Environmental and Distortionary Taxes in Perspective

William K. Jaeger
Department of Economics
Williams College
Williamstown MA 01267
Wjaeger@williams.edu
(413) 597-3213

August 9, 2001

I. Introduction

Recent contributions to environmental tax theory have concluded that in a second-best setting with preexisting distortionary taxes, the optimal environmental tax will generally lie below the marginal social damage from pollution, even if the revenues are employed to cut preexisting taxes (Bovenberg and de Mooij 1994). This finding has been reinforced by numerical estimates and optimal tax rules suggesting that the Pigouvian Principle requires modification in the presence of distortionary taxes (Bovenberg and Goulder 1996). These results have also been interpreted as demonstrating that the "double dividend hypothesis"—when interpreted as implying that increased revenue requirements should be met by raising the environmental tax—is flawed (Fullerton 1997; Bovenberg and de Mooij 1997).

The term "double dividend" became widely used following its use by Pearce (1991) to refer to the idea that a revenue-neutral carbon tax would reduce the marginal excess burden from distortionary taxes, and that this would make the effective carbon tax 20 to 50 percent lower than its nominal value. This idea is consistent with earlier literature making a similar point (see for example Terkla 1984; Lee and Misiolek 1986). The reasoning underlying the double dividend idea is straightforward. If, in a first-best world, the optimal environmental tax should equal the marginal social damage from pollution, then in a second-best world with preexisting revenue-motivated taxes, the revenues from a pollution tax can be used to lower the distortionary effects of these preexisting taxes, and this additional benefit, or "double dividend," makes it optimal to raise the environmental tax above the marginal social damage.[1]

[1] Goulder (1995) proposes an alternative definition of the double dividend claim focused on the total rather than the marginal welfare effects of revenue-neutral environmental taxation. He separates the environmental benefits of green tax reform from changes in what he calls the "gross" (but non-environment) costs associated with taxation. This "strong form" of the double dividend claim, however, differs from the suggestions made by Pearce (1991). With the present model it is also unclear whether separating the environmental gains from non-environmental "gross costs" is either possible, or instructive.

In contrast to this double dividend reasoning, the recent literature suggests that when faced with an increase in revenue requirements, government will actually find it optimal to reduce pollution taxes because tax distortions have been exacerbated by a previously unrecognized "tax interaction effect" (Bovenberg and de Mooij 1994; Parry 1995; Bovenberg and Goulder 1996; Fullerton 1997). These authors recognize the validity of the notion about using revenues from environmental taxes to lower the cost of the tax system (see Bovenberg and de Mooij 1994), but the intuitive inference that the optimal environmental tax will then lie above marginal social damage is contradicted by their results indicating that the optimal environmental tax lies below the marginal social damage. If correct these results not only caste doubt on the double dividend hypothesis, but they are seen as having profound implications not just for environmental policy but for many other non-environmental policy issues as well (Parry and Oates 2000).

This note argues that the findings and subsequent inferences in the recent tax interaction literature are due to the use of a particular definition of "marginal social damage," one that is not based on the marginal rate of substitution between income and the environment for the social planner's problem. In particular, this definition values an incremental unit of income in a way that does not reflect its Pareto efficient use. The present analysis finds that when the social marginal utility of income is used as the numeraire unit of value, so that marginal social damage reflects the marginal rate of substitution between income and the environment, the relationship between the optimal environmental tax and marginal social damage is consistent with the intuition of the double dividend hypothesis: the optimal environmental tax will generally lie above the marginal social damage, and it will rise with increased revenue requirements.

I. The model

Two endowments, household time, T, and the environment, E, are allocated in a model similar to Bovenberg and de Mooij (1994) such that time, in the form of labor supply, L, is the only input into a linear production technology given by NL = NC + ND + G. Labor productivity is assumed to be constant and equal to one for N identical households, and where output can be used for public consumption, G, as well as for consumption of "clean" and "dirty" private consumption commodities denoted by C and D, respectively. Units are defined so that the constant rates of transformation between the three produced commodities and leisure are unity. All private commodities are expressed in per capita terms.

Household utility is given as U = u(C, D, V, G, E) where V is leisure and E is the environment. Households take the supply of G and E as given, and optimize over V, C, and D. Full income for each household is their time endowment, T, which can be allocated between leisure and labor supply such that L=T-V. Environmental quality is a function of the consumption of the polluting good, D, where

$$E = e(ND), \qquad \frac{de}{d(ND)} < 0 \tag{1}$$

The household maximization problem is written as

$$\begin{aligned} &\underset{C,D,V}{MAX}: \quad u(C,D,V,G,E) \\ &s.t.\ (\lambda) \quad (T-V)(1-t_L) = C + (1+t_D)D \end{aligned}$$

where distortionary taxes may be introduced to finance G and corrective taxes to improve the environment. Preferences are restricted in this model so that a single tax on labor supply may substitute for, and be equivalent to, optimal (uniform) revenue-motivated taxes on C and D. Specifically, the utility function is assumed to be homothetic in C and D, and separable in leisure and environmental quality. With these assumptions in place, the optimal tax on D will correspond to the environmental component of the overall optimal tax program (Bovenberg and de Mooij 1994; Bovenberg and Goulder 1996). Note that with a labor tax, the household's budget constraint has been renormalized to reflect units of net income rather than gross income. This change does not alter any real variables or the optimum allocation (Fullerton 1997; Schöb 1997), and the social planner's problem remains one of allocating units of gross income among private and public uses independent of whether taxes are imposed at the household level on income or expenditures.[2]

To maximize social welfare, W, the planner's problem is one of solving:

$$\begin{aligned} \underset{t_L, t_D}{MAX}: \quad & N\left[\max_{C,D,V} : u(C,D,V,G,E) \;\; s.t. \;\; (T-V)(1-t_L) = C + (1+t_D)D\right] \\ s.t. \;\; (\mu) \quad & N(t_L(T-V) + t_D D) = G \\ (\pi) \quad & E = e(ND) \end{aligned} \tag{2}$$

where μ denotes the implied Lagrange multiplier for the revenue constraint corresponding to the shadow value of a unit of public revenue. Similarly π denotes the implied Lagrange multiplier for the constraint on environmental quality so that the social marginal utility of environmental quality can be expressed as $\pi = N\partial u/\partial E$.

From (2) we derive an expression for the social marginal utility of income, $\alpha \equiv \partial W/\partial T$, as

$$\alpha = \tilde{\lambda}(1-t_L) + \mu\left(t_L \frac{\partial L}{\partial T} + t_D \frac{\partial D}{\partial T}\right) + N\frac{\partial U}{\partial E}\frac{de}{d(ND)}\frac{\partial D}{\partial T} \tag{3}$$

This expression is analogous to the standard definition for the social marginal utility of income (Diamond 1985), but with the addition of the environmental damage term. The Lagrange multiplier

[2] For a normalization involving only expenditures taxes, the relevant comparison for our purposes would be between the marginal social damage from pollution and the difference between the optimal taxes on the two good (i.e., $t^*_D - t^*_C$), which would correspond to the environmental component of the optimal tax on D.

on the household's budget is denoted by $\tilde{\lambda}$ and interpreted as the private marginal utility of income. A tilde over a variable indicated that the normalization is in net income units rather than gross income units due to the use of the income tax t_L. The correspondence between gross and net income units is a function of the income tax rate where, for example, $\lambda=\tilde{\lambda}(1-t_L)$. In this setting a one unit change in net income corresponds to a greater than one unit change in gross income. The Lagrange multiplier on the social planner's revenue constraint, μ, is not affected by the income tax normalization of the household budget constraint; it continues to reflect the shadow value of public revenues in gross income units.[3]

The marginal rate of substitution between income and the environment is defined here as the marginal utility of a unit increase in environmental quality divided by the marginal utility of a unit increase in income. This definition implicitly takes a unit of income to be the numeraire. Given (3) and the identity $\pi = N\partial u/\partial E$, this marginal rate of substitution can be expressed as

$$MRS_{TE} = \frac{\partial W / \partial E}{\partial W / \partial T} = \frac{N \dfrac{\partial u}{\partial E}}{\alpha} \tag{4}$$

For the incremental environmental damage associated with a unit of D, we can utilize (1), (3) and (4) to define the marginal social damage (MSD) as

$$MSD = MRS_{TE} \frac{-de}{d(ND)} = \frac{N \dfrac{\partial u}{\partial E} \dfrac{-de}{d(ND)}}{\tilde{\lambda}(1-t_L) + \mu\left(t_L \dfrac{\partial L}{\partial T} + t_D \dfrac{\partial D}{\partial T}\right) + N \dfrac{\partial U}{\partial E} \dfrac{de}{d(ND)} \dfrac{\partial D}{\partial T}} \tag{5}$$

[3] To be clear on this potentially confusing point, consider a model expressed entirely in gross income units with one good, C, a uniform commodity tax t_C and labor supply L. Given a consumer price $P_C=1+t_C$, household consumption $C=(1/(1+t_C))L$. Total revenues are expressed as $G = t_C C = N(t_C/(1+t_C))L$. Now consider an equivalent model in net income units where a labor tax, t_L, replaces the commodity tax such that $(1- t_L)=1/(1+t_C)$. The government budget constraint in this case, as with the model in (1), is expressed as $G= N(t_L L)$. The expression for the equivalence of these two normalizations, $(1- t_L)=1/(1+t_C)$, can also be written as $t_L=tc/(1+t_C)$. In which case we can see that $N(t_L L)= N(t_C/(1+t_C))L$. This verifies that revenues have the same numerical value with either normalization of the household budget constraint. The government budget constraint is expressed in gross income units for both.

In contrast to (5), the recent "tax interaction" literature has defined the marginal damage from pollution in a way that corresponds to the sum of households' "marginal private damages"(MPD). This definition can be expressed as

$$MPD = \frac{N \frac{\partial u}{\partial E} \frac{-de}{d(ND)}}{\tilde{\lambda}} . \qquad (6)$$

The entire denominator in (6) corresponds to the first of three terms in the denominator of (5), although they reflect different normalizations given $\lambda = \tilde{\lambda}(1-t_L)$.[4] Both terms reflect the private marginal utility of income, or the direct increase in utility to households from a unit increase in income. The second term in the denominator of (5) reflects the utility from the increased tax revenues due to additional labor supply and expenditure. And the third term reflects the environmental consequences of a unit change in income pertaining to the marginal propensity to pollute.

In the first-best setting the two expressions in (5) and (6) will be equal. To see this, note that in the absence of a binding revenue requirement, $t_L=0$ and $\lambda=\tilde{\lambda}$, so that the expression in (6) will equal marginal social damage or the first-best Pigouvian tax rate, $t^*_D=N(\partial u/\partial E)(-de/d(ND))/\lambda$. For the expression in (5), note also that in first-best $\mu=\lambda$ and that the pollution tax, t_D, must equal the Pigouvian tax, $N(\partial u/\partial E)(-de/d(ND))/\lambda$. Substituting this expression for t_D in the denominator of (5), the second and third terms cancel leaving only λ. Thus (5) and (6) are identical in the first-best case.

In the second-best setting, marginal social damage as expressed in (5) will differ from the first-best expression due to the presence of a labor tax and the consequent divergence between λ

---

[4] Differences in the normalization of income units must also be born in mind when interpreting expressions such as $\mu/\tilde{\lambda}$, frequently referred to as the marginal cost of public funds, where μ reflects units of gross income while $\tilde{\lambda}$ reflects net income units.

and μ. The definition of MPD in (6) is based on the summing of marginal private damages across *N* households; it does not emerge directly in the course of solving the social planner's optimal tax problem. And while the summation of values across households reflects recognition of the non-rival characteristic of the environment, this is but one source of divergence between the social and private marginal utility of either environmental quality or income.

Notably in both Bovenberg and de Mooij (1994) and Bovenberg and Goulder (1996), this definition is used interchangeably to represent marginal social damages in both the first-best and second-best settings. In Bovenberg and de Mooij, the welfare effect for a revenue-neutral change in the tax mix (i.e., $dG=0$ and $t_L>0$) demonstrates that welfare is raised when t_D is lowered below the "Pigouvian rate" as defined by MPD. This conclusion is no doubt correct. However, MPD exceeds MSD because the social marginal utility of income exceeds the private marginal utility of income.

Direct comparison of the denominators in (5) and (6) is complicated by the different normalizations. Since households are paying the environmental tax out of net income, we want to express all values in these units. We can express α in net income units as

$$\tilde{\alpha} \equiv \frac{\alpha}{(1-t_L)} = \tilde{\lambda} + \frac{\mu}{(1-t_L)}\left(t_L \frac{\partial L}{\partial T}\right) + \frac{1}{(1-t_L)}\left(\mu t_D + N\frac{\partial u}{\partial E}\frac{de}{d(ND)}\right)\frac{\partial D}{\partial T}. \qquad (7)$$

The second term on the right-hand side of (7) is positive given a positive marginal propensity to supply labor if C and D are normal goods. The sign of the third term on the right-hand side depends on the value of the environmental tax. Setting t_D equal to (6), as Bovenberg and de Mooij have done, we can deduce that the expression in parentheses will also be positive since $\mu/\tilde{\lambda}>1$, making the first (positive) term within the parentheses greater than the second (negative) term. Hence the entire third term on the right-hand side of (7) will be positive provided D is a normal good, so that $\partial D/\partial T>0$. This confirms that the social marginal utility of income exceeds the private marginal utility of income (whether normalized in gross or net income units), making MPD>MSD. The analysis in Bovenberg and de Mooij, therefore, begs the question of whether the optimal environmental tax will be below marginal social damage if it is defined based on the social marginal utility of income.

II. The optimal environmental tax and marginal social damage

Extending the analysis in Bovenberg and de Mooij, Bovenberg and Goulder (1996) derive an optimal environmental tax expression that can be written for the current model and notation as

$$t_D^* = \left[\frac{N \frac{\partial u}{\partial E} \frac{-de}{d(ND)}}{\tilde{\lambda}} \right] \frac{\tilde{\lambda}}{\mu} . \tag{8}$$

Bovenberg and Goulder identify the bracketed term as the "marginal social benefit" of emissions reduction, or the "Pigouvian rate," expressed, appropriately, in net income units since t^*_D is to be paid by households out of net income. The marginal cost of public funds (MCPF) is defined as $\mu/\tilde{\lambda}$ and is assumed to be greater than one. It follows therefore that $\tilde{\lambda}/\mu<1$ so that t^*_D < MPD. On this basis, Bovenberg and Goulder conclude that "the presence of distortionary taxes requires a modification of the Pigouvian principle" and that "the higher the MCPF, the greater the cost of public consumption goods, including the public good of environmental quality. When these goods are more costly, the government finds it optimal to cut down on public consumption of the environment by reducing the pollution tax" (Bovenberg and Goulder 1996, p. 987).

To express the optimal environmental tax in terms of MSD, (8) can be modified. Substituting the identities $\tilde{\lambda}/\tilde{\lambda} = \tilde{\alpha}/\tilde{\alpha}$ and $\tilde{\mu} = \mu(1-t_L)$, we can represent (8) as

$$t_D^* = \left[\frac{N \frac{\partial u}{\partial E} \frac{-de}{d(ND)}}{\tilde{\alpha}} \right] \frac{\tilde{\alpha}}{\tilde{\mu}(1-t_L)} \tag{9}$$

where now t^*_D is defined in relation to MSD, and all terms are expressed in net income units. An equivalent expression can be derived directly from the kind of model evaluated by Sandmo (1975) (see Appendix).

The expression in (9) suggests that the optimal environmental tax may be higher or lower than MSD given the relationships $\tilde{\mu} > \tilde{\alpha}$ and $(1-t_L)<1$. Thus, it remains an empirical question; it will depend on the actual tax rates and the marginal utility of income and public revenues.

By substituting the optimal tax expression from (9) for t_D into (7) and rearranging, the third term on the right-hand side of (7) reduces to

$$\frac{1}{(1-t_L)} \left(N \frac{\partial u}{\partial E} \frac{-de}{d(ND)} + N \frac{\partial u}{\partial E} \frac{de}{d(ND)} \right) \frac{\partial D}{\partial T}$$

which by inspection must equal zero. This result provides an intuitive way to understand optimal taxation by thinking about the allocation of an incremental unit of income: at the optimum, an incremental unit of income should be allocated so that the marginal utility derived from additional environmental tax revenues is just equal to the social disutility of the incremental environmental damage.

At the optimum we can drop the third term in (7) and divide both sides by $\tilde{\mu}$ to get

$$\frac{\tilde{\alpha}}{\tilde{\mu}} = \frac{\tilde{\lambda}}{\mu}(1-t_L) + \left(t_L \frac{\partial L}{\partial T} \right). \tag{10}$$

Different numerical models will produce different estimates of the relationship between t^*_D and MSD. One such estimation can be easily made using (9) and (10), and drawing on the numerical model in Bovenberg and Goulder (1996). For Bovenberg and Goulder’s central case where $t_L = 0.4$, they report that $\mu/\tilde{\lambda}$ =1.16, and that the uncompensated and compensated labor supply elasticities are 0.15 and 0.94, respectively. From the Slutsky equations, the difference between the compensated and uncompensated labor supply elasticities imply a value $\partial L/\partial T$= 0.79. With these numerical values we can evaluate the right-hand side of (10) to get $\tilde{\alpha}/\tilde{\mu}$ =0.833. Plugging this result into (9) gives us the result that the optimal environmental tax will exceed MSD by 38 percent (t^*_D=1.38($\pi/\tilde{\alpha}$). Note that this relationship is based on the same model and parameters, indeed the same optimal tax with which Bovenberg and Goulder concluded that the optimal environmental tax fell short of MPD by 12 percent (t^*_D=0.88($\pi/\tilde{\lambda}$). If follows that MPD =1.57MSD. Thus, it is the definition of marginal environmental damages that distinguishes the current interpretations from those in the tax interaction literature, not differences in the optimal environmental tax.

The inference that a rise in revenue requirements will lower the optimal environmental tax is not supported by the current analysis. The source of confusion on this point can be seen by comparing the first-best case where t^*_D=MSD=MPD, to a model with a 40 percent income tax where MPD is estimated to be 57 percent higher than MSD due to the decline in the private marginal utility of income relative to the social marginal utility of income. Although the use of MPD as a benchmark may make it appear as though the optimal environmental tax declines with rising revenue requirements, in fact both the optimal environmental tax and MPD will rise. The appearance of a reduction in the optimal environmental tax when the revenue requirement increases is due to the fact that the value of MPD rises faster than the optimal environmental tax.[5]

[5] The relationship between MPD and the optimal environmental tax also depends on the type of externality. In the case of a production externality where environmental quality affects labor productivity, MPD is found to be lower rather than higher than MSD, and as a result of this the optimal environmental tax is estimated to be 53 percent higher than MSD but 73 percent higher than MPD based on a numerical model for the US economy (Jaeger 2001).

III. Conclusion

When marginal social damage is defined using the social marginal utility of income as the numeraire, the present analysis finds that the relationship between the optimal environmental tax and the marginal social damage will depend on preexisting tax rates and the ratio of the marginal utility of income to the marginal utility of public revenues. Based on estimates of these parameters for the U.S. economy found in the literature, the optimal environmental tax is shown to exceed marginal social damages by 38 percent—consistent with the estimate from Pearce (1991), and with the reasoning in the early literature (see Tullock 1967; Terkla 1984; Lee and Misiolek 1986).

Since, in the first-best case with no binding revenue requirement, the optimal environmental tax equals marginal social damage, it follows that the optimal environmental tax rises with increasing and binding revenue requirements over the range under consideration. Intuitively, like other goods and services in the economy, environmental services should be priced at their social cost in a first-best setting, and priced above their social cost in a second-best setting in which Ramsey taxes have been added to raise revenue.

This intuitive result appears to be contradicted in the recent literature where marginal social damage has been defined using the private marginal utility of income. By valuing income in private terms rather than social terms, the numeraire in these analyses does not reflect a Pareto efficient allocation: it does not correspond to the difference between *ex ante* and *ex post* Pareto states for a unit change in income. Indeed, for a decrement in income, this definition implies an infeasible combination of reductions in consumption of taxed goods and labor supply, while maintaining the original provision of public goods (or transfers) despite a loss of tax receipts.

In standard optimal tax models, we understand that as revenue requirements rise the proportion of incremental income paid in tax receipts rises and the proportion going to private consumption declines, thereby causing the private marginal utility of income to fall relative to the social marginal utility of income. In the case of an amenity externality, holding marginal environmental damage constant in utility units, a rising revenue requirement and falling private marginal utility of income causes the marginal private damage to rise by more than the increase in the optimal environmental tax, which creates the appearance that the optimal environmental tax falls with increased revenue requirements. However, when marginal social damage is viewed from society's collective perspective this apparent relationship is reversed and the results anticipated by the double dividend hypothesis are revealed.

The theoretical issued raised here and in the recent literature serve as a reminder of the important distinction between the private and shadow value of income. As an empirical matter, valuing environmental damages constitutes a challenge to economists for a variety of reasons. The shadow pricing of income represents yet another consideration to be attended to. Although measuring marginal private damages may be simpler in some situations where values have been estimated using contingent valuation techniques or hedonic studies, even these methods may reflect values based on tax-deductible expenditures, pre-tax income or gross rather than net wage differentials. In such cases the conversion to equivalent shadow values may be useful simply to maintain comparability among these different methods. Indeed, for cases where environmental damage reduces the productivity of labor or other assets, these losses typically measure gross income costs, including the tax payments that would otherwise not be included as part of marginal private damages. These measurement issues notwithstanding, the perspectives offered here are consistent with the important notion that the provision of public goods and the protection of environmental quality are complementary rather than competing goals of government, and that increased revenue requirements should be met by raising, not lowering, the environmental tax.

References

Auerbach, A. "The Theory of Excess Burden and Optimal Taxation" in A. Auerbach and M. Feldstein, eds., Handbook of public economics, Vol. 1. Amsterdam: North-Holland, 1985.

Bovenberg, A. L. and L. H. Goulder. "Optimal Environmental Taxation in the Presence of Other Taxes: General Equilibrium Analysis." American Economic Review, September 1996, 86(4): 985-1000.

Bovenberg, A. L. and R. A. de Mooij. "Environmental Levies and Distortionary Taxation." American Economic Review, September 1994. 84(4): 1085-89.

Bovenberg, A. L. and R. A. de Mooij. "Environmental Levies and Distortionary Taxation: Reply." American Economic Review, March 1997. 87(1): 252-253.

Diamond, P.A. A many-person Ramsey tax rule. Journal of Public Economics 1975, 4: 335-342.

Fullerton, D. "Environmental Levies and Distortionary Taxation: Comment." American Economic Review, March 1997, 87(1): 245-51.

Goulder, L.H. Environmental taxation and the double dividend: a reader's guide. International Tax and Public Finance, 1995, 2: 157-83.

Jaeger, William K. Carbon taxation when climate affects productivity. *Land Economics*, 2001 forthcoming.

Lee, D.R. and W. S. Misiolek. 1986."Substituting pollution taxation for general taxation: some implications for efficiency in pollution taxation." *Journal of Environmental Economics and Management* 13(4): 228-47.

Parry, I.W.H. 1995. "Pollution taxes and revenue recycling." *Journal of Environmental Economics and Management* 29(3): 564-77.

Parry, I.W.H. and W. E. Oates. 2000. "Policy analysis in the presence of distorting taxes." *Journal of Policy Analysis and Management* 19(4): 603-13.

Pearce, D. 1991. "The role of carbon taxes in adjusting to global warming." *Economic Journal* 101(407): 938-48.

Sandmo, A. "Optimal Taxation in the Presence of Externalities." Swedish Journal of Economics, 1975. Vol. 77, pp. 86-98.

Schöb, R. "Environmental taxes and pre-existing distortions; the normalization trap." International Tax and Public Finance, 1997, 4:167-176.

Terkla, D. 1984. "The Efficiency Value of Effluent Tax Revenues." *Journal of Environmental Economics and Management* 11(2): 107-23.

Appendix

For notational convenience, the general derivations below are based on a model with *n*+1 goods. This is followed by additional derivations for the special case of one clean and one polluting good, and where restrictions on preferences have been assumed. The first part generates implicit optimal tax rules taking an approach parallel to Sandmo (*Swedish J. of Economics,* 1975).

The general problem can be stated as one where *m* identical individuals maximize utility U = u(X_0, X_1, …, X_Z,…, X_n, E) for goods j = 0, …, *n*, where leisure is X_0 and labor is supplied from a time endowment normalized to equal one so that labor supply equals 1-X_0. Units are chosen for goods and income so that all pre-tax prices equal one, and where there are *n*-1 non-polluting *X* goods (excluding leisure) and one good X_z which produces an environmental externality. The consumption of X_Z erodes the environment, *E*, where $E = e(mX_Z)$ and $de/d(mX_Z) < 0$.

We define aggregate output as $mh(1\text{-}X_0) = \Sigma m X_i$, where labor productivity, *h*, is constant and where transfers to households of *mG* are financed through collection of tax revenues. The environment E enters the utility function directly.

Each household's maximization problem can be stated as

$$\underset{X_0 \ldots X_n}{Max} : \quad u(X_0, X_1, \ldots X_n, E)$$

$$s.t. \quad h(1 - X_0) + G = \sum_{j=1}^{n} (1 + t_j) X_j \qquad \text{[A1]}$$

The Lagrangian expression for each household taking E and G as given is thus

$$\mathcal{L} = u(X_0, X_1, \ldots X_n, E) + \lambda \left[h(1 - X_0) + G - \sum_{j=1} (1 + t_j) X_j \right] \text{ for } j = 1, \ldots, z, \ldots, n. \qquad \text{[A2]}$$

In this case the social planners problem is

$$\underset{t_1 \ldots t_n}{Max}: \quad m\left[u(X_0, X_1, \ldots X_n, E) \quad s.t. \quad h(1 - X_0) + G = \sum_{j=1}^{n} (1 + t_j) X_j \right]$$

$$s.t. \qquad m\sum_{j=1}^{n} t_j X_j = G \qquad \text{[A3]}$$

$$E = e(mX_Z)$$

The household's indirect utility function is $V(p_0, p_1, \ldots p_n, E) = u(x(p_0, p_1, \ldots p_n), E)$, so we can state the social optimization problem as the Lagrangian equation

$$\mathcal{L} = mu\big(v(p_0, p_1, \ldots p_n, E)\big) + \mu\left[m\sum_{j=1}^{n} t_j X_j - G \right]$$

The planner's first-order conditions are

$$-\lambda X_j + \mu\left[\sum_i t_i \frac{\partial X_i}{\partial p_j} + X_j \right] + m\left[\frac{\partial U}{\partial E} \frac{de}{d(mX_Z)} + \mu \sum_i t_i \frac{\partial X_i}{\partial E} \frac{de}{d(mX_Z)} \right] \frac{\partial X_Z}{\partial p_j} = 0 \quad \forall\ j \neq 0 \qquad \text{[A4]}$$

so that the marginal damage denoted as π is defined as

$$\pi = m\left[\frac{\partial U}{\partial E} \frac{de}{d(mX_Z)} + \mu \sum_i t_i \frac{\partial X_i}{\partial E} \frac{de}{d(mX_Z)} \right]. \qquad \text{[A5]}$$

In this case we see that the marginal social damage will include both the direct loss of income to households and the loss of revenues due to changes in consumption. The latter of these two effects will be zero if E is assumed to be separable in utility, as has been the common assumption in the recent literature.[6]

Simplifying the notation in [A4] using [A5] we have

$$-\lambda X_j + \mu\left[\sum_{j=1}^{n} t_i \frac{\partial X_i}{\partial p_j} + X_j \right] + \pi \frac{\partial X_Z}{\partial p_j} = 0 \qquad \forall\ j \neq 0 \qquad \text{[A6]}$$

---

[6] Sandmo does not explicitly consider the effect of changes in environmental quality on demands for other goods, so the second term in brackets on the right-hand side of [A5] is omitted in his analysis. However, given the highly stylized representation of an environmental externality, one may assume that Sandmo has either assumed the effects to be zero, or incorporated them as indirect components of the cross-price effects with respect to the polluting good.

Derivations of optimal tax rules often include substitution of the Slutsky equation in such a way that the social marginal utility of income, α, is represented along with the shadow cost of raising an additional dollar of revenue (Auerbach, *Handbook of Public Economics*, 1985). Diverging from the approach taken by Sandmo, we rearrange the planner's first-order conditions and use the Slutsky equation to split the cross-price effects into compensated effects (superscript *U*) and effects on income, *Y*, as $\frac{\partial X_Z}{\partial p_i} = \frac{\partial X_Z^U}{\partial p_i} - X_i \frac{\partial X_Z}{\partial Y}$. We substitute α to obtain

$$-\alpha X_j + \mu \sum_i t_i \frac{\partial X_i^U}{\partial p_j} + \mu X_j + \pi \left(\frac{\partial X_Z^U}{\partial p_j} - X_j \frac{\partial X_Z}{\partial Y} \right) = 0 \qquad \forall\ j \neq 0\,. \qquad \text{[A7]}$$

We define **Š** as the determinant of the Slutsky matrix of compensated demands, so that $\mathbf{S_{ij}}$ is the cofactor of the element for the *jth* row (price) and *ith* column (quantity). Using Cramer's rule we can solve for the optimal taxes

$$t_j = \frac{(\mu - \alpha)}{\mu} \frac{\sum_{i=1}^{n} X_i S_{ij}}{\check{S}} + \frac{\pi}{\mu} \frac{\sum_{i=1}^{n} \left(\frac{\partial X_Z^U}{\partial p_i} - X_i \frac{\partial X_Z}{\partial Y} \right) S_{ij}}{\check{S}} \qquad \text{[A8]}$$

where the second term on the right-hand side is the environmental component of the tax. From theorems about the expansion of determinants, we know that

$$\sum_{i=1}^{n} \frac{\partial X_Z^U}{\partial p_i} S_{ij} = \begin{cases} 0 & \text{for } \ j \neq Z \\ \check{S} & \text{for } \ j = Z \end{cases}.$$

Let R denote the "Ramsey term" for compensated demands or $R \equiv \sum_{i=1}^{n} X_i S_{ij} \Big/ p_j \check{S}$ reflecting the revenue generating potential for a marginal change in the tax on X_i due to the direct and indirect effects on consumption for all goods. Further simplify the notation by defining the income effect on the environment as $\theta = \pi \sum_{i=1}^{n} X_i \frac{\partial X_Z}{\partial Y}$. We can thus rearrange terms and simplify so that the optimal tax expressions can then be written as

$$\frac{t_j}{(1+t_j)} = \frac{(\mu - \alpha + \theta)}{\mu} R \qquad \forall j \neq z \qquad \text{[A9]}$$

and

$$\frac{t_j}{(1+t_j)} = \frac{(\mu - \alpha + \theta)}{\mu} R + \frac{\pi}{\mu(1+t_j)} \qquad \forall j = z \qquad \text{[A10]}$$

These implicit solutions are difficult to interpret by inspection, in part because of the lack of transparency in interpreting R. Moreover, although the environmental component of the tax in [A10] appears to be separable from the standard formula, the independence is illusory both because of the denominator (1 + t_z) is endogenous and because the actual level of the externality depends on the actual equilibrium and hence the optimal tax rates; the same is true in the other direction (Auerbach 1985).

The results differ from the expressions obtained by Sandmo involving uncompensated demands. Sandmo concluded that the environmental damages of X_Z "does not enter the tax formulas for the other commodities, regardless of the pattern of complementarity and substitutability" (1975, p. 92). In this alternative derivation, we see that the numerator in the first term on the right-hand side includes θ, a term involving π, indicating that the presence of an externality raises the optimal tax on all goods due to their income effect: by reducing real income, all taxes discourage consumption of the externality-producing good to some extent, and these optimal tax rates will be higher as a result. These two versions of the optimal tax results are not in conflict: in the model involving ordinary demands, the income effects are implicit.

Substituting the identity MSD = π/α and considering a model with one polluting good, D, and one non-polluting good, C, we can write [A9] and [a10] as

$$\frac{t^*_C}{(1+t^*_C)} = \left(\frac{\mu-\alpha+\theta}{\mu}\right)R \qquad [A11]$$

and

$$\frac{t^*_D}{(1+t^*_D)} = \left(\frac{\mu-\alpha+\theta}{\mu}\right)R + \frac{\alpha}{\mu}\frac{MSD}{(1+t_D)} \qquad [A12]$$

where R is the "Ramsey term" reflecting the revenue generating potential for a marginal change in the tax on any good due to the direct and indirect effects on consumption for all goods.

These expressions, which are similar to those derived by Sandmo, are difficult to interpret by inspection, in part because of the lack of transparency in R. Nor can the two components of the optimal tax rule for D be evaluated separately by inspection, since the denominator is a function of both terms.[7] The problem can be simplified by assuming that C and D are similar from a revenue raising perspective. This has also been done in the recent literature by restricting preferences so that utility is homothetic in consumption, and weakly separable in leisure, environmental quality, and government consumption (Bovenberg and de Mooij 1994; Bovenberg and Goulder 1996). Here, however, we assume only that the Ramsey terms for both goods are equal, which allows us to derive an expression for the environmental component of the optimal tax.

For the polluting good, D, we rearrange [A12] to get

$$t^*_D = \frac{\left(1-\frac{\alpha+\theta}{\mu}\right)R}{\left(1-\left(1-\frac{\alpha+\theta}{\mu}\right)R\right)} + \frac{\alpha MSD}{\mu\left(1-\left(1-\frac{\alpha+\theta}{\mu}\right)R\right)} \qquad [A13]$$

From [A11] we can express the Ramsey term as

[7] Fullerton (1997), Schöb (1997), and Bovenberg and de Mooij (1997) have suggested that Sandmo's formula indicates that the optimal pollution tax should be less than marginal social damages, but this interpretation overlooks the endogeneity of the denominator on the left-hand side: one cannot infer by inspection that the differential between the optimal tax on a polluting good versus a similar non-polluting good will simply equal the value of the second term (since the first term equals the tax on non-polluting goods).

$$R = \frac{\frac{t^*_C}{(1+t^*_C)}}{\left(1-\frac{\alpha+\theta}{\mu}\right)} \qquad \text{[A14]}$$

To evaluate the optimal tax t^*_D relative to MSD, we substitute [A14] into the second term of [A13] and rearrange to get

$$t^*_D = \frac{\left(1-\frac{\alpha+\theta}{\mu}\right)R}{\left(1-\left(1-\frac{\alpha+\theta}{\mu}\right)R\right)} + \frac{\alpha(1+t_C)MSD}{\mu}. \qquad \text{[A15]}$$

We can evaluate the relationship between the optimal environmental tax and MSD in one of two ways; either evaluating the difference between the optimal tax expressions for C and D or by differentiating t_D with respect to MSD. Taking the differentiation approach, we obtain

$$\frac{\partial t^*_D}{\partial(MSD)} = \frac{(1+t_C)\alpha}{\mu}. \qquad \text{[A16]}$$

Comparing this expression with (9) in the text, and recognizing the identity $(1+t_C)=1/(1-t_L)$ for the two different normalizations, we see that these are equivalent. The expression in [A16] indicates that the optimal environmental tax may rise by more or less than MSD depending on the relationship between the revenue-motivated tax rate t^*_C and α/μ. Based on parameters equivalent to those used in the text (where t_C=0.667), the right-hand side of [A16] equals 1.38.