Public Expenditures and Environmental Protection

When Is the Cost of Funds Irrelevant?

Gunnar S. Eskeland

Pigou's conjecture was that under costly taxation public expenditures should not reach the point where marginal benefits equal marginal costs. In the treatment here, public expenditures (and environmental protection) may provide public goods for consumption but also collective inputs for production. When the benefits are in production, the cost of funds is irrelevant. Why? Collective inputs benefit goods that are taxed, while for public goods the shadow price of funds reduces provision as if they were.

The World Bank
Development Research Group
Public Economics
December 2000
Summary findings

Assume that a public program—whether in the form of public expenditures or regulation of private activities—provides not only a public good to consumers but also a collective input (say, a less polluted water source for brewers, or better roads for their trucks).

In a context of optimal taxation and constant returns to scale, Eskeland shows that only the direct benefits to consumers in the form of a public good are adjusted by the shadow price of public revenue (typically downward, as Pigou conjectured) before benefits are aggregated to establish optimal provision. When public programs benefit productive sectors through cost savings, the marginal cost of provision is in optimum equal to the marginal cost savings in the benefiting sectors. The reason that programs that benefit production are not scaled down by the shadow price of public revenue is that the benefits are derived from markets that are otherwise taxable. Government can capture those cost savings at no distortionary cost by increasing the tax rates for each good, to match the cost savings provided.

In practice, do public programs to protect the environment benefit mostly consumers or mostly producers? Eskeland suggests that environmental protection has direct value for consumers and indirect value, as inputs, for producers. In the case of programs to reduce emissions of global greenhouse gases, for instance, most of the benefits appear to be in agriculture, a productive sector. Public programs in general provide a combination of public and private benefits: the share of commercial vehicles on roads is typically high in poor countries.

In related papers, “Externalities and Production Efficiency” (Policy Research Working Paper 2319) and “Environmental Protection and Optimal Taxation” (Policy Research Working Paper 2510), Eskeland shows that under optimal taxation, marginal abatement costs should be the same for polluting government, polluting producers, and polluting consumers, rich and poor.

This paper—a product of Public Economics, Development Research Group—is part of a larger effort in the group to provide guiding principles in public finance and environmental protection. Copies of the paper are available free from the World Bank, 1818 H Street NW, Washington, DC 20433. Please contact Hedy Sladovich, room MC2-609, telephone 202-473-7698, fax 202-522-1154, email address hsladovich@worldbank.org. Policy Research Working Papers are also posted on the Web at www.worldbank.org/research/workingpapers. The author may be contacted at geskeland@worldbank.org. December 2000. (32 pages)
Public Expenditures and Environmental Protection
When is the Cost of Funds Irrelevant?

Gunnar S. Eskeland*
Summary

Assume that a public program - whether in the form of government provision or as environmental protection induced by the government’s emission taxes - provides not only a public good to consumers but also a collective input (say, for brewers: a less polluted water source, or better roads for their trucks). In a context of optimal taxation and constant returns to scale, we show that only the direct benefits to consumers in terms of a public good are adjusted by the shadow price of public revenue (typically by downward, as Pigou conjectured) before benefits are aggregated to establish optimal provision. This holds also for optimal environmental protection, which is our example. Put differently, while the Samuelson (1954) condition for optimal provision of public goods applies only when taxation is costless, the analogous rule for collective inputs applies even under costly revenue generation.

One example is greenhouse gas limitations, believed in large part to benefit agriculture. Another is infrastructure investments such as roads, to the extent that they provide services to firms, rather than public goods for households. Optimal provision in these types of programs requires that marginal costs are equal to the cost savings for productive sectors - in a way independent of the shadow price of public revenue. The intuition is that output markets enable the government to capture these cost savings through taxation, so it is not worth it to have inefficient programs for revenue reasons.

This result is independent of whether the good in question is provided by government (widening a road, say) or it is damaged/generated by firms or households, rich or poor (say, emitting pollution/engaging in pollution control). Also, it is independent of whether the final users of the outputs from the benefiting industries are rich or poor, government or firms.

Apart from being of policy relevance in itself, our results assist in the interpretation of established principles. The well-known adjustment of benefits in the form of public goods plays the role of taxing a good that otherwise would not be taxed (King, 1986). In contrast, benefits in the form of cost savings in production are derived from demand in markets that can be taxed directly, and thus need no such adjustment. Christiansen (1981) makes a finding much like ours - but with a very different twist on standard assumptions.
Interestingly, the intuition that commodity taxes make the benefits taxable supports both findings.

Results along lines similar to the ones shown here are given in Bovenberg and van der Ploeg (1994) and in Williams (2000). Amongst the contributions of the present paper is to highlight how these results are closely related to Diamond and Mirrlees' (1971) result on production efficiency. This link to central questions in the optimal tax literature is seen more easily in our treatment since our analysis is more general, and in particular because we treat in a parallel fashion programs pursued through government expenditures and programs pursued by soliciting efforts from the private sector (as with pollution abatement). For applied analysis as well as for principles, an advantage of our approach is to highlight when the tax system is better suited for redistribution than is the modification of public programs.
I. Introduction

In the public finance literature, environmental quality is typically introduced as a pure public good in the sense of Samuelson (1954), directly represented in consumer preferences.\(^1\) This public good is provided at a certain level by nature and in addition influenced by human activities, most typically depleted as a negative external effect of consumption and/or production activities. Important contributions that illustrate this approach are Sandmo (1975) and - more recently - Cremer et al. (1998).

In our treatment, there are a number of parallels between provision via public expenditures - e.g. to procure public goods - and provision via other government powers, as when regulation and taxation is used by the government to protect the environment.\(^2\) We shall use the term public program for the general case when provision is either by government directly - as with a public road - or induced by government policies - as when government acts with emission standards or taxation to stimulate emission reductions from the private sector.

Our more important generalization, however, is in terms of how the provided items are useful. We let the public program generate cost reductions to producers in addition to the benefits enjoyed directly by consumers in the form of a pure public good. We conduct this analysis while describing the public program as providing an environmental good - i.e. addressing a negative externality by inducing emission reductions. However, our central result applies equivalently when a good that is collectively available to households and/or producers is provided directly by government.

It is not new to suggest that publicness - in the sense of nonrivalry amongst beneficiaries - can apply to a factor of production as well as to goods that play a role in

\(^1\) Samuelson's term was collective consumption goods "which all enjoy in common in the sense that each individual's consumption of such a good leads to no subtraction from any other individual's consumption of that good". Later, the terms "pure public good" has come to be accepted for goods that are non-rivalrous in consumption, and terms like "non-rivalry" or "non-exhaustability" are often used as criteria for pure public goods (see, for instance, Laffont, 1988). Interestingly, in Samuelson's article, the word 'pure' was used to qualify his contribution to pure theory, not to describe the type of goods in question.

\(^2\) One of these parallels are shown in Eskeland (2000), in which the environment is a pure public good. Pollution abatement by the firms (firms, households) and by government is shown in optimum to yield the same marginal emission reductions. For firms and government, this is also an implication of Diamond and Mirrlees' (1971) result on aggregate production efficiency. Government abatement is equivalent to provision of a public good via public expenditures. Thus, externalities and government provided public goods differ in the means of intervention but not in the more basic optimality conditions, such as the wedge - if any - between marginal rates of transformation and marginal benefits.
consumer preferences. First, as A. C. Pigou introduced what later came to be known as “external effects”, or externalities, his words were: “a person A, in the course of rendering some service, for which payment is made, to a second person, B, incidentally also renders services” .. “to other persons (not producers of like services), of such sort that payment cannot be exacted from the benefited parties...”. Thus, with his parenthesis, Pigou included producers in his notion of beneficiaries (Pigou, 1932).3 Further, in his concrete examples, the affected third parties were producers whose costs are affected as well as households whose wellbeing are affected: The lighthouse generates benefits to ships, and smoke from a chimney results in nuisance and higher costs for neighboring households as well as firms (Pigou 1932, pages 183-184; 1947, pages 94-95). Second, when Meade (1952) gave name to the influential concept of “atmospheric externalities”, he used effects between producers as examples. Third, among others, Sandmo (1972) and Oakland (1987) have treated “collective factors” of production as a case parallel to pure public goods. In a context without costly revenue generation, the result is an intuitively appealing rule for optimal provision analogous to the Samuelson condition (1954) for optimal provision of a pure public good, with vertical summation over beneficiaries of marginal benefits.

In a context of costly revenue generation, in contrast, treatments of external effects and of optimal provision have emphasized pure public goods directly represented in consumer preferences, as opposed to collective factors of production. Thus our contribution lies in setting collective factors of production in a context of distortionary revenue generation.

We should conclude this introduction by visiting the literature that feeds our priors. First, we should highlight what may be called Pigou's conjecture: That public expenditures should be lower in the context of costly revenue generation than they would be under lump sum taxation (Pigou, 1947, page 34). Atkinson and Stern (1974), following Stiglitz and Dasgupta (1971), qualified Pigou's conjecture by classifying the exceptions that apply for the case with a pure public good. Second, a more recent

---

3 We follow the tradition in the literature not to press sign generalizations into the text: the word tax is used even if it may be negative (a subsidy). Similarly, we may talk about the choices of polluters as benefiting the polluted (polluters are potential providers of pollution reductions, i.e. of the public good). Thereby, we
literature focuses on the interaction between an environmental pure public good and costly revenue generation (See, for instance, Bovenberg and de Mooij, 1994). What differentiates that problem is additional restrictive assumptions about preferences and technology; the fact that provision is induced by taxation turns out not to be important in terms of basic principles, as long as the government still needs to resort to costly revenue generation.

To this author's knowledge, no direct analysis has been made of Pigou's conjecture in the case when the public expenditures benefit production with a "collective factor of production" (Sandmo, 1972). Diamond and Mirrlees' result on production efficiency (obtained under general assumptions equivalent to ours), and also the treatment of Dasgupta and Stiglitz (1972) would seem to indicate that "effects" between producers, and between producers and government entities should reflect fully the benefits to the recipients - with no distorting wedges - whether the goods/inputs "delivered" are private or non-rivalrous to the benefiting entities. But what if a public program benefits producing sectors as well as consumers, as does for instance the expansion of a road? If in addition the good in question can be generated by consumers and producers, as with emission reductions, the answer does not seem to be indicated by the result on aggregate production efficiency.

In benefit cost analysis, two factors may have contributed to the fact that benefits to firms have not been highlighted. First, there is a tradition in that literature to model net effects, the analyst taking program benefits through producing sectors to evaluate them as they give households expanded consumption opportunities. Our treatment supplements this approach by highlighting what it implies in terms of valuation of the more direct cost savings accruing in production, and also provides a critique in terms of policy principles. Second, cost benefit analysis is often not conducted under the assumptions that lead to the production efficiency result; it may not even be assumed that the tax structure is optimal.

facilitate the analogy with public goods provision. Here, we skipped Pigou's sign generalization: "services or disservices".
II. The Model

We shall modify the traditional model of an environmental good by allowing not only consumers but also productive sectors to benefit from pollution reductions (equation 3, below). Since other aspects of the model are the same as in Eskeland, 2000, we shall be brief in our description of the model.

Preferences. A consumer g's utility $u^e$ depends on her consumption $x^e_j$ of $n + 1$ private goods, $j = 0,1,\ldots,n$, as well as on the aggregate level of emissions, $e$, in her environment. $e$ is a public good (we might have called it a "public bad", since it has a negative effect on utility, see below), experienced at the same level by all consumers. For convenience of notation, we assume that the $h$ consumers are identical, and we may thus suppress personal superscripts (we briefly treat the case with heterogenous consumers in section IV). Thus,

$$u = u(x_0, x_1, \ldots, x_n, e) = u^e(x^e_0, x^e_1, \ldots, x^e_n, e).$$

We assume that the utility function is continuous, twice differentiable, quasiconcave, and that $u_e \leq 0$.\(^4\) Government consumption is assumed constant and is therefore suppressed in (1).

Emissions, aggregated across agents, is linked linearly by emission factors $f_j$, $j=1,\ldots,n$ to total output in each sector (i.e. to consumption or production - or any combination - of each good). We shall introduce endogenous emission factors later (in equation 32), and initially simplify by assuming exogenously given emission factors.

$$e = \sum_j f_j \cdot (hx_j + x^e_j),$$

where $x^e_j$ is the exogenously given government consumption (or use) of good $j$.\(^5\) To simplify, we assume that the numeraire good is not polluting: $f_0 = 0$, and we describe all other goods as possibly polluting, $f_j \geq 0, j = 1,\ldots,n$.

Producers are harmed if there is pollution in the environment, but apart from this, production possibilities are characterized by constant marginal rates of transformation.

---

\(^4\) Whenever possible without risking confusion, we shall use subscripts to denote partial derivatives.
\(^5\) If not otherwise indicated, summation indicated for instance by $\Sigma_j$ is over goods $j=1,\ldots,n$. 
\[(3) \quad h_{x_0} + x_0^p + \sum_{j=1}^{n} c_j(e) \cdot (h_{x_j} + x_j^p) = K,\]

where \(K\) is a constant. (3) can be thought of as a general conversion technology possessed by many producers, and we shall assume perfectly competitive behavior.

It is in (3) that the somewhat novel element of this optimal tax analysis is introduced. \(c_j(e)\), continuous, twice differentiable and convex, describes how the unit cost (and marginal cost – they are equal) in each sector (i.e. for each good) depends on the aggregate level of pollution. Sectors may vary in their sensitivity to pollution, but we assume for simplicity that \(c_j = \partial c_j / \partial e \geq 0\), all \(j = 1, \ldots, n.\) We describe pollution as not affecting sector zero, but this is no restrictive assumption, since the role of \(c\) simply is to describe the \(n\) marginal rates of transformation between the \(n+1\) goods, and how these change as the level of pollution changes.

The consumer has a nontaxable lump sum income \(I\) and faces consumer prices \(q = q_0, q_1, \ldots, q_n\) equal to producer prices plus linear commodity taxes

\[(4) \quad q_0 = 1, \quad q_i = c_i(e) + t_i, \quad i = 1, \ldots, n.\]

We assume that consumers and producers take as given the levels of pollution and government revenue. The first-order conditions for the consumer's individual optimum are her budget constraint \(\sum_{j=0}^{n} q_j x_j = I\) and, for \(j = 1, \ldots, n,\)

\[(5) \quad \frac{u_j}{u_0} = q_j.\]

Let

\[(6) \quad x_j = x_j(q, I, e)\]

---

\(6\) To clarify, we retain the classical description of what causes emissions: each polluter derives benefits from her own individual emissions, exploiting the environment's services as a waste recipient. The disservices thus provided by polluters are analogous to the services from those who provide a traditional public good in the sense that both are rivalrous. For polluters, when one emits more, another must emit less, if environmental quality is not to fall. For providers, when one provides less, another must provide more, if provision of the public good is not to fall. The novel aspect introduced here is that producers are "harmed" by the aggregate level of pollution in the environment, a producer analogy to the consumer side of a traditional public good, which is nonrivalrous.
denote the Marshallian demand function consistent with the consumer's first order conditions for optimum.

A benevolent planner uses linear commodity tax rates \( t = t_1, \ldots, t_n \) to maximize the indirect utility function \( v(q, I, e) \) corresponding to (1), (4) and (6) subject to the government's budget constraint \( h \sum_j t_j x_j = x^p_0 + \sum_j c_j(e)x^p_j \). The Lagrangian of the planner's maximization problem is

\[
L = hv(q, I, e) + \mu \sum_j [ht_j x_j - (x^p_0 + c_j(e)x^p_j)].
\]

III. Optimal provision

We may partially differentiate (7) with respect to the \( n \) tax rates to find first order conditions for welfare optimum. For all \( i = 1, \ldots, n \), we have:

\[
\frac{\partial L}{\partial t_i} = -\beta \left[ \sum_j x_j \frac{dq_j}{dt_i} + \alpha \frac{de}{dt_i} \right] + \mu \left[ h \left( x_i + \sum_i t_i \frac{dx_i}{dt_i} \right) - \sum_j x^p_j c_j \frac{de}{dt_i} \right] = 0,
\]

where \( \beta \) is the marginal value of income (in terms of the numeraire) to the consumer and \( \alpha \) is the consumer's marginal willingness to pay for pollution reductions, \( \alpha = -\frac{\partial v}{\partial e} \frac{\partial v}{\partial I} \).

Using (4), we have

\[
\frac{dq_i}{dt_i} = 1 + c_{i e} \frac{de}{dt_i}, \text{ and } \frac{dq_j}{dt_i} = c_{j e} \frac{de}{dt_i}, \text{ for } j \neq i.
\]

From (2) and the assumption that government consumption is exogenously determined,

\[
\frac{de}{dt_i} = h \sum_j f_j \frac{dx_j}{dt_i}.
\]

Substituting (9) and (10) into (8), we have, for all \( i = 1, \ldots, n \)

\[
\frac{\partial L}{\partial t_i} = -\beta \left[ x_i + (\alpha + \sum_j x_j c_{j e}) h \sum_j f_j \frac{dx_j}{dt_i} \right] + \mu \left[ x_i + \sum_i t_i \frac{dx_i}{dt_i} - \sum_j x^p_j c_j \sum_j f_j \frac{dx_i}{dt_i} \right] = 0.
\]

Reordering to have the tax rates and their coefficients on the left hand side, we have

\[\text{In the case of fixed emission coefficients, assumed here (equation 2), the role of emission taxes can be assumed by commodity taxes (see Sandmo, 1975, and Eskeland, 2000). Emission taxes play a role, however, in the case with endogenous emission coefficients, later in section III.}\]
It will prove useful to have introduced matrix notation. Let $G$ be the following matrix of total derivatives:

$$
G = 
\begin{bmatrix}
\frac{dx_1}{dt} & \cdots & \frac{dx_n}{dt} \\
\frac{dx_1}{dt} & \cdots & \frac{dx_n}{dt} \\
\vdots & \ddots & \vdots \\
\frac{dx_1}{dt} & \cdots & \frac{dx_n}{dt}
\end{bmatrix}
$$

We can now write (12) as

$$
G \cdot t = \left( \frac{\beta}{\mu} - 1 \right) x + G \cdot f \left[ h \frac{\beta}{\mu} \left( \alpha + \sum_j x_j c_{je} \right) + \sum_j x_j^p c_{je} \right],
$$

where $t$ and $f$ are $n$-vectors of tax rates and emission factors, respectively. Assuming that $G$ is non-singular (to be explored below, footnote 9), we may pre-multiply both sides by the inverse of $G$ to have the tax rates in a more explicit form:

$$
G^{-1} G t = t = G^{-1} \left( \frac{\beta}{\mu} - 1 \right) x + f \left[ h \frac{\beta}{\mu} \left( \alpha + \sum_j x_j c_{je} \right) + \sum_j x_j^p c_{je} \right].
$$

(14) characterizes the optimal tax structure.

To obtain further insight, we need to analyze the matrix $G$, and to do this, we shall assume that the environmental good is separable in preferences (1) from the $n+1$ market goods, so that $x_{je} = 0$, all $j = 1, \ldots, n$ (see equation 6). Let us examine the differentials $dx_j / dt$. Using separability, (6) and (9),

$$
\frac{dx_k}{dt} = \sum_i \frac{dx_k}{dq_j} \frac{dq_j}{dt} = x_{ki} + \sum_j x_{ji} c_{je} \frac{de}{dt}.
$$

---

8 This separability assumption, typically invoked in models with public goods, is discussed further in the annex to this essay, and Eskeland (2000b). Under nonseparability, optimal provision of the public good, minus $e$, is expanded (the emission tax is raised) to the extent that provision contributes to demand for taxed goods ($\sum_j x_{je} < 0$) and vice versa. When the public good is provided directly by government (or when an emission tax is used to induce environmental protection), the expressions for optimal commodity taxes are not changed. The plausibility of the separability assumption clearly depends on the externality. Consider two negative externalities from driving; pollution and congestion. Separability appears more plausible for the former.
For illustration, let us briefly visit the example with three private goods, or two relative prices; \( x_j = x_j(q_1, q_2) \) and two tax rates. Only the goods that are taxed and pollute appear in \( G \), and we have two equations in two unknowns, \( dx_1/dt \) and \( dx_2/dt \) and two additional equations for \( dx_1/dt_1, dx_2/dt_1 \). The first of these four equations is, as an example:

\[
\frac{dx_1}{dt_1} = x_{11} + x_{11}c_{ie} \frac{de}{dt_1} + x_{12}c_{2e} \frac{de}{dt_1},
\]

or on a more standardized form:

\[
\frac{dx_1}{dt_1} \left( 1 - h_f \sum_{j=1}^{2} x_{1j}c_{je} \right) + \frac{dx_2}{dt_1} \left( - h_f \sum_{j=1}^{2} x_{1j}c_{je} \right) = x_{11}. \]

The system has one solution if and only if the coefficient matrix \( (F, \text{below}) \) is nonsingular. Making this assumption, let us display the solution for one example of these four total differentials:

\[
\frac{dx_1}{dt_1} = - h_f c_{2e} (x_{11}x_{22} - x_{12}x_{21}) \frac{1}{1 - h_f \sum_{j=1}^{2} f_i \sum_{j=1}^{2} x_{ij}c_{je}},
\]

where the denominator, the determinant of the coefficient matrix, is written in a form applicable for a general number of goods.

For the case of a general number of goods, let \( A \) be the transpose of the Jacobian matrix of the Marshallian demand functions for goods \( 1, \ldots, n \):

\[
(16) \quad A = \begin{bmatrix}
    x_{11} & \cdots & x_{1n} \\
    \vdots & \ddots & \vdots \\
    x_{n1} & \cdots & x_{nn}
\end{bmatrix}.
\]

Working with the individual elements \( dx_j/dt_i \) in the general case, we find that

\[
(17) \quad GF = A,
\]

where

\[
F = \begin{bmatrix}
    1 - h_f \sum_{j=1}^{n} x_{1j}c_{je} & - h_f \sum_{j=1}^{n} x_{2j}c_{je} & \cdots & - h_f \sum_{j=1}^{n} x_{nj}c_{je} \\
    - h_f \sum_{j=1}^{n} x_{1j}c_{je} & 1 - h_f \sum_{j=1}^{n} x_{2j}c_{je} & \cdots & - h_f \sum_{j=1}^{n} x_{nj}c_{je} \\
    \vdots & \vdots & \ddots & \vdots \\
    - h_f \sum_{j=1}^{n} x_{1j}c_{je} & - h_f \sum_{j=1}^{n} x_{2j}c_{je} & \cdots & 1 - h_f \sum_{j=1}^{n} x_{nj}c_{je}
\end{bmatrix}.
\]

Assuming that \( A \) and \( F \) are nonsingular, we have

\[
GFF^{-1} = AF^{-1} \Rightarrow G = AF^{-1} \Rightarrow G^{-1} = (AF^{-1})^{-1} \Rightarrow
\]

\[
G^{-1} = FA^{-1}.
\]

The inverse of \( A \) involves its determinant and the transpose of the matrix of its cofactors:
where $A_y$ is the cofactor of $x_y$ in $A$. 9 Thus

$$A^{-1} = \frac{1}{|A|} \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix},$$

where $A_y$ is the cofactor of $x_y$ in $A$. 9 Thus

$$G^{-1} = \frac{1}{|A|} \begin{bmatrix} 1 - h f_1 \sum_j x_{ij} c_{je} & \cdots & - h f_1 \sum_j x_{nj} c_{je} \\ \vdots & \ddots & \vdots \\ - h f_n \sum_j x_{ij} c_{je} & \cdots & 1 - h f_n \sum_j x_{nj} c_{je} \end{bmatrix} \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix}.$$

We may write $F$ as the sum of the identity matrix and one with elements $- h f_i \sum_j x_{ij} c_{je}$ in row $i$, column $k$. We then obtain:

$$G^{-1} = \frac{1}{|A|} \left\{ \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix} + \begin{bmatrix} - f_i \sum_j c_{je} \sum_i A_{il} x_{ij} & \cdots & - f_i \sum_j c_{je} \sum_i A_{il} x_{ij} \\ \vdots & \ddots & \vdots \\ - f_n \sum_j c_{je} \sum_i A_{il} x_{ij} & \cdots & - f_n \sum_j c_{je} \sum_i A_{il} x_{ij} \end{bmatrix} \right\}.$$

The elements of the last matrix are of the form $f_m \sum_j c_{je} \sum_i A_{il} x_{ij}$. By central results for expansion of determinants by cofactors, $\sum_i A_{il} x_{ij} = |A|$ for $j = k$, and $\sum_i A_{il} x_{ij} = 0$ for $j \neq k$, so

---

9 From $G = AF$ we can elaborate on the condition that $G$ is nonsingular. First, $G$ here plays a role analogous to the one played by $A$ in the more traditional problem of optimal commodity taxes with a public good separable from private goods in preferences. $A$ is negative semi-definite by the second order conditions for consumer optimum, and the additional assumption that $A$ is of full rank ($|A| \neq 0$) is standard and ensures that there are $n$ independent instruments in the $n+1$ commodity problem. The interesting question is thus about the transformation performed by $F$ of $A$ into $G$. The determinant of $F$ is $1 - h \sum_j c_{je} \sum_i A_{il} x_{ij}$, and if externalities are "big", $|F|$ can thus be zero. If $|F|$ is zero, $G$ is of less than full rank despite $A$ being of full rank, and in that case the optimal tax problem has at most $n-1$ independent instruments. An illustration of this problem is as follows: Say only good one is polluting and only good one is affected (as with congestion). Then, since $(dx_1/dt_1) (1 - h f_1 x_{11} c_{1e}) = x_{11}, |F| = 1 - h f_1 x_{11} c_{1e} = 0$ would mean that congestion influences demand for good one so strongly that it cancels the direct effect of a tax on good one, and changes in the tax on good one does not influence demand for good one. While we do not pursue this question further, to avoid anomalies, stronger conditions than $|F| \neq 0$ are required. In this case, $1 > h f_1 x_{11} c_{1e}$ would ensure that $dx_1/dt_1$ and $x_{11}$ take the same sign (We here benefit from insights in an inquiry started by Buchanan and Kafoglis, 1963, continued by Diamond and Mirrlees, 1973, Diamond, 1973, and Sandmo, 1980).
\[ G^{-1} = \frac{1}{|A|} \begin{bmatrix} A_{11} & \ldots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \ldots & A_{nn} \end{bmatrix} - h \begin{bmatrix} f_1 c_{1e} |A| & \cdots & f_1 c_{ne} |A| \\ \vdots & \ddots & \vdots \\ f_n c_{1e} |A| & \cdots & f_n c_{ne} |A| \end{bmatrix} \]

\[ = A^{-1} - h \begin{bmatrix} f_1 c_{1e} & \cdots & f_1 c_{ne} \\ \vdots & \cdots & \vdots \\ f_n c_{1e} & \cdots & f_n c_{ne} \end{bmatrix} = A^{-1} - h f_c e', \]

where \( c_e \) is the row vector \( c_{1e}, \ldots, c_{ne} \). Using (21) in (14), we have

\[ t = A^{-1} \left( \frac{\beta}{\mu} - 1 \right) x - h f_c e' \left( \frac{\beta}{\mu} - 1 \right) x + f \left[ \frac{h \beta}{\mu} (\alpha + \sum_j x_j c_{je}) + \sum_j x_j^p c_{je} \right], \]

or for all \( k = 1, \ldots, n \):

\[ t_k = \left( \frac{\beta}{\mu} - 1 \right) \frac{\sum_j A_{kj} x_j}{|A|} - \left( \frac{\beta}{\mu} - 1 \right) h f_k \sum_j x_j c_{je} + f_k \left[ \frac{h \beta}{\mu} (\alpha + \sum_j x_j c_{je}) + \sum_j x_j^p c_{je} \right]. \]

Elements in the second term cancel against elements in the third term, and the expression simplifies to

\[ t_k = \left( \frac{\beta}{\mu} - 1 \right) \frac{\sum_j A_{kj} x_j}{|A|} + f_k \left( \frac{\beta}{\mu} h \alpha + \sum_j (hx_j + x_j^p) c_{je} \right) \text{ all } k = 1, \ldots, n. \]

This corresponds to Sandmo's (1975) formula (and the generalization in Eskeland, 2000) in the case when \( c_{je} = 0, \text{ all } j = 1, \ldots, n \). The optimal tax structure combines the formulas from the traditional problem of optimal taxation without external effects with an externality motivated term \( f_k \left( \frac{\beta}{\mu} h \alpha + \sum_j (hx_j + x_j^p) c_{je} \right) \), where the latter is included for each polluting good according to the emission factor \( f_k \). Thus, the emission motivated term gives the same inducement – in terms of dollars per unit of emissions – across all polluting goods.

An alternative implementation

As with the optimal allocation in Eskeland (2000), the one supported by \( n \) commodity taxes in (24) can be implemented in a more intuitive manner by an alternative
set of $n+1$ instruments: an emission tax $\tau$ levied uniformly on emissions from all activities, in combination with $n$ commodity taxes:

\begin{equation}
\tau = \frac{h\beta\alpha}{\mu} + \Sigma_j (hx_j + x_j^p)c_{je}, \quad \text{and}
\end{equation}

\begin{equation}
t_k = \left( \frac{\beta}{\mu} - 1 \right) \frac{\Sigma A_{jk}x_j}{|A|}, \quad \text{all } k = 1, \ldots, n.
\end{equation}

(26) is equivalent to the solution given by Samuelson (1951) for the problem of optimal linear commodity taxes in the problem without an environmental externality. Special cases of (26) are known as “inverse elasticity” rules and Corlett and Hague’s (1953) rule, which emphasizes complementarity with the untaxed good.

To see that the system (25), (26) is consistent with optimum, proceed as follows:

Restate the Lagrangian (7) to have proceeds from the emission tax $\tau$ included in the government budget constraint,

\begin{equation}
L = hv(q, I, e) + \mu \Sigma_j [h(t_j + \sigma_j)x_j - (x_j^p + c_j(e))x_j^p].
\end{equation}

Nothing has changed with the underlying economic problem, of course: We still assume it is of rank $n$, so there is redundancy in instruments and one can be chosen arbitrarily. We shall proceed by taking a choice for the emission tax as given by (25), to check whether this in combination with (26) is consistent with the first order conditions for maximum of (27). Partially differentiating (27) with respect to the $n$ commodity tax rates, the first order conditions for maximum of (27), corresponding to (12) are:

\begin{equation}
\sum_{k} t_k \frac{dx_k}{dt_i} + \tau \sum_{k} f_{ik} \frac{dx_k}{dt_i} = \left( \frac{\beta}{\mu} - 1 \right) x_i + \Sigma_{k} f_{ik} \left[ h + \Sigma_j x_j c_{je} \right], \quad \text{all } i = 1, \ldots, n.
\end{equation}

As we substitute (25) into (28), several terms cancel, and we obtain

\begin{equation}
\sum_{k} t_k \frac{dx_k}{dt_i} = \left( \frac{\beta}{\mu} - 1 \right) x_i + \Sigma_{k} f_{ik} \left( \frac{\beta}{\mu} - 1 \right) h \Sigma_j x_j c_{je}, \quad \text{all } i = 1, \ldots, n, \text{ or}
\end{equation}

\begin{equation}
G \cdot t = \left( \frac{\beta}{\mu} - 1 \right) x + G \cdot f \left( \frac{\beta}{\mu} - 1 \right) h \Sigma_j x_j c_{je} \Rightarrow
\end{equation}

\begin{equation}
t = A^{-1} \left( \frac{\beta}{\mu} - 1 \right) x - h c x, \quad \text{and}
\end{equation}

\begin{equation}
t = A^{-1} \left( \frac{\beta}{\mu} - 1 \right) x + f \left( \frac{\beta}{\mu} - 1 \right) h \Sigma_j x_j c_{je} = A^{-1} \left( \frac{\beta}{\mu} - 1 \right) x,
\end{equation}

\begin{equation}
A^{-1} \left( \frac{\beta}{\mu} - 1 \right) x + f \left( \frac{\beta}{\mu} - 1 \right) h \Sigma_j x_j c_{je} = A^{-1} \left( \frac{\beta}{\mu} - 1 \right) x,
\end{equation}

\begin{equation}
A^{-1} \left( \frac{\beta}{\mu} - 1 \right) x + f \left( \frac{\beta}{\mu} - 1 \right) h \Sigma_j x_j c_{je} = A^{-1} \left( \frac{\beta}{\mu} - 1 \right) x,
\end{equation}

\begin{equation}
A^{-1} \left( \frac{\beta}{\mu} - 1 \right) x + f \left( \frac{\beta}{\mu} - 1 \right) h \Sigma_j x_j c_{je} = A^{-1} \left( \frac{\beta}{\mu} - 1 \right) x,
which is the same as (26), so (25),(26) implements the optimal allocation.

In the more general model with endogenous abatement (see below) the implementation with an emission tax (25), (26) represents an essential implementation mechanism, rather than an alternative one. In the present model with exogenous emission factors, it contributes by separating the instrument implementing environmental protection from the commodity tax rates, thereby emphasizing parallels to the more basic problems in the literature. For these reasons, we view the implementation with an emission tax (25), (26) as the reference point from now on.

Abatement technologies available

We now consider the more realistic and interesting case when it is possible in each activity \(i=1,\ldots,n\) to reduce emissions per unit by expending resources on abatement, \(b_i\), so \(f_i = f_i(b_i)\). This model is described in greater detail in Eskeland, 2000.\(^{10}\) An important result provided there is that abatement shall be efficient across abatement opportunities for producers, consumers and government, and across polluting goods and activities. This implies that emission reductions from consumers and producers can be stimulated by a uniform tax levied on emissions where they occur. We use this result here by describing abatement as performed by the user of a polluting good, focusing first on the consumer who faces an emission tax. This covers the general case if we interpret \(b_i\) as the sum of abatement by the producer and the consumer of a unit of good \(i\), and \(f_i\) as the sum of emissions from the maker and the user. We allow the government as a user to choose its abatement level independently (equation 32 and 46, below).

For emissions, recalling that subscripts index goods, \(i,j=1,\ldots,n\), we now have

\[
(32) \quad \epsilon = \sum_j \left( h f_j (b_j) x_j + f_j^p (b_j^p) x_j^p \right)
\]

replacing (2). We assume the marginal cost of emission reductions, \(-1/f_i\), is positive and increasing in abatement. As we let the "all-inclusive consumer prices" include abatement and the emission tax \(\tau\), (4) is replaced by

\(^{10}\) A familiar example is emissions from motor vehicles: Makers of cars and fuels can reduce the car’s emission factors by modifying the vehicle and its fuels; drivers/owners can reduce emission factors for driving through maintenance, tuning and style of driving.
(33) \[ q_i = c_i(e) + t_i + \tau f_j(b_j) + b_i, \quad i=1,\ldots,n. \]

The polluter (now represented by the consumer) is assumed to abate to minimize the "all-inclusive consumer price" (33). This implies that abatement is a function of the emission tax only, and marginal costs of emission reductions equal the emission tax:

(34) \[ \tau = -1/f_{ib}(b_i), \quad i=1,\ldots,n. \]

The Lagrangian (27) still applies (modified with government abatement, assumed given), and partial differentiation with respect to the \( n \) commodity taxes yields \( n \) expressions identical to (28). Moreover, upon examination, (9) and (10) still apply, and the proof from the previous section that the \( n \) first-order conditions for the commodity taxes are satisfied with tax rates (25), (26) still applies.

With abatement, however, the underlying problem has changed. There is no redundancy in a set of \( n+1 \) instruments, since \( \tau \) influences abatement, and no other instrument does. We shall now need:

(35) \[ \frac{dq_m}{d\tau} = f_m + c_{me} \frac{de}{d\tau}, \quad \text{all } m=1,\ldots,n, \text{ from (33), and the envelope theorem, and} \]

(36) \[ \frac{dx_i}{d\tau} = \sum_m x_{im} \frac{dq_m}{d\tau}, \quad \text{all } i=1,\ldots,n, \text{ from (6), and separability.} \]

Partial differentiation of (27) under the assumption that abatement is available yields the following first order condition for instrument number \( n+1 \):

(37) \[ -h \beta \left[ \sum_j x_j f_j + (\alpha + \sum_j x_j c_{je}) \frac{de}{d\tau} \right] + \mu \left[ h \left( \sum_j x_j f_j + \sum_i t_i \frac{dx_i}{d\tau} \right) + (\tau + \sum_j x_j^p c_{je}) \frac{de}{d\tau} \right] = 0. \]

It remains to be checked whether (25), (26) satisfies (37). (37)=>

(38) \[ h \sum_i t_i \frac{dx_i}{d\tau} = h \left( \frac{\beta}{\mu} - 1 \right) \sum_j x_j f_j + \frac{de}{d\tau} \left[ h \frac{\beta}{\mu} (\alpha + \sum_j x_j c_{je}) - \tau + \sum_j x_j^p c_{je} \right]. \]

We substitute in (25) and simplify to obtain

(39) \[ \sum_i t_i \frac{dx_i}{d\tau} = \left( \frac{\beta}{\mu} - 1 \right) \sum_j x_j \left[ f_j + c_{je} \frac{de}{d\tau} \right]. \]

Using (35) and (36), we have

(40) \[ \sum_i t_i \frac{dx_i}{d\tau} = \sum_i t_i \sum_m x_{im} \frac{dq_m}{d\tau} = \sum_i t_i \sum_m x_{im} \left( f_m + c_{me} \frac{de}{d\tau} \right), \quad \text{and thus (39) =>} \]
For both sides, let us take one element in the sum over $m=1,..,n$, looking only at the first term in the brackets:

$$
\sum_i t_i \sum_m x_{im} \left[ f_m + c_{me} \frac{de}{d\tau} \right] = \left( \frac{\beta}{\mu} - 1 \right) \sum_m x_m \left[ f_m + c_{me} \frac{de}{d\tau} \right].
$$

If we divide by $f_m$ on both sides, we have one of the first order conditions in the traditional problem of optimal commodity taxation (without external effects) - the simple problem that is solved by (26). Thus, (42) is satisfied as (26) is satisfied, and we may subtract (42) from (41) for $m=1,..,n$. What remains is

$$
\sum_i t_i x_{im} c_{me} \frac{de}{d\tau} = \left( \frac{\beta}{\mu} - 1 \right) x_m c_{me} \frac{de}{d\tau}.
$$

Here, we conduct similar steps again: For each element in the sum over $m=1,..,n$, we have

$$
\sum_i t_i x_{im} c_{me} \frac{de}{d\tau} = \left( \frac{\beta}{\mu} - 1 \right) x_m c_{me} \frac{de}{d\tau},
$$

which is just $c_{me} \frac{de}{d\tau}$ times (42) divided by $f_m$.

Thus, we have shown that a tax structure satisfying the set (25), (26) satisfies the first order conditions for optimum also in the case with endogenous abatement.

We can now state the condition for optimal environmental protection given by (25) and (34):

$$
-\frac{1}{f_{ib}} = h\alpha \frac{\beta}{\mu} + \Sigma(hx_j + x_j^p) c_{je}, i=1,..,n.
$$

On the left hand side is the marginal rate of transformation between emission reductions and the numeraire good. On the right hand side is an aggregation of marginal benefits which – for the benefits accruing as a public good to consumers - includes an adjustment factor $\beta/\mu$, the inverse of the marginal cost of funds (see below).

We can now use a finding from Eskeland (2000) to show that an alternative implementation of this allocation is one with emission standards or abatement standards in combination with commodity taxes like those of Sandmo (1975), which include presumptive emission taxes. For a particular structure of monitoring costs, plausible for
cars, for instance, emission standards or abatement standards are feasible when emission
taxes are not, so this alternative implementation can be an attractive one.

Direct provision of a public good
We only state briefly that the same condition applies for government provision of a
public good, formulated here as government abatement of emissions (\(b_i^p\) in equation 32).
This formulation is unusual in the way public provision is linked to public consumption
of \(j, x_j^p\) (assumed given), but the result is easily seen to be general:

\[
\frac{-1}{f_i^p(b_i^p)} = h\alpha \frac{\beta}{\mu} + \Sigma (hx_j + x_j^p)c_j, \text{ all goods } i=1,..,n.
\]

Thus, marginal costs of provision are equal to those for consumers and producers in (45),
and we may note that this equality does not depend on the government emission function
\(f_i^p\) being equal to those of producers or consumers (see equation 32).

In the current context, it is obvious that the marginal costs of emission reductions
will also be the same across consumers (consumers are identical and are exposed to the
same emission tax). Eskeland (2000) shows that the planner would want emission taxes
to be uniform across polluting goods even with heterogenous consumers as long as
consumers have the same access to abatement technologies (i.e. that \(f_i^p(b_i^x) = f_i^h(b_i^h)\) for
\(b_i^x = b_i^h\), so that if Peter and Paul install the same catalytic converter, then their emissions
are reduced by the same amount).

We are now ready to state our main result:

Proposition:
When optimal public provision or environmental protection is benefiting
A) consumers with a pure public good, a wedge (the ratio of the marginal utility of
income to the shadow price of public revenue to) separates the Samuelsonian sum of
marginal benefits and the marginal rate of transformation;
B) production with a nonrivalrous input, so that provision reduces marginal costs in
production, there is production efficiency in the sense that no wedge applies between
the sum of marginal benefits and the marginal rate of transformation.
Equations (45) and (46) present the optimality condition for a public program that combines these two types of benefits. The observation that benefits in terms of productivity will not be adjusted by the shadow price of public funds has also been made - in more restrictive models - by Bovenberg and van der Ploeg (1994) and Williams (2000). However, these do not place the result in the context of production efficiency or of Pigou's conjecture (see below).

Discussion

Sandmo (1972) and others have shown that a condition analogous to the Samuelson condition for optimal provision of public goods applies to “collective factors of production” under lump sum taxation. We know (from Stiglitz and Dasgupta, 1971, for instance) that the Samuelson condition for optimal provision of public goods applies only after an adjustment by the ratio \( \beta/\mu \) in the case with distortionary taxation. Thus, a way of restating our result is that the condition for optimal provision of a collective factor of production – in contrast to the condition for public goods - extends without any adjustment to the case with distorsionary taxation.

The factor \( \beta/\mu \) in (25) which adjusts the benefits associated with consumer preferences \( h\alpha \) but not the benefits associated with production \( \Sigma_j (h_j x_j + x'_j)c_{je} \) is the subject of a literature on a topic we may call Pigou's conjecture. Pigou’s much cited statement was:

"Where there is indirect damage" (from mobilizing revenue, Eskeland’s remark) “it ought to be added to the direct loss of satisfaction involved in the withdrawal of the marginal unit of resources by taxation, before this is balanced against the satisfaction yielded by the marginal expenditure. It follows that, in general, expenditure ought not be carried so far as to make the real yield of the last unit of resources expended by the

---

11 This may be the right place to apologize that we use an adjustment factor \( \beta/\mu \), rather than its better known inverse, often called the marginal cost of funds, MCF. MCF typically adjusts the cost side of the Samuelson condition for optimal provision of public goods (see, for instance, Auerbach, 1985). In our case, the domain of benefits is expanded beyond public goods, and then only a strict subset of benefits is to be adjusted (45), so the adjustment cannot equivalently be done on the cost side of the optimality condition. Besides, as seen in the following, adjusting benefits “as if they were taxed” (King, 1986) contributes to intuition. Sandmo (1998) highlights that some aspects of what is sometimes referred to with the MCF concept relates to whom the beneficiaries are, another reason to view adjustment of the cost side as too restrictive (see also section IV, below).
government equal to the real yield of the last unit left in the hands of the representative citizen”.

Starting with Stiglitz and Dasgupta (1971), Pigou’s conjecture that optimal provision should be lower when revenue generation is costly has been analyzed assuming the program provides pure public goods (see Atkinson and Stern, 1974, or Auerbach, 1985). For programs providing pure public goods, the exceptions to the rule that provision be adjusted downward when taxation is costly are now well known. We have assumed separability \((x_{je} = 0, \text{ all } j, \text{ see the annex})\), which rules out one class of exceptions. The remaining class is when taxed goods are predominantly inferior goods, in such a way as to make \(\sum_j t_j \frac{\partial x_j}{\partial l} < 0\). In that case, the excess burden may be falling in the level of taxation due to an income effect shifting consumption towards taxed goods. In other, perhaps more frequently observed cases, the shadow price of public revenue plays the role of cutting optimal provision short of the point where the sum of marginal benefits equal marginal costs.

The discovery in the present context is that such an adjustment (whether upward or downward) does not apply to benefits derived as cost savings in production sectors. This result is in line with Diamond and Mirrlees' classical result on production efficiency, but there are a number of reasons not to see this as a widely recognized consequence. First, our result applies independently of whether consumers or producers are the providers in the public program (e.g. whether it is consumers or producers that can reduce emission factors). Thus, if we thought of the production efficiency result as saying there should not be wedges between producers, we might have expected a wedge to apply in the case when consumers are providers – say of pollution abatement - and producers are benefiting. Diamond and Mirrlees themselves (1971) venture to discuss externalities only to the extent that they occur between consumers. Second, such an implication has to this author's knowledge not been highlighted, and studies such as Bovenberg and van der Ploeg (1994) and Williams (2000) do not make this connection. Third, applied analysis, say benefit cost analysis of roads, to this author’s knowledge does not typically distinguish between benefits to consumers and benefits to producers, even when it applies a shadow price of public funds in the analysis. It follows from the current analysis that if the benefits of a road project accrue directly to households as a pure public good and to
producing sectors in terms of reduced transportation costs, then Pigou's conjectured adjustment applies to the direct household benefits but not to transportation cost savings of firms and government.\footnote{We say typically (the correction factor may be greater than one, see below), just as we think of the vector of commodity taxes as “typically” having positive elements (though goods may be taxed at negative rates). In a context with an income tax, this becomes simpler: King, 1986, notes “Treasury should instruct those responsible for project appraisal to calculate benefits as if they were taxed at the same rate as private incomes”.

12 If households benefit not through a pure public good, but with savings of a taxed good (say gasoline, as the road improves), then those savings are to be valued by the producer price (Christiansen, 1981).

We venture now to provide intuition to this result. When the public program provides a pure public good for consumers, the role of the correction factor $\beta/\mu$ in the principle for optimal provision is typically to reduce provision of a good that would otherwise not have been taxed.\footnote{13 When the public program provides an input into production, in contrast, the benefits from provision are derived from a taxable market. Cost savings can be captured as government revenue at no distortionary costs (matching the savings by tax increases) so that distorting a program providing such benefits is redundant and costly. Put differently, for the program providing cost savings, the benefits are taxed, and for the program providing a pure public good, provision is adjusted according to the shadow price of public funds, to the same effect.} When the public program provides an input into production, in contrast, the benefits from provision are derived from a taxable market. Cost savings can be captured as government revenue at no distortionary costs (matching the savings by tax increases) so that distorting a program providing such benefits is redundant and costly. Put differently, for the program providing cost savings, the benefits are taxed, and for the program providing a pure public good, provision is adjusted according to the shadow price of public funds, to the same effect.

IV Heterogenous individuals and distribution

We briefly address the implications of heterogenous consumers. In this case, the welfare function takes the form $w(v^1, v^h)$, with consumer prices and the environment experienced at the same level by all, so $v^h = v^h(q, I^h, e)$. Now with $\beta^h = \partial w^h / \partial v^h \cdot \partial v^h / \partial I^h$, and $\alpha^h = \partial v^h / \partial e / \partial v^h / \partial I^h$, the expressions equivalent to (25) and (26) take the form:

\begin{align*}
\tau_e &= \frac{\sum_h \beta^h \alpha^h \beta^x}{\mu} + \sum_j (h x_j + x_j^p) c_{je}, \\
\sum_{s} \sum_{j} \left( \frac{\beta^x}{\mu} - 1 \right) A_{jk} x_j^e = & \frac{h - 1}{|I|} t_k, \quad \text{all } k=1,...,n,
\end{align*}
where $\bar{x}_j$ denotes average consumption of good $j$, $A$ now consists of average responses for consumers $g = 1, \ldots, h$, $\Sigma_g x_g^e / h$, and $A_{jk}$ is the cofactor of column $k$, row $j$, of matrix $A$.

Thus, the weighting scheme that applies in the condition for optimal provision (47) applies to the individual marginal benefits for the public good, but not to the distribution of consumption for the goods that benefit through cost reductions. The way distributional considerations apply in the commodity tax structure (48) is well known (See, for instance, Feldstein, 1972, or Sandmo, 1975). It is useful to notice, however, that it matters who consumes how much of a good, but demand responsiveness matters only in aggregate across consumers.

So optimal provision - or more precisely the divergence between aggregate marginal benefits and the marginal costs of provision - is independent of who consumes the goods that benefit in terms of cost reductions. The intuition behind this is exactly the same as in the preceding case with identical consumers, and illustrates how the result on aggregate production efficiency applies: To the redistributive planner, of course, it matters who consumes which goods, since the distributional characteristics of goods give her means with which to redistribute, in this case with linear commodity taxes, as given by the equations (26). Nevertheless, precisely because the planner has available a tax instrument for each commodity, programs that influence costs in industries do not need to be distorted by distributional considerations.\footnote{We may here insert a qualification that applies more generally: The program is influenced indirectly by distributional considerations (and in the previous section, by the shadow price of public revenue), in the sense that all “parameters” of our solution are functions, and thus endogenous. However, the program is not influenced directly - i.e. in terms of the intervention in the market itself. For instance, we point out that marginal costs equal aggregate marginal cost savings in industries - so there is no wedge applying in that market. But there are many such points - and the point associated with the optimal allocation depends, inter alia, on the shadow price of public revenue, and on distributional considerations. For a public good, provision may be expected to fall (rise) with redistribution to the poor if the elasticities with respect to income of willingness to pay for the public good is higher (lower) than one. Cost savings to industry $j$ are worth more the higher is output of $j$, and provision is thus influenced by redistribution according to the income elasticities for good $j$.}

This particular implication will easily get lost in applied cost benefit analysis if the analysis is conducted by taking the cost reductions directly through the productive sectors and to households (with distributional weights) according to their consumption of goods from benefiting industries. That approach may have a role to play if tax rates are considered exogenously given, or if for other reasons the assumptions behind our analysis
do not apply. However, under the assumptions of the present analysis, benefits accruing through cost reductions in productive sectors shall be accounted for equally whether the goods in question are eventually consumed by households that are rich or poor, or by government.

V. Summary and conclusions

We have analyzed optimal provision of pure public goods and nonrivalrous inputs in a context of constant returns to scale and optimal taxation. When public provision benefits productive sectors with cost savings, we show that provision shall be such that the marginal costs of provision is equal to the marginal reduction in costs in benefiting sectors. That result is as can be expected in a context of Diamond and Mirrlees' result on production efficiency.

Thus, programs which benefit production shall not be scaled down by the "penalty" from the shadow price of public revenue. This is in contrast to programs providing pure public goods (i.e. valued directly by consumers), for which an adjustment of benefits "as if they were taxed" is appropriate, typically leading to reduced provision, according to Pigou's conjecture.

The intuition behind the absence of such an adjustment when program benefits are in the form of cost savings is that these benefits are derived from markets that are otherwise taxable. The cost savings can be captured by government at no distortionary cost by increasing the tax rates for each good to match the provided cost savings.

As we have shown in Eskeland (2000), optimal abatement requires that marginal abatement costs are the same for polluting consumers, polluting producers, and government. Thus, our rules for optimal abatement in the private sector hold irrespective of whether it is consumers who pollute and abate or producers who pollute and abate - or a combination. The equivalence in marginal abatement costs between polluting producers and government (and also the equivalence with more traditional government provision, such as for roads) is a natural implication of Diamond and Mirrlees' result from 1971 on aggregate production efficiency. The equivalence between the marginal costs of emission reductions and benefits in terms of cost reductions (observed also by Bovenberg and van der Ploeg, 1994, Williams, 2000) also is such a natural implication (though not
highlighted before, we believe), since production efficiency implies that no wedges shall apply within the aggregate sector of government and production. The results for optimal provision by consumers - in the form of pollution abatement, for instance - are to our knowledge new.

We may conclude with some reflection over whether benefits of public programs are mostly for consumers or mostly benefitting producers. For environmental protection, most problems that we can think about have aspects of both: Viable ecosystems may play a role directly in our preferences, with non-use values as well as use-values, but also as inputs in the production of commodities such as timber and fish, and pharmaceuticals; present or future. Clean air and water, similarly, have value both directly to consumers and as inputs for producers. Studies of global climate change (the greenhouse problem) seem to reflect a general assumption that a significant proportion of the implications of global climate change will be reflected in production systems, and in particular in damage to agriculture (See, for example, Mendelsohn, Nordhaus and Shaw, 1994). Our approach indicates, thus, that it is important to expand the normal treatment of “the environmental good” to separate public goods in preferences from benefits in production systems.

For non-environmental programs too, an approach limited to public goods provision often will be too narrow, though stylized facts likely differ according to context. For roads, for instance, the share of commercial vehicles (trucks, buses) is predominant in poor countries. Road provision should - in an optimal tax context - not be reduced by the shadow price of public revenue to the extent that traffic is from commercial vehicles. To the extent that this is provocative (poor countries often are assumed to have high marginal costs of public revenue), the explanation comes along with the qualification: The assumption is that taxation is optimal, and that the markets served by the commercial vehicles - or the vehicles themselves - can be taxed. In some countries (in Africa, in particular) with low funding to roads as well as to other sectors, potholes are the size of small cars and truckers have accepted – or even proposed – that price increases for fuels etc. be used to fund road improvements via earmarked taxes. Thus, as a model assuming a benevolent planner and optimal taxation should be seen
merely as shedding light from one particular angle, forces may emerge in her absence to cut into gross inefficiencies, as much as they may emerge to create inefficiencies.

It may be worthwhile in the future to pursue formulations that further examine “intermediate cases” between goods that benefit consumers directly and those that reduce production costs. Christiansen (1981) provides one important step in this direction. Interesting examples would be health and education, both major public programs in all countries. Health effects of pollution, for instance, may be described partly as raising the costs of keeping us healthy, partly as simply reducing the quality of life (interestingly, the applied literature on benefit estimation uses both perspectives). With a different angle, education probably can be described partly as directly improving the well-being of the educated individual, partly as instilling social values and finally as raising the individual’s labor productivity. The latter will in part raise the individual’s earnings, but to the extent the price of human capital falls (it likely will, if rising education is wide-spread) - the gains are passed on via firms as cost reductions for their outputs. The latter case, if important, could indicate that broad public investments in human capital not be directly burdened by the shadow price of public revenue. We leave these as ideas for pursuit in the future.

Acknowledgements
The author thanks S. Devarajan, B. Jack, W. Wane, S. Rajkumar and especially Agnar Sandmo and Vidar Christiansen for comments and discussions.
REFERENCES


Williams, Robert C. III. 2000. “Environmental Tax Interactions When Pollution Affects Health or Productivity.” Mimeographed, University of Texas, Austin.
Annex to Essay III: Non separability, or when pollution influences demands

We relax the assumption of separability in preferences between the set of private goods and the environmental good in a model with four goods: \( u = u(x_0, x_1, x_2, e) \). For simplicity of exposition, we describe the problem in terms of a representative consumer, with \( h \) identical consumers. \( \beta / \mu \) is the value of income to the consumer relative to shadow price of public funds, and \( \alpha \) is the marginal rate of substitution (in preferences) between environmental quality (or emission reductions) and the numeraire good. Good one pollutes and the emission coefficient is endogenous: \( e = f_i(b_i) h x_i \). Policy instruments are commodity taxes on goods 1 and 2 and an emission tax\(^{15} \). In the case with endogenous emission coefficients, instrument number \( n+1 \), the emission tax, is not redundant, since abatement, \( b_i \), responds to the emission tax, \( \tau_e \), but not to the \( n \) commodity tax instruments (see equations 33 and 34). The first order conditions for the commodity taxes are identical to the case with exogenous emission coefficients, and thus correspond to (12):

\[
\frac{\partial L}{\partial t_i} = -\beta \left( x_i + h \alpha \frac{de}{dt_i} \right) + \mu \left[ x_i + t_1 \frac{dx_1}{dt_i} + t_2 \frac{dx_2}{dt_i} + \tau_e \frac{de}{dt_i} \right] = 0, \quad i=1,2.
\]

The first order condition for the emission tax, corresponding to (37) is:

\[
\frac{L}{\partial \tau_e} = -\beta \left( f_i x_i + h \alpha \frac{de}{d\tau_e} \right) + \mu \left[ f_i x_i + t_1 \frac{dx_1}{d\tau_e} + t_2 \frac{dx_2}{d\tau_e} + \tau_e \frac{de}{d\tau_e} \right].
\]

Rewriting these three equations to have tax rates on the left hand side, we have

\[
t_1 \frac{dx_1}{dt_1} + t_2 \frac{dx_2}{dt_1} + \tau_e \frac{de}{dt_1} = \left( \frac{\beta}{\mu} - 1 \right) x_i + h a \frac{\beta}{\mu} \frac{de}{dt_1},
\]

\[
t_1 \frac{dx_1}{dt_2} + t_2 \frac{dx_2}{dt_2} + \tau_e \frac{de}{dt_2} = \left( \frac{\beta}{\mu} - 1 \right) x_i + h a \frac{\beta}{\mu} \frac{de}{dt_2},
\]

and

\[
t_1 \frac{dx_i}{d\tau_e} + t_2 \frac{dx_2}{d\tau_e} + \tau_e \frac{de}{d\tau_e} = \left( \frac{\beta}{\mu} - 1 \right) f_i x_i + h \frac{\beta}{\mu} \frac{de}{d\tau_e}.
\]

Let us name the coefficient matrix \( A \):

\[^{15}\text{Generalizations (more goods, more polluting goods) should be confirmed but appear to apply (examine matrix } A).\]
These total derivatives describe how the quantities $x_1, x_2$ and $e$ respond to tax changes, with three underlying functional relationships: First, the uncompensated demand functions now are defined over the all-inclusive consumer prices, emissions and private income: $x_i = x_i(1, q_1, q_2, e, I)$. Second, there is the emission function, above, and finally there is equation (33) describing how the all-inclusive consumer prices for good one include tax rates and abatement, $t_1, \tau_e f_1, b_1$, and for good two includes $t_2$. Solving (first for the differentials of $q$, then for the quantities as functions of $q$ and $\tau_e$), the elements in $A$ are as follows:

$$A = \begin{bmatrix}
\frac{dx_1}{dt_1} & \frac{dx_2}{dt_1} & \frac{dx_e}{dt_1} \\
\frac{dx_1}{dt_2} & \frac{dx_2}{dt_2} & \frac{dx_e}{dt_2} \\
\frac{dx_1}{d\tau_e} & \frac{dx_2}{d\tau_e} & \frac{dx_e}{d\tau_e}
\end{bmatrix}.$$  

It can be shown that the determinant of $A$ is

$$|A| = \frac{f_{1b} b_1 x_1}{1 - f_{1b} x_{1e}} (x_{11} x_{22} - x_{12} x_{21}).$$  

For the commodity tax rate $t_1$, we have

$$t_1 = \frac{\beta}{\mu - 1} \begin{bmatrix} x_1 & \frac{dx_2}{dt_1} & \frac{dx_e}{dt_1} \\
\frac{dx_1}{dt_2} & \frac{dx_2}{dt_2} & \frac{dx_e}{dt_2} \\
\frac{dx_1}{d\tau_e} & \frac{dx_2}{d\tau_e} & \frac{dx_e}{d\tau_e}
\end{bmatrix} + \frac{h a \beta}{\mu} \begin{bmatrix} \frac{de}{dt_1} & \frac{dx_2}{dt_1} & \frac{dx_e}{dt_1} \\
\frac{dx_1}{dt_2} & \frac{dx_2}{dt_2} & \frac{dx_e}{dt_2} \\
\frac{dx_1}{d\tau_e} & \frac{dx_2}{d\tau_e} & \frac{dx_e}{d\tau_e}
\end{bmatrix} \begin{bmatrix} x_1 & \frac{dx_2}{dt_1} & \frac{dx_e}{dt_1} \\
\frac{dx_1}{dt_2} & \frac{dx_2}{dt_2} & \frac{dx_e}{dt_2} \\
\frac{dx_1}{d\tau_e} & \frac{dx_2}{d\tau_e} & \frac{dx_e}{d\tau_e}
\end{bmatrix}^{-1} \begin{bmatrix} x_1 & \frac{dx_2}{dt_1} & \frac{dx_e}{dt_1} \\
\frac{dx_1}{dt_2} & \frac{dx_2}{dt_2} & \frac{dx_e}{dt_2} \\
\frac{dx_1}{d\tau_e} & \frac{dx_2}{d\tau_e} & \frac{dx_e}{d\tau_e}
\end{bmatrix}.$$
The determinant in the numerator can be shown to be \( f_{13} b_{11} x_1 (x_1 x_{22} - x_2 x_{21})/(1 - f_{x e}) \), so

\[
\begin{align*}
\tau_1 &= \frac{(\beta - 1)}{\mu} \frac{(x_1 x_{22} - x_2 x_{21})}{x_1 x_{22} - x_{12} x_{21}}, \\
\tau_2 &= \frac{(\beta - 1)}{\mu} \frac{(x_2 x_{11} - x_1 x_{12})}{x_1 x_{22} - x_{12} x_{21}}.
\end{align*}
\]

Thus, in terms of the formulas for optimal commodity taxation, the proposition that emissions and their taxation do not interfere with commodity tax principles generalizes to the case of an environmental good nonseparable from private goods.

We go through similar steps to arrive at an expression for the emission tax:

\[
\tau_e = \frac{h \alpha \beta}{\mu} + \frac{-x_{1e} \left(\frac{\beta - 1}{\mu} \right) (x_1 x_{22} - x_2 x_{21}) - x_{2e} \left(\frac{\beta - 1}{\mu} \right) (x_1 x_{12} - x_2 x_{11})}{x_1 x_{22} - x_{12} x_{21}}.
\]

Additional insight can be gained by substituting the expressions for the commodity tax rates (A.7) into (A.8):

\[
\tau_e = h \alpha \frac{\beta}{\mu} + (t_1 \cdot (-x_{1e}) + t_2 \cdot (-x_{2e})).
\]

In (A.9), we have placed \(-x_{1e}\) inside an inner parenthesis, to see the tax rates multiplied by the response to a public good, as opposed to a bad. (A.9) shows quite clearly how adjustments of the emission tax from Pigovian levels arise because of the effects that public goods provision has on proceeds from commodity taxes. If we think of the finance ministry and the environment ministry as two departments in a corporation, then it appears as if the environment ministry is credited with the support that its product (environmental quality, or minus \( e \)) gives to the revenues from commodity taxes. Using the corporation analogy further, one could envisage the corporation using such principles in its internal processes to make sure that appropriate resources are allocated to a department marketing the brand name of the corporation as a whole. The analogy between the benevolent planner's public sector and the corporation departs at one point:
For the planner, these effects via revenues represent adjustments to a Pigovian principle, while for the corporation, the effects via product revenues is the whole story. For the corporation, the separability assumption would demolish the budget for a corporate brand name: if the value of the program is not captured through profits from sold products, then it is of no value to the corporation.

Using the Slutsky equation and the symmetry of the compensated demand derivatives $s_{12} = s_{21}$, the commodity tax structure (A.7) implies:

$$
\frac{\sum_j t_j s_{1j}}{x_1} = \frac{\sum_j t_j s_{2j}}{x_2} = \frac{\beta}{\mu} - 1 + \sum_j t_j x_{ji}, \ j = 1, 2,
$$

where $x_{ji}$ is the marginal propensity to spend on good $j$ out of personal income.

(A.10) states that for an allocation near the first best (i.e. near the one implemented when there are no distortions $\beta/\mu = 1$), the compensated demand reductions caused by the commodity taxes shall be in equal proportions for the taxed goods (as shown by Samuelson, 1951).

Also for the emission tax, let us examine implications further by replacing the demand derivatives with the Slutsky equation:

$$
A_{10} = -f_{1b} + \frac{1}{\mu} \left[ t_1 (s_{1e} - h\alpha x_{1e}) + t_2 (s_{2e} - h\alpha x_{2e}) \right]
$$

Using symmetry $s_{ie} = s_{ei}$ yields

$$
A_{13} = \frac{t_1 s_{1e} + t_2 s_{2e}}{h\alpha} = \frac{1}{\mu} \left[ -f_{1b} + t_1 x_{1e} + t_2 x_{2e} \right].
$$

The left hand side of this expression is analogous to the left hand side of (A.10) above. It expresses the proportionate reduction in compensated willingness to pay for the public good (defining the good as a reduction in pollution) that is caused by the commodity

---

16 This is traditional, from Samuelson, 1951, so we suppress deduction.
taxes. The right hand side is similar – but not equal – to the right hand side of (A.10), but the two are close (as noted by King, 1986, in the case of a government provided public good), when we examine allocations close to the first best equilibrium. In first best \( \tau_e = -1/f_{1b} = h\alpha \), the two right hand sides are equal, and the commodity taxes cause the same proportionate reduction in willingness to pay for the public good as in demand for the taxed goods. In the case of a public good provided directly by government, the allocation near the first best would have to be supported by exogenous government revenue. In the case of pollution, positive revenue can be generated by the instrument(s) implementing the allocation (as pointed out by Sandmo, 1975), and we could imagine the proceeds redistributed as lump sum transfers or being used to provide another public good.
<table>
<thead>
<tr>
<th>Title</th>
<th>Author</th>
<th>Date</th>
<th>Contact for paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>WPS2493 Annuity Markets in Comparative Perspective: Do Consumers Get Their Money's Worth?</td>
<td>Estelle James, Dimitri Vittas</td>
<td>November 2000</td>
<td>A. Yaptenco</td>
</tr>
<tr>
<td>WPS2494 The Relevance of Index Funds for Pension Investment in Equities</td>
<td>Ajay Shah, Kshama Fernandes</td>
<td>November 2000</td>
<td>A. Yaptenco</td>
</tr>
<tr>
<td>WPS2495 The Australian Annuity Market</td>
<td>David M. Knox</td>
<td>November 2000</td>
<td>A. Yaptenco</td>
</tr>
<tr>
<td>WPS2496 Perspectives on the Sources of Heterogeneity in Indian Industry</td>
<td>Somik V. Lall, G. Chris Rodrigo</td>
<td>November 2000</td>
<td>R. Yazigi</td>
</tr>
<tr>
<td>WPS2499 The Treatment of Non-Essential Inputs in a Cobb-Douglas Technology: An Application to Mexican Rural Household-Level Data</td>
<td>Isidro Soloaga</td>
<td>December 2000</td>
<td>L. Tabada</td>
</tr>
<tr>
<td>WPS2500 Investigating Corruption</td>
<td>Canice Prendergast</td>
<td>December 2000</td>
<td>H. Sladovich</td>
</tr>
<tr>
<td>WPS2501 Anti-Corruption Policies and Programs: A Framework for Evaluation</td>
<td>Jeff Huther, Anwar Shah</td>
<td>December 2000</td>
<td>A. Santos</td>
</tr>
<tr>
<td>WPS2502 Implications of the Currency Crisis for Exchange Rate Arrangements in Emerging East Asia</td>
<td>Masahiro Kawai, Shigeru Akiyama</td>
<td>December 2000</td>
<td>A. Azarcon</td>
</tr>
<tr>
<td>WPS2503 Proposed Strategy for a Regional Exchange Rate Arrangement in Post-Crisis East Asia</td>
<td>Masahiro Kawai, Shinji Takagi</td>
<td>December 2000</td>
<td>A. Azarcon</td>
</tr>
<tr>
<td>WPS2504 Social Capital and the Quality of Government: Evidence from the United States</td>
<td>Stephen Knack</td>
<td>December 2000</td>
<td>P. Sintim-Aboagye</td>
</tr>
<tr>
<td>Title</td>
<td>Author</td>
<td>Date</td>
<td>Contact</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>--------------</td>
<td>---------------</td>
<td>----------------</td>
</tr>
<tr>
<td>WPS2505 Family Altruism and Incentives</td>
<td>Roberta Gatti</td>
<td>December 2000</td>
<td>R. Bonfield</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>31248</td>
</tr>
<tr>
<td>WPS2506 Ethnicity and Wage Determination in Ghana</td>
<td>Abigail Barr</td>
<td>December 2000</td>
<td>R. Bonfield</td>
</tr>
<tr>
<td></td>
<td>Abena Oduro</td>
<td></td>
<td>31248</td>
</tr>
</tbody>
</table>