda} \frac{\partial d_{si}^{com}}{\partial q_k} \tag{A11}$$

$k = 1, .., m$. A10 and A11 can then be transformed to matrix notation as given by equation 11 from the text.

References

Atkinson, A.B. and Stiglitz, J.E. (1976), The design of tax structure: direct versus indirect taxation, *Journal of Public Economics* 6, 55-75.

Cremer, H., Gahvari, F. and Ladoux N. (1998), Externalities and optimal taxation, *Journal of Public Economics* 70, 343-365.

Edwards, J., Keen, M. and Tuomala M. (1994), Income tax, commodity tax and public good provision: a brief guide, *Finanzarchiv* 51, 472-487.

Kopczuk, L. (2003), A note on optimal taxation in the presence of externalities, *Economics Letters* 80, 81-86.

Meade, J.E. (1952), External economies and diseconomies in a competitive situation, *Economic Journal* 62, 54-67.

MICHELETTO, L. (2008), Redistribution and optimal mixed taxation in the presence of consumption externalities, *Journal of Public Economics* 92, 2262-2274.

MIRRLEES, J.A. (1971), An Exploration in the theory of optimal income taxation, *Review of Economic Studies* 38, 175-208.

PIRTTILÄ, J. and TUOMALA, M. (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.