Pigouvian Taxation in a Ramsey World

Robin Boadway[a*] and Jean-François Tremblay[b]

[a]*Queen's University, Canada*

[b]*University of Ottawa, Canada*

Abstract

This paper studies the optimal Pigouvian tax for correcting pollution when the government also uses distortionary taxes to raise revenues. When preferences are quasilinear in leisure and additive, the Pigouvian tax can be separated from the Ramsey revenue-raising tax. We characterize the relationship between the Pigouvian tax and marginal social damages in a variety of circumstances. In a setting with homogeneous households, the Pigouvian tax exceeds marginal damages if goods have inelastic demands, and vice versa. When households are heterogeneous so taxes can be redistributive, the Pigouvian tax gives more weight to damages suffered by low-income persons. The analysis is extended to allow for costly abatement. In general, corrective taxes have to be applied to both emissions and output of the polluting good.

Keywords: Pigouvian Tax, optimal taxes, pollution tax

JEL Classification: H21, H23

1. Introduction

This paper explores the issue of the optimal tax for correcting externalities from pollution when the government is raising revenues using distortionary taxes. In particular, will the Pigouvian component of the tax on a polluting good equal the marginal damages to households (as it does in a first-best world), and how will the Pigouvian tax change as revenue requirements rise? The general analysis of optimal taxes in the presence of externalities has been available in the literature at least since Sandmo's (1975) seminal paper. Cremer et al (1998) and Cremer and Gahvari (2001) have refined the analysis to take account of nonlinear income taxation and the existence of an abatement technology. However, their results are somewhat formal and the precise applicability of them to environmental taxation remains unclear. Partly this is because optimal revenue-raising taxes and Pigouvian taxes interact in a complicated way in optimal tax characterizations since closed-form solutions are generally not possible: formulas for optimal tax rates are generally only in implicit terms. This lack of clarity

* Corresponding author: Robin Boadway, Department of Economics, Queen's University, Kingston, ON K7L 3N6, Canada. Email: boadwayr@econ.queensu.ca Fax number: +1 613-533-6668.

has been stressed in the recent literature on the double dividend from environmental taxes, which has cast doubt on whether there is, in fact, a full double dividend: one dividend from correcting externalities, the other from the revenue raised. Bovenberg and de Mooij (1994), for example, suggest that environmental taxes might exacerbate the tax distortions that already exist. This would suggest that the environmental (Pigouvian) tax should be less than marginal damages to households. Others, such as Jaeger (2002), have argued the opposite.

Part of the problem, as noted by Cremer et al (2001), lies with identifying precisely what is the Pigouvian component of the tax on the polluting good. In an optimal tax world, the tax on a polluting good will comprise both a Ramsey (revenue-raising) effect and a Pigouvian (corrective) effect, and disentangling them will generally not be possible (since the imposition of a corrective tax will affect optimal revenue-raising taxes, and vice versa). Moreover, as Auerbach and Hines (2002) have stressed in their recent survey of optimal taxation, the interpretation of Pigouvian commodity taxes becomes confounded depending on whether commodity taxes or income taxes are used for revenue raising. In particular, this affects the marginal utility of income for the consumer, and therefore the value of the numeraire for measuring marginal damages.

To address these issues, we adopt a simple formulation for the household utility function that allows for a natural separation between Ramsey and Pigouvian taxes, similar in spirit to that used by Cremer et al (2001). They characterize the optimal tax on a polluting good in a Ramsey setting with many non-polluting goods and show that, with separable preferences, all nonpolluting goods are taxed at the same rate. As a result, they define the corrective tax as the tax differential between polluting and nonpolluting goods. Cremer et al (2001) also examine the implications for the optimal tax on a polluting good of different preference separability assumptions, when agents are heterogeneous and an optimal nonlinear income tax system, is used.

In order to derive explicit solutions to optimal tax rates, we assume the utility function to be quasilinear in leisure, so that the demand for goods depends only on own prices relative to the wage rate, and not on either income or other goods' prices. We explore the size of the Pigouvian tax relative to the size of marginal damages to households in a variety of settings. These include the basic case where only commodity taxes are used, the case where a wage income tax is used, the case where there is an abatement technology, and the case where households have different wage rates and an optimal income tax, either linear or nonlinear is used.

2. The Basic Ramsey Optimal Commodity Tax Model

The model we use for our benchmark analysis is the simplest and most transparent one for our purposes. It consists of a population of identical households with quasilinear preferences in labor/leisure from whom the government must extract a given amount of tax revenues at the least cost using distorting commodity taxes. It differs from the standard optimal commodity tax model in that the consumption of one good emits some harmful environmental externality. The choice of quasilinear preferences implies that goods' demands depend only on own prices and not on the prices of other goods or

income, effectively leading to a quasi-partial equilibrium setting. This formulation is chosen partly because the disaggregation of taxes into Ramsey (revenue-raising) and Pigouvian (corrective) components can be made as clearly as possible, and also because closed-form solutions for optimal taxes can be derived.

More formally, the economy consists of N households, each of whose utility function takes the quasilinear form $U_C(C) + U_D(D) + E - L$ where $U_C(C)$ and $U_D(D)$ are increasing and strictly concave functions. Using mnemonic notation, good C is a clean good, good D is a dirty good, E is the quality of the environment and L is labor supplied.[1] One can think of labor supply as being $L = T - H$, where T is time available and H is leisure. The quality of the environment is given by $E = \overline{E} - \delta ND$, where $\overline{E}$ is its quality in the absence of pollution and δ is the marginal damage to the environment per unit of consumption of the dirty good, with δ assumed for simplicity to be constant. Thus, each person's consumption of D affects the quality of the environment enjoyed by all N households. We suppose that N is large enough that each person treats the quality of the environment E as given, and thus independent of their own consumption of D.

Production is linear in this economy and we normalize the producer prices of the two goods to unity. The wage rate is taken to be w, although it too could be normalized as desired. It is useful for expositional purposes not to set it to unity. The government requires an amount of resources valued at R, so the economy's aggregate production constraint is:

$$NC + ND + R = NwL \tag{1}$$

Following the Ramsey optimal tax approach (e.g., Atkinson and Stiglitz, 1980), the government cannot levy lump-sum taxes on consumers so must impose taxes on transactions. In this economy with three goods, C, D and leisure or labor, the government could use three taxes. However, it is well-known that one tax rate is always redundant (e.g., a tax on labor is equivalent to proportional taxes on goods). In our basic model, we assume that the government dispenses with taxes on labor and uses only taxes on the two goods, denoted t_C and t_D. Consumer prices are then given by $1 + t_C$ and $1 + t_D$. The tax on D takes account of the externality associated with its consumption.

Given that households have no initial endowments, their individual budget constraint can be written:

$$(1 + t_C)C + (1 + t_D)D = wL \tag{2}$$

Combining the aggregate of (2) for all N households with (1), we immediately obtain the government revenue constraint:

$$Nt_C C + Nt_D D = R \tag{3}$$

Thus, only two of the economy's three constraints, (1), (2) and (3), need to be taken account of explicitly, and the other is implied.

---

[1] More generally, our analysis applies when there are many clean goods, as in Cremer et al (2001). Since little insight is gained by having more than one clean good, we aggregate them all into one composite good C.

To solve the government's optimal tax problem, consider first household behavior. Households maximize utility, taking E as given, subject to their budget constraint (2). The Lagrangian is:

$$\Omega(C,D,L,\alpha) = U_C(C) + U_D(D) + E - L - \alpha\left((1+t_C)C + (1+t_D)D - wL\right)$$

The first-order conditions on C, D and L are:

$$U_C'(C) - \alpha(1+t_C) = 0,\ U_D'(D) - \alpha(1+t_D) = 0,\ -1 + \alpha w = 0$$

From these, we obtain the demands for C and D and the marginal utility of income, α:

$$C\left(\frac{1+t_C}{w}\right),\ D\left(\frac{1+t_D}{w}\right),\ \alpha = \frac{1}{w}$$

where $C'(\cdot) < 0$ and $D'(\cdot) < 0$. The fact that α is constant is a consequence of the quasilinear utility function, and is useful for purposes of interpretation. Notice that because of the additive utility function, not only do cross-price and income effects disappear, but also goods demands are independent of environmental quality, E.

The indirect utility function can be written:

$$V\left(\frac{1+t_C}{w}, \frac{1+t_D}{w}\right) + E$$

where, by the envelope theorem, we have:

$$\frac{\partial V(\cdot)}{\partial t_C} \equiv V_{t_C}(\cdot) = -\alpha C(\cdot) = -\frac{C(\cdot)}{w} \text{ and } \frac{\partial V(\cdot)}{\partial t_D} \equiv V_{t_D}(\cdot) = -\alpha D(\cdot) = -\frac{D(\cdot)}{w} \tag{4}$$

Given the behavior of the households, the government's problem is to choose tax rates to maximize the sum of indirect utilities subject to its revenue constraint (3) and treating the quality of the environment E as endogenous. The government's Lagrangian is:

$$\Lambda(t_C, t_D, \lambda) = N\left(V(\cdot) + \overline{E} - \delta N D(\cdot)\right) + \lambda\left(N t_C C(\cdot) + N t_D D(\cdot) - R\right)$$

The first-order conditions on t_C and t_D, using the envelope conditions (4), are:

$$-\frac{C}{w} + \lambda\left(C + t_C \frac{C'}{w}\right) = 0 \tag{5}$$

$$-\frac{D}{w} - \frac{N\delta D'}{w} + \lambda\left(D + t_D \frac{D'}{w}\right) = 0 \tag{6}$$

Given that demands for C and D depend only on own relative prices, which depend on own tax rates, (5) and (6) implicitly determine tax rates as a function of the shadow

price of government revenue, $t_C(\lambda)$ and $t_D(\lambda)$. We shall exploit these relationships below.

First, note that (5) and (6) can be rearranged as follows, using $\alpha = 1/w$:

$$t_C = \frac{\alpha - \lambda}{\lambda}\frac{Cw}{C'} = \frac{\alpha - \lambda}{\lambda}\frac{1+t_C}{\eta_C} \tag{7}$$

$$t_D = \frac{\alpha - \lambda}{\lambda}\frac{Dw}{D'} + \frac{N\delta}{\lambda} = \frac{\alpha - \lambda}{\lambda}\frac{1+t_D}{\eta_D} + \frac{N\delta}{\lambda} \tag{8}$$

where, η_C and η_D are elasticities of demand:

$$\eta_C \equiv \frac{C'}{C}\frac{1+t_C}{w} < 0, \quad \eta_D \equiv \frac{D'}{D}\frac{1+t_D}{w} < 0 \tag{9}$$

In the absence of environmental externalities, $\delta = 0$, (7) and (8) are the familiar inverse elasticity rules for optimal taxation:

$$\frac{t_C}{1+t_C} = \frac{\alpha - \lambda}{\lambda}\frac{1}{\eta_C}, \quad \frac{t_D}{1+t_D} = \frac{\alpha - \lambda}{\lambda}\frac{1}{\eta_D}$$

where the tax rates are expressed as *ad valorem* taxes based on consumer prices.

Note that $\lambda > \alpha$ when the government is using distortionary taxes. If the government had access to non-distortionary taxes so $\lambda = \alpha = 1/w$, the tax on good D by (8) would be the first-best Pigouvian tax, $t_D = N\delta/\alpha$. That is, t_D would be set equal to marginal damages measured in terms of household income.

When taxes are distortionary, so $\lambda > \alpha$, (8) corresponds to the optimal tax expression obtained by Cremer et al (1998), albeit here in a much simpler context with homogeneous households. (Their analysis involves heterogeneous households and nonlinear income taxation as well as commodity taxes.) The optimal tax on D consists of two separate components, one seeming to involve a 'Ramsey' component and the other a 'Pigouvian' component. However, the decomposition is incomplete and ambiguous for a couple of reasons. For one, the total tax rate t_D, which includes both the Ramsey and the Pigouvian component, also appears on the right-hand side of (8), including as an argument of η_D. As well, the second term involving the pollution externality is not exactly the marginal damages suffered by consumers since it is discounted by the shadow price of funds to the government, λ. In order to explore this more carefully, we put slightly more structure on the problem in the following section by assuming elasticities of demand are constant.

3. Pigouvian Taxes in the Basic Model: A Special Case

To facilitate the interpretation of Pigouvian taxes in the above Ramsey optimal commodity tax model, let us make the following two special assumptions. Assume that the elasticities of demand for the two goods are constant, and assume that they are equal to one another. Thus, $\eta = \eta_C = \eta_D$, where η is a constant. These assumptions are

pedagogically useful since they allow us to solve for the properties of corrective taxes explicitly. Moreover, in this context, the meaning of the Ramsey component of the tax structure is apparent.[2] If the elasticities of demand are equal, the Ramsey component of the commodity tax structure would be uniform. We can therefore interpret the difference in commodity tax rates as the Pigouvian component of the tax on the dirty good, denoted $t_P = t_D - t_C$.

Given these assumptions, (8) becomes:

$$t_C + t_P = \frac{\alpha - \lambda}{\lambda}\frac{1 + t_C + t_P}{\eta} + \frac{N\delta}{\lambda}$$

Using (7) with $\eta_C = \eta$, this can be solved for t_P :

$$t_P = \frac{\eta N\delta}{(\eta + 1)\lambda - \alpha} \tag{10}$$

Thus, our assumptions have allowed us to obtain an explicit solution for t_P which we can use to obtain two relevant properties of the Pigouvian tax.

First, note that if $\eta = -1$, (10) reduces to $t_P = N\delta/\alpha$. That is, the Pigouvian tax is equal to marginal social damages measured in terms of household utility. Furthermore, differentiating (10) with respect to η, we obtain:

$$\frac{\partial t_P}{\partial \eta} = \frac{(\lambda - \alpha)N\delta}{\left((\eta + 1)\lambda - \alpha\right)^2} > 0 \tag{11}$$

This implies that t_P increases as $|\eta|$ decreases, implying that $t_P > N\delta/\alpha$ for $-\eta < 1$ and vice versa. We summarize this in the following proposition.

Proposition 1: *In the basic model with elasticities of demand constant for both goods, the Pigouvian component of the tax on good D will be greater (less) than marginal social damages to the households if goods are inelastic (elastic) in demand.*

The intuition for this seems to be as follows. Imposing a Pigouvian tax on good *D* requires deviating from equal taxes on both goods, which involves an efficiency cost. The optimal tax must trade off this efficiency cost against the benefits of decreasing environmental damages. Since deviating from equal taxes on both goods involves a greater efficiency cost when demands are more elastic, the Pigouvian tax will be smaller in this case.

Next, recall that in this case with quasilinear preferences, the first order conditions (5) and (6), or equivalently (7) and (8), implicitly determine the tax rates t_C and t_D as a function of λ respectively, that is, the size of revenue requirements. In the special case where the elasticities of demand are constant and equal, we can investigate how the

[2] Our analysis also applies in the case where η_C and η_D are constant but different in size, but the analysis is simpler when they are equal. With different elasticities the Ramsey taxes would be proportional to the ratio of elasticities rather than being identical to one another.

Pigouvian component of the tax on D, t_P, varies with revenue requirements. To do so, differentiate (10) with respect to λ to obtain:

$$\frac{\partial t_P}{\partial \lambda} = -\frac{N\delta\eta(\eta+1)}{((\eta+1)\lambda-\alpha)^2} >,=,< 0 \text{ as } 1>,=,<-\eta \tag{12}$$

Thus, t_P will increase with revenue requirements if the demand for goods is inelastic, and decrease if demand is elastic.

This leads to the following proposition:

Proposition 2: *In the basic model with elasticities of demand constant for both goods, the Pigouvian component of the tax on good D will diverge more from marginal social damages (positively in the case of inelastic demands, negatively in the case of elastic demands) as revenue requirements increase.*

The results of this section apply when the government uses only commodity taxes. In the special case studied here where demand elasticities for both goods are constant and equal, the government could apply a proportional tax on the consumer purchases of both goods, such as a uniform VAT, and accompany it with a Pigouvian tax on the consumption of good D. The Pigouvian tax would then be greater or less than marginal damages to the consumer according to whether the demands are inelastic or elastic. Next, we investigate the case where the uniform tax is applied to income (wages) rather than consumption.

4. Pigouvian Taxes Combined with a Wage Tax

We continue to assume that preferences are quasilinear and that demand elasticities are equal and constant. However, instead of imposing taxes on the two goods, we use a tax on labor income along with a tax on good D. Let t_w be the tax rate on wage income and t_P be the tax on the dirty good, which in this case we can think of as the Pigouvian tax. Then, the household budget constraint is $C+(1+t_P)D=(1-t_w)wL$, and the Lagrangian function for the household problem becomes:

$$\Omega(C,D,L,\alpha)=U_C(C)+U_D(D)+E-L-\alpha(C+(1+t_P)D-(1-t_w)wL)$$

The first-order conditions on C, D and L are now:

$$U_C'(C)-\alpha=0,\ U_D'(D)-\alpha(1+t_P)=0,\ -1+\alpha(1-t_w)w=0$$

which yield the demands for C and D (and implicitly the supply of labor) and the marginal utility of income α:

$$C\left(\frac{1}{(1-t_w)w}\right),\ D\left(\frac{1+t_P}{(1-t_w)w}\right),\ \alpha=\frac{1}{(1-t_w)w}$$

The indirect utility function is:

$$V\left(\frac{1}{(1-t_w)w}, \frac{1+t_P}{(1-t_w)w}\right) + E$$

and the envelope theorem gives:

$$V_{t_p}(\cdot) = -\alpha D(\cdot) = -\frac{D(\cdot)}{(1-t_w)w} \text{ and } V_{t_w}(\cdot) = -\alpha w L(\cdot) = -\frac{L(\cdot)}{1-t_w}$$

Turning to the government's problem, we first obtain an expression for the government's budget constraint. To do so, we can use the resource constraint for the economy, which as above is $NC + ND + R = NwL$. Combining this with the aggregate budget constraint of the consumers, we obtain $Nt_P D + Nwt_w L = R$. From the household budget constraint, labor income wL is given by:

$$wL = \frac{C}{1-t_w} + \frac{1+t_P}{1-t_w} D$$

Inserting this into the government budget constraint, we obtain:

$$N\frac{t_w}{1-t_w} C(\cdot) + N\frac{t_w + t_P}{1-t_w} D(\cdot) = R \tag{13}$$

Given the government's budget constraint (13), the Lagrangian function for the government's optimal tax problem can be written:

$$\Lambda(t_P, t_w, \lambda) = N\left(V(\cdot) + \overline{E} - \delta N D(\cdot)\right) + \lambda\left(N\frac{t_w}{1-t_w} C(\cdot) + N\frac{t_P + t_w}{1-t_w} D(\cdot) - R\right)$$

The first-order conditions on t_w and t_p are:

$$-N\alpha wL - \frac{N^2\delta(1+t_P)D'}{(1-t_w)^2 w} + \lambda\left(N\frac{C}{(1-t_w)^2} + N\frac{t_w}{1-t_w}\frac{C'}{(1-t_w)^2 w} + N\frac{1+t_P}{(1-t_w)^2} D\right.$$

$$\left. + N\frac{t_P + t_w}{1-t_w}\frac{1+t_P}{(1-t_w)^2 w} D'\right) = 0$$

$$-N\alpha D - \frac{N^2\delta D'}{(1-t_w)w} + \lambda\left(N\frac{D}{1-t_w} + N\frac{t_P + t_w}{1-t_w}\frac{D'}{(1-t_w)w}\right) = 0$$

Using the second equation to eliminate terms from the first equation, the value of L from the household budget, these two equations simplify to:

$$-\frac{C}{w} + \lambda\left(C + \frac{t_w}{1-t_w}\frac{C'}{w}\right) = 0, \quad -\frac{D}{w} - \frac{N\delta D'}{w} + \lambda\left(D + \frac{t_P + t_w}{1-t_w}\frac{D'}{w}\right) = 0 \tag{14}$$

Analogously to the basic model, we can use (14) to get expressions for the Pigouvian tax t_p. First note that, using the relative prices in this case, the common elasticity of demand η can be written as follows:

$$\eta = \frac{C'}{C}\frac{1}{(1-t_w)w} = \frac{D'}{D}\frac{1+t_P}{(1-t_w)w}$$

Using these expressions and the fact that $1/w = (1-t_w)\alpha$, the equations in (14) can be rewritten:

$$-(1-t_w)\alpha + \lambda(1+\eta t_w) = 0, -(1-t_w)\alpha - \frac{N\delta\eta(1-t_w)}{1+t_P} + \lambda\left(1+\frac{t_P+t_w}{1+t_P}\eta\right) = 0 \qquad (15)$$

Combining these two expressions and solving for t_p, we obtain:

$$t_P = \frac{\eta N\delta(1-t_w)}{(\eta+1)\lambda - (1-t_w)\alpha} \qquad (16)$$

which is the analog of (10) in this case.

Proceeding as before, we see that when $\eta = -1$, $t_P = N\delta/\alpha$ so the Pigouvian tax equals marginal social damages to the household. Moreover, differentiating (16) by η, we find that:

$$\frac{\partial t_P}{\partial_\eta} = \frac{(\lambda - (1-t_w)\alpha)N\delta(1-t_w)}{((\eta+1)\lambda - (1-t_w)\alpha)^2} > 0$$

This is the analog of (11), so Proposition 1 applies. Similarly, differentiating (16) with respect to λ, we obtain:

$$\frac{\partial t_P}{\partial\lambda} = -\frac{N\delta\eta(1-t_w)(\eta+1)}{((\eta+1)\lambda - (1-t_w)\alpha)^2} >, =, < 0 \text{ as } 1 >, =, < -\eta$$

This is equivalent to (12) implying that Proposition 2 also applies. Not surprisingly, our results are independent of whether the government uses a set of differential commodity taxes or a wage tax combined with a tax on good D.[3]

5. Pigouvian Taxes with Pollution Abatement

So far we have assumed that the externality is proportional to the output of the dirty good. Let us now suppose that pollution emissions arising from the use of good D can be reduced by an abatement technology, as in Cremer and Gahvari (2001). Pollution abatement is modeled in the simplest way consistent with making the point. Let A be

[3] The analysis would be slightly more complicated if the elasticities of demand were different for the two goods since then the Ramsey taxes would not be equal to one another.

total abatement such that $ND - A$ are total emissions of pollution by the dirty good industry. The quality of the environment then becomes $E = \overline{E} - \delta(ND - A)$, following the above notation. The total cost of abatement borne by producers is given by $Z(A)$, where $Z'(A), Z''(A) > 0$. We revert to the commodity tax model of section 3, but in addition to commodity taxes t_c and t_D imposed on the sales of goods C and D, there is also a per unit tax t_E imposed on emissions by producers.

Given the level of abatement A, the tax t_E applies on incremental production of output D. Assuming $A < ND$, the producer price of D is therefore $1 + t_E$. Producer costs include inputs into the production of D, which given unit costs are simply ND, as well as taxes on emissions and the cost of abatement. Given the producer price, $1 + t_E$, profits of the producers in industry D are therefore:

$$\Pi \equiv (1 + t_E)ND - ND - t_E(ND - A) - Z(A) = t_E A - Z(A) \tag{17}$$

Producers will choose A to maximize their profits, which leads to the first-order condition $t_E = Z'(A)$, whose solution is $A(t_E)$ with $A'(t_E) > 0$. We assume for simplicity that the government taxes these profits fully.

Consumer prices for C and D are $1 + t_C$ and $1 + t_D + t_E$, and as before the wage rate is fixed at w. The consumer budget constraint is $(1 + t_C)C + (1 + t_D + t_E)D = wL$. Consumers maximize utility, given E, subject to their budget constraint. The Lagrangian is:

$$\Omega(C, D, L, \alpha) = U_C(C) + U_D(D) + E - L - \alpha\left((1 + t_C)C + (1 + t_D + t_E)D - wL\right)$$

Proceeding as earlier, the solution gives the demand functions and indirect utility function:

$$C\left(\frac{1 + t_C}{w}\right), \quad D\left(\frac{1 + t_D + t_E}{w}\right) \text{ and } V\left(\frac{1 + t_C}{w}, \frac{1 + t_D + t_E}{w}\right) + E$$

with $\alpha = 1/w$. Applying the envelope theorem yields:

$$V_{t_C} = -\alpha C = -\frac{C}{w} \text{ and } V_{t_D} = V_{t_E} = -\alpha D = -\frac{D}{w} \tag{18}$$

The resource constraint for the economy is $NC + ND + Z(A) + R = NwL + \Pi$, where Π satisfies (17). Together with the aggregate household budget constraint, this leads to the government revenue constraint $Nt_C C + Nt_D D + t_E(ND - A) + \Pi = R$. Given this, the Lagrangian expression for the problem of the government, using (17), is:

$$\Lambda = NV(\cdot) + N\left(\overline{E} - \delta ND(\cdot) + \delta A(t_E)\right) + \lambda\left(Nt_C C(\cdot) + N(t_D + t_E)D(\cdot) - Z(A(t_E)) - R\right)$$

The first-order conditions for t_C, t_D and t_E, using $t_E = Z'(A)$ from the abatement decision of producers, are:

$$-\frac{C}{w} + \lambda\left(C + \frac{t_C C'}{w}\right) = 0 \tag{19}$$

$$-\frac{D}{w}-\frac{\delta N D'}{w}+\lambda\left(D+\frac{\left(t_D+t_E\right)D'}{w}\right)=0 \tag{20}$$

$$-N\frac{D}{w}-\frac{\delta N^2 D'}{w}+N\delta A'\left(t_E\right)+\lambda\left(ND+\frac{N\left(t_D+t_E\right)D'}{w}-t_E A'\left(t_E\right)\right)=0 \tag{21}$$

Assume that the elasticities of demand for C and D are identical and constant, so

$$\eta=\frac{C'}{C}\frac{1+t_C}{w}=\frac{D'}{D}\frac{1+t_D+t_E}{w}$$

Then, using $\alpha=1/w$, we can rewrite conditions (19) and (20), respectively, as:

$$t_C=\frac{\alpha-\lambda}{\lambda}\frac{1+t_C}{\eta} \tag{22}$$

$$t_D+t_E=\frac{\alpha-\lambda}{\lambda}\frac{1+t_D+t_E}{\eta}+\frac{N\delta}{\lambda} \tag{23}$$

which are analogous to (7) and (8) in our basic case. In this case, the total tax on good D is t_D+t_E, and following the earlier logic, the Pigouvian tax can be defined as $t_P\equiv t_D+t_E-t_C$. By (23), we have:

$$t_C+t_P=\frac{\alpha-\lambda}{\lambda}\frac{1+t_C+t_P}{\eta}+\frac{N\delta}{\lambda}$$

which reduces using (22) to:

$$t_P=\frac{\eta N\delta}{(\eta+1)\lambda-\alpha} \tag{24}$$

This Pigouvian component of the total tax on good D is the same as in the basic case without abatement, given by (10). The same analysis as before leads to the analogs of Propositions 1 and 2 in this case, given the definition of the Pigouvian tax used here.

In this case, the Pigouvian component is more complicated than earlier since now it includes the tax on emissions, t_E, as well as the difference in commodity taxes, $t_D - t_C$. We can use the first-order condition (21) on t_E to give further insight. Substituting (20) into (21), we obtain immediately:

$$t_E=\frac{N\delta}{\lambda} \tag{25}$$

This component of the Pigouvian tax is lower than marginal damages $N\delta/\alpha$, since $\lambda>\alpha$. The other component t_D-t_C is given by combining (24) and (25) to yield:

$$t_D-t_C=\frac{N\delta}{\lambda}\left(\frac{\alpha-\lambda}{\lambda\eta-(\alpha-\lambda)}\right)$$

Thus, $t_D > t_C$ if $\lambda\eta - (\alpha - \lambda) < 0$ (since the numerator is negative). In fact, this condition is satisfied as can be seen from (22), which can be written:

$$\frac{t_C}{1+t_C} = \frac{\lambda - \alpha}{-\lambda\eta}$$

Therefore, assuming $t_C > 0$, this implies that $\lambda - \alpha < -\lambda\eta$, which in turn implies that $t_D > t_C$. This means that the Pigouvian tax incorporates not just a tax on emissions, but also a differential tax on good *D*.

Note further that if the government could levy a non-distorting lump-sum tax, then $\alpha = \lambda$, so there would be no differential tax on D $(t_C = t_D)$, and the tax on emissions would equal marginal damages to the consumers $t_E = N\delta/\alpha = N\delta/\lambda$. We can summarize these results in the following proposition.

Proposition 3: *When producers can reduce emissions of pollution by costly abatement,*

1. *the Pigouvian component of the second-best commodity tax system includes both a tax on emissions and a differential tax on sales of the dirty good,*
2. *the second-best tax on emissions is less than its first-best level, while the differential tax on D exceeds its first-best level, which is zero, and*
3. *the composite Pigouvian tax satisfies Propositions 1 and 2.*

A final apparent implication of this analysis is that if the government uses a wage tax rather than commodity taxes on t_D and t_C, it can only achieve the second-best optimum by imposing a two-component Pigouvian corrective tax. One component is the tax on emissions, while the other is a tax on sales of good *D*.

6. Pigouvian Taxes and Linear Progressive Taxation

Suppose now that households are heterogeneous. Following the optimal income tax literature, we assume that they differ only in their wage rates, but otherwise the model used above applies. It suffices to restrict our attention to two wage-types, w_1 and w_2 with $w_2 > w_1$, where N_1 and N_2 are the numbers of the two types. The government is assumed to be able to observe income, $w_i L_i$, $i = 1,2$, but neither the wage rate, w_i, nor labor supply, L_i. We begin with the case where the government uses a linear progressive income tax as well as a Pigouvian tax, t_P, on the polluting good, *D*. The income tax consists of a constant marginal tax rate, t_w, combined with an equal per capita subsidy, *s*. The budget constraint for a household with wage rate w_i becomes $C_i + (1+t_P)D_i = (1-t_w)w_i L_i + s$. The household maximizes utility subject to this budget constraint, and the solution gives analogous demand functions and marginal utility of income to the wage tax case:

$$C_i\left(\frac{1}{(1-t_w)w_i}\right),\quad D_i\left(\frac{1+t_P}{(1-t_w)w_i}\right),\quad \alpha_i = \frac{1}{(1-t_w)w_i}$$

The indirect utility function for the type-*i* household is now:

$$V^i\left(\frac{1}{(1-t_w)w_i}, \frac{1+t_P}{(1-t_w)w_i}, s\right) + E$$

where E is now $\overline{E} - \delta(N_1 D_1 + N_2 D_2)$. The envelope theorem gives:

$$V^i_{t_P}(\cdot) = -\alpha_i D_i(\cdot) = -\frac{D_i(\cdot)}{(1-t_w)w_i}, V^i_{t_w}(\cdot) = -\alpha_i w_i L_i(\cdot) = -\frac{L_i(\cdot)}{1-t_w}, V^i_s = \alpha_i$$

Using the aggregate household budget constraint and the economy's resource constraint as before, the government revenue constraint can be written:

$$\frac{t_w}{1-t_w}(N_1 C_1 + N_2 C_2) + \frac{t_w + t_P}{1-t_w}(N_1 D_1 + N_2 D_2) = R + \frac{N_1 + N_2}{1-t_w} s$$

The government problem is treated as a Pareto maximizing one, and we can take the objective function to be $\rho_1 N_1 (V^1(\cdot) + E) + \rho_2 N_2 (V^2(\cdot) + E)$. The social weights ρ_1, ρ_2 are arbitrary, although it is useful to suppose that they are chosen such that the government wants to redistribute from the high-wage to the low-wage types. That implies that the marginal social utility of income of the latter exceeds the marginal social utility of income of the former, or $\rho_1 \alpha_1 > \rho_2 \alpha_2$ in the optimum. If the government could use lump-sum taxation, it would equate these marginal social utilities in a first-best optimum. The latter serves as a useful benchmark below.

The government maximizes its objective function subject to its revenue constraint. Using the first-order condition on *s*, the first-order conditions with respect to t_w and t_P reduce to the following after some simplification:

$$-\frac{\rho_1 N_1 C_1}{w_1} - \frac{\rho_2 N_2 C_2}{w_2} + \lambda\left(N_1 C_1 + N_2 C_2 + \frac{t_w}{1-t_w}\left(\frac{N_1 C_1'}{w_1} + \frac{N_2 C_2'}{w_2}\right)\right) = 0 \quad (26)$$

$$-\frac{\rho_1 N_1 D_1}{w_1} - \frac{\rho_2 N_2 D_2}{w_2} - \overline{N}\delta\left(\frac{N_1 D_1'}{w_1} + \frac{N_2 D_2'}{w_2}\right)$$
$$+\lambda\left(N_1 D_1 + N_2 D_2 + \frac{t_w + t_P}{1-t_w}\left(\frac{N_1 D_1'}{w_1} + \frac{N_2 D_2'}{w_2}\right)\right) \equiv 0 \quad (27)$$

where $\overline{N} = \rho_1 N_1 + \rho_2 N_2$. These are the analogs of (14) above in the wage tax case. Let us again assume a constant elasticity of demand, so:

$$\eta \equiv \frac{C_1'}{C_1}\frac{1}{(1-t_w)w_1} = \frac{C_2'}{C_2}\frac{1}{(1-t_w)w_2} = \frac{D_1'}{D_1}\frac{1+t_P}{(1-t_w)w_1} = \frac{D_2'}{D_2}\frac{1+t_P}{(1-t_w)w_2}$$

Using these definitions, (26) and (27) can be written:

$$-(1-t_w)\overline{\alpha}_C + \lambda(1+t_w\eta) = 0, -(1-t_w)\overline{\alpha}_D - \overline{N}\delta\eta\frac{(1-t_w)}{1+t_P} + \lambda\left(1 + \frac{t_P + t_w}{1+t_P}\eta\right) = 0 \quad (28)$$

where:

$$\bar{\alpha}_C = \frac{\rho_1\alpha_1 N_1 C_1 + \rho_2\alpha_2 N_2 C_2}{N_1 C_1 + N_2 C_2}, \quad \bar{\alpha}_D = \frac{\rho_1\alpha_1 N_1 D_1 + \rho_2\alpha_2 N_2 D_2}{N_1 D_1 + N_2 D_2} \tag{29}$$

These are weighted averages of the marginal social utilities of income of the two wage types, weighted by their shares of consumption of the two goods.[4]

Given our assumption about utility functions, preferences are homothetic in C and D and separable from leisure, which implies that $C_1/D_1 = C_2/D_2$, so that $\bar{\alpha}_C = \bar{\alpha}_D \equiv \bar{\alpha}$. Then, the equations in (28) are the same as those in (15) in the homogeneous-consumer case when $\bar{\alpha}$ replaces α and $\bar{N}$ replaces N, so the same derivation yields the analog of (16):

$$t_P = \frac{\eta\bar{N}\delta(1-t_w)}{(\eta+1)\lambda - (1-t_w)\bar{\alpha}} \tag{30}$$

The analogs of Propositions 1 and 2 then apply. In particular, when $\eta = -1$, $t_P = \bar{N}\delta/\bar{\alpha}$, while t_P is increasing in η. As well, $\partial t_P/\partial\lambda >, =, < 0$ as $1 >, =, < -\eta$.

To interpret this, consider the first-best outcome with lump-sum redistributive taxes and a Pigouvian tax. The lump-sum taxes are used to equate marginal social utilities of income to the shadow price of government revenue, while the Pigouvian tax is set equal to the sum of marginal damages to the households:

$$\rho_1\alpha_1 = \rho_2\alpha_2 = \lambda, t_P = \frac{N_1\delta}{\alpha_1} + \frac{N_2\delta}{\alpha_2} \tag{31}$$

With linear progressive taxes, marginal social utilities of income cannot be equated. In these circumstances, we can use $t_P = \bar{N}\delta/\bar{\alpha}$ obtained when $\eta = -1$ as a benchmark and interpret it by rewriting it in the following way, using $\bar{N} = \rho_1 N_1 + \rho_2 N_2$:

$$t_P\big|_{\eta=-1} = \frac{N_1\delta}{\bar{\alpha}/\rho_1} + \frac{N_2\delta}{\bar{\alpha}/\rho_2}$$

Using (29) for $\bar{\alpha}$ and recalling that $\rho_1\alpha_1 > \rho_2\alpha_2$, we have:

$$\frac{\bar{\alpha}}{\rho_1} = \alpha_1 \frac{N_1 C_1 + \frac{\rho_2\alpha_2}{\rho_1\alpha_1} N_2 C_2}{N_1 C_1 + N_2 C_2} < \alpha_1, \quad \frac{\bar{\alpha}}{\rho_2} = \alpha_2 \frac{\frac{\rho_1\alpha_1}{\rho_2\alpha_2} N_1 C_1 + N_2 C_2}{N_1 C_1 + N_2 C_2} > \alpha_2$$

The implication is that, compared with the first best, the Pigouvian tax puts more weight on marginal damages to the low-wage persons than to the high-wage persons, thus taking on some redistibutive role to complement the linear income tax system. This is reminiscent of the result of Sandmo (2006) that in the absence of full international redistributive transfers, low-income countries should have lower pollution taxes than high-income countries. We can summarize these results in the following proposition, which is related to Propositions 1 and 2.

[4] The expressions $\bar{\alpha}_C$ and $\bar{\alpha}_D$ are analogous to the distributive weights for optimal commodity taxes (or public sector prices) defined by Feldstein (1972).

Proposition 4: *Suppose households differ in wage rates but have the same quasilinear-in-leisure preferences, and the government uses a linear progressive income tax.*

1. *When the elasticity of demand for goods is unity,* $t_P = \rho_1 N_1 \delta / \overline{\alpha} + \rho_2 N_2 \delta / \overline{\alpha}$, *which differs from social marginal damages by putting relatively more weight on damages to low-wage persons and less weight on damages to high-wage persons.*
2. *The Pigouvian tax,* t_P, *falls with the absolute value of the elasticity of demand.*
3. *The Pigouvian tax increases with revenue requirements if* $\eta < -1$, *and vice versa.*

7. Pigouvian Taxes and a Nonlinear Income Tax

Suppose now that the government levies a nonlinear tax on income. Let the before-tax income of a household with wage rate w_i be $Y_i = w_i L_i$. Given that the government can observe Y_i, it is useful to transform the utility function into one involving income rather than labor supply as follows:

$$V^i(C,D,Y) + E \equiv U_C(C) + U_D(D) - Y / w_i + E \tag{32}$$

where V^i is type-specific.

While the government can observe income, it cannot observe individual purchases of C and D. It can therefore levy a non-linear income tax on Y, but only impersonal indirect taxes on commodity purchases. From an analytical perspective, it suffices to consider only a tax on D: proportionate taxes on C and D can be subsumed into the income tax function so we can normalize one commodity tax rate to be zero. Let t_D be the per unit tax rate on commodity D, so the consumer price is $q_D = 1 + t_D$. Suppose that $T(Y)$ is a nonlinear tax function applied to pre-tax income. We can then define I as disposable income (total consumption expenditures), where $I = Y - T(Y)$. Although the government cannot observe how I is divided between C and D, it does know the consumer utility function. If it knew a person's type, it could infer a household's labor supply from its income and would therefore know how each household spends its income. However, households of one type can mimic one another's (I,Y) bundle, which precludes the government from inferring L and constrains government policy. All this is well known from optimal income tax theory (e.g., Stiglitz, 1987).

As above, we first characterize household behavior and then turn to the government. A household of type i maximizes $V^i(C,D,Y) + E$ subject to a budget constraint $C + q_D D = I = Y - T(Y)$. It turns out to be useful for the purposes of analysis to disaggregate household behavior into two successive stages. In the first stage, they choose their labor supply, and therefore the combination of Y and I. In the second stage, they allocate I between C and D. The analysis of household behavior is considered in reverse order.

In stage 2, I and Y have been determined and are taken as given. From the household's budget constraint, $C = I - q_D D$. Therefore, the problem of a representative household of a given type can be written (with superscripts suppressed): $\max_{\{D\}} V(I - q_D D, D, Y) + E$.

The first-order condition is $-q_D V_C + V_D = 0$, which yields the demand for D with the following properties (where E is separable so does not affect demand):

$$D(q_D, I, Y): \frac{\partial D}{\partial q_D} < 0, \frac{\partial D}{\partial I} > 0, \frac{\partial D}{\partial Y} = 0$$

The sign of $\partial D / \partial Y$ depends on the substitute/complement relations between D and L. Since preferences are separable, labor supply and thus income do not affect preferences for C and D. So, given I, a change in Y will not affect the demand for D. Therefore, we can write the demand for D as simply $D(q_D, I)$.

The maximum value function for the consumer—the indirect utility function—is denoted $W(q_D, I, Y) + E$. Applying the envelope theorem to this problem, we obtain:

$$W_{q_D} = -DV_C = -DU_C'(C), W_I = V_C = U_C'(C), W_Y = V_Y = -\frac{1}{w} \tag{33}$$

For a given value of q_D, consumer indifference curves in Y and I can be drawn. They have a slope of $-W_Y / W_I = 1/(wU_C')$ and satisfy the single crossing property: $-W_Y^1 / W_I^1 > -W_Y^2 / W_I^2$.

As mentioned, households of a given type may mimic the consumption-income (I,Y) bundles of the other type. If we assume that the government will redistribute from the high- to the low-wage types starting in the laissez faire, the case of interest for us is where the high-wage types are tempted to mimic the low-wage types. Let $\hat{V}^2(I_1 - q_D D, D, Y_1) + E$ denote the utility of the type 2's when they are mimicking the consumption-income bundle of the type 1's. The problem for the high-wage mimickers in stage 2 is then: $\max_{\{D\}} \hat{V}^2(I_1 - q_D D, D, Y_1) + E$. This yields the demand function $\hat{D}_2(q_D, I, Y)$ which has the same properties as for the non-mimickers. The indirect utility function for the mimickers is denoted $\hat{W}^2(q_D, I_1, Y_1) + E$.

The relationship between D_1 and $\hat{D}_2$ is important in what follows. Both type 1s and mimicking type 2s will have the same consumption expenditures and incomes, while labor supply is lower for the mimicker (since the same income can be earned with lower L). Therefore, since consumption goods are separable from leisure in the utility function, $D_1 = \hat{D}_2$. If D had been more complementary than C with leisure, $D_1 < \hat{D}_2$, and vice versa.

Turn now to stage 1. In this stage, households choose their labor supply, or equivalently, their income. In effect, they choose the most preferred bundle (I,Y) from the budget constraint $I = Y - T(Y)$. Given this behavior, the government chooses its tax policies, both $T(Y)$ and t_D. Following the optimal income tax approach, instead of analyzing the government choice of $T(Y)$, we use the direct approach and let the government offer the bundles (I_1, Y_1), (I_2, Y_2) for the two household types. Then households faced with the bundles offered choose their most preferred.

As before, we characterize a Pareto optimal tax policy, focusing on those Pareto optimal outcomes for which the incentive constraint on the type 2s is binding. (This will be consistent with a social welfare function that exhibits aversion to inequality.) The optimal policy problem for the case where the incentive constraint applies only to the high-ability types is as follows, where $\rho_1 + \rho_2 = 1$ with no loss of generality:

$$\underset{\{I_i,Y_i,t_D\}}{Max}\ \rho_1 N_1\left(W^1(q_D,I_1,Y_1)+E\right)+\rho_2 N_2\left(W^2(q_D,I_2,Y_2)+E\right)$$

subject to:

$$W^2(q_D,I_2,Y_2)+E\geq \hat{W}^2(q_D,I_1,Y_1)+E \qquad (\gamma)$$

$$N_1\left(Y_1-I_1+t_D D_1(q_D,I_1)\right)+N_2\left(Y_2-I_2+t_D D_2(q_D,I_2)\right)=R \qquad (\lambda)$$

where, recall, $q_D=1+t_D$ and $E=\overline{E}-\delta\left(N_1D_1(q_D,I_1)+N_2D_2(q_D,I_2)\right)$, while γ and λ refer to the Lagrange multipliers in the respective constraints.

The first-order conditions on I_1, Y_1, I_2, Y_2 and t_D are:

$$\rho_1 N_1 W_I^1-\gamma\hat{W}_I^2-\overline{N}\delta N_1\frac{\partial D_1}{\partial I_1}-\lambda\left(N_1-N_1t_D\frac{\partial D_1}{\partial I_1}\right)=0 \qquad (34)$$

$$\rho_1 N_1 W_Y^1-\gamma\hat{W}_Y^2+\lambda N_1=0 \qquad (35)$$

$$(\rho_2 N_2+\gamma)W_I^2-\overline{N}\delta N_2\frac{\partial D_2}{\partial I_2}-\lambda\left(N_2-N_2t_D\frac{\partial D_2}{\partial I_2}\right)=0 \qquad (36)$$

$$(\rho_2 N_2+\gamma)W_Y^2+\lambda N_2=0 \qquad (37)$$

$$\rho_1 N_1 W_{q_D}^1+(\rho_2 N_2+\gamma)W_{q_D}^2-\gamma\hat{W}_{q_D}^2-\overline{N}\delta\left(N_1\frac{\partial D_1}{\partial q_D}+N_2\frac{\partial D_2}{\partial q_D}\right)$$

$$+\lambda\left(N_1D_1+N_1t_D\frac{\partial D_1}{\partial q_D}+N_2D_2+N_2t_D\frac{\partial D_2}{\partial q_D}\right)=0 \qquad (38)$$

where, as before, $\overline{N}=\rho_1N_1+\rho_2N_2$. These conditions give the structure of the nonlinear income tax system as well as the optimal tax on D, t_D. Given the assumptions about the form of the utility function, the so-called Atkinson and Stiglitz (1976) Theorem applies. That is, in the absence of environmental externalities, there would be no differential tax on goods. That implies that we can interpret t_D as the Pigouvian tax.

To determine the value of t_D, combine (34), (36) and (38), using $W_{q_D}^i=-D_iW_I^i$ from (33) and $D_1=\hat{D}_2$:

$$\left(\lambda t_D-\overline{N}\delta\right)\left(N_1\left(\frac{\partial D_1}{\partial q_D}+D_1\frac{\partial D_1}{\partial I_1}\right)+N_2\left(\frac{\partial D_2}{\partial q_D}+D_2\frac{\partial D_2}{\partial I_2}\right)\right)=0$$

Since the terms in the large brackets are non-zero, this implies that the Pigouvian tax satisfies:

$$t_D=\frac{\overline{N}\delta}{\lambda}=\frac{(\rho_1N_1+\rho_2N_2)\delta}{\lambda} \qquad (39)$$

In words, the Pigouvian tax is the marginal damages $(\rho_1 N_1 + \rho_2 N_2)\delta$ in terms of the marginal value of government revenue λ. We return to the interpretation of this below.

First, consider the marginal tax rate on the high-wage households. Condition (36) can be written:

$$(\rho_2 N_2 + \gamma) W_I^2 - \lambda N_2 - (\overline{N}\delta - \lambda t_D) N_2 \frac{\partial D_2}{\partial I_2} = 0$$

Using (39), this becomes $(\rho_2 N_2 + \gamma) W_I^2 - \lambda N_2 = 0$. Combining this with (37), we obtain $-W_Y^2 / W_I^2 = 1$, implying that the marginal tax rate at the top is zero. This familiar condition thus continues to apply when pollution taxes are set optimally. Similar reasoning shows that the marginal tax rate on the low-wage types has the standard form: it is positive but less than 100 percent.

Consider now the interpretation of the Pigouvian tax, t_D, given by (39). As a benchmark, it is useful to recall the first-best case where taxes are non-distortionary. In this context, the incentive constraint is not binding, so $\gamma = 0$. Then, from the first-order conditions (35) and (37) and using the envelope conditions (33), we obtain:

$$\lambda = -\rho_1 V_Y^1 = -\rho_2 V_Y^2 = \frac{\rho_1}{w_1} = \frac{\rho_2}{w_2} = \rho_1 \alpha_1 = \rho_2 \alpha_2$$

Thus, the government equates social utility of income for both types. Condition (39) can therefore be written:

$$t_D = \frac{N_1 \delta}{\alpha_1} + \frac{N_2 \delta}{\alpha_2} \tag{40}$$

The optimal Pigouvian tax equals the sum of marginal damages measured in terms of the income of each consumer.

When the income tax is distorting so the incentive constraint binds, the government can no longer equate marginal social utilities of income because $\gamma \neq 0$. Combining (35) and (37) now yields:

$$\rho_1 N_1 W_Y^1 + \rho_2 N_2 W_Y^2 + \gamma (W_Y^2 - \hat{W}_Y^2) = -\lambda (N_1 + N_2)$$

By (33), $W_Y^2 = \hat{W}_Y^2 = -1/w_2$ and $W_Y^1 = -1/w_1$. Therefore, we obtain:

$$\lambda = -\frac{\rho_1 N_1 W_Y^1 + \rho_2 N_2 W_Y^2}{N_1 + N_2} = \frac{\rho_1 N_1 / w_1 + \rho_2 N_2 / w_2}{N_1 + N_2} = \frac{N_1 \rho_1 \alpha_1 + N_2 \rho_2 \alpha_2}{N_1 + N_2}$$

Thus, the shadow value of public funds, λ, is a weighted average of marginal social utilities of income.

To compare the second-best Pigouvian tax with the first best, rewrite (36) as follows:

$$t_D = \frac{N_1 \delta}{\lambda / \rho_1} + \frac{N_2 \delta}{\lambda / \rho_2}$$

Using the above expression for λ and noting that in the second best, marginal social utility of income will be higher for the low-wage types, $\rho_1\alpha_1 > \rho_2\alpha_2$, we obtain:

$$\frac{\lambda}{\rho_1} = \alpha_1 \frac{N_1 + \frac{\rho_2\alpha_2}{\rho_1\alpha_1}N_2}{N_1 + N_2} < \alpha_1, \quad \frac{\lambda}{\rho_2} = \alpha_2 \frac{\frac{\rho_1\alpha_1}{\rho_2\alpha_2}N_1 + N_2}{N_1 + N_2} > \alpha_2$$

The implication is that, compared with the first best, the Pigouvian tax puts more weight on marginal damages to the low-wage persons than to the high-wage persons, thus taking on some redistributive role to complement the nonlinear income tax system. This might be contrasted with the results of the basic model whereby the Pigouvian tax deviates systematically from marginal damages depending on the elasticity of demand.

With more general preferences, the tax on good D would deviate from zero according to the substitute-complement relationship between the demand for D and leisure, as analyzed by Edwards et al (1994) and Nava et al (1996). In this case, the formula for t_D would include not just a Pigouvian component but also a component reflecting the role of the indirect tax system as a part of the government's redistribution policy, as in Cremer et al (1998). The issue of disaggregating the Pigouvian tax component and the redistributive component would raise issues similar to those in earlier sections.

Finally, suppose we introduce the possibility of costly abatement into the nonlinear tax model using the same abatement technology as in section 5. As before, the producer price in industry D is $1+t_E$, and aggregate profits are $\Pi = t_E A - Z(A)$, which are taxed away by the government. The consumer price for D is $q_D = 1 + t_D + t_E$. The outcome of stage 2 of the consumers' utility maximization is exactly as before, yielding consumer demands $D_i(q_D, I_i)$ for $i = 1,2$, demand $\hat{D}_2(q_D, I_1)$ for the mimicker, indirect utilities $W^i(q_D, I_i, Y_i)$ and $\hat{W}^2(q_D, I_1, Y_1)$, and the envelope results (33).

The government policy problem is the same as before with $t_D + t_E$ replacing t_D and profits added to the government budget constraint. The Lagrangian expression becomes:

$$\begin{aligned}\mathcal{L} = {} & \rho_1 N_1 W^1(q_D, I_1, Y_1) + \rho_2 N_2 W^2(q_D, I_2, Y_2) + \overline{N}\left(\overline{E} - \delta N_1 D_1(q_D, I_1) - \delta N_2 D_2(q_D, I_2)\right) \\ & + \overline{N}\delta A(t_E) + \gamma\left(W^2(q_D, I_2, Y_2) - \hat{W}^2(q_D, I_1, Y_1)\right) \\ & + \lambda\left(N_1(Y_1 - I_1 + (t_D + t_E)D_1(q_D, I_1)) + N_2\left(Y_2 - I_2 + (t_D + t_E)D_2(q_D, I_2)\right) - Z(A(t_E)) - R\right)\end{aligned}$$

The first-order conditions on I_1, Y_1, I_2, Y_2 and t_D are the same as (34)–(38) with $t_D + t_E$ replacing t_D, where $q_D = 1 + t_D + t_E$. The first-order condition on t_E is:

$$\begin{aligned}& \rho_1 N_1 W^1_{q_D} + (\rho_2 N_2 + \gamma) W^2_{q_D} - \gamma \hat{W}^2_{q_D} - \overline{N}\delta\left(N_1 \frac{\partial D_1}{\partial q_D} + N_2 \frac{\partial D_2}{\partial q_D} - A'(t_E)\right) \\ & + \lambda\left(N_1 D_1 + N_1(t_D + t_E)\frac{\partial D_1}{\partial q_D} + N_2 D_2 + N_2(t_D + t_E)\frac{\partial D_2}{\partial q_D} - Z'(A)A'(t_E)\right) = 0 \qquad (41)\end{aligned}$$

Substituting the first-order condition for t_D (the analog of (38)) into (41) and using $t_E = Z'(A)$, we obtain:

$$t_E = \frac{\overline{N}\delta}{\lambda} \tag{42}$$

Next, proceed exactly as in the previous section by combining the first-order conditions on I_1, I_2 and t_D to obtain the analog of (39):

$$t_D + t_E = \frac{\overline{N}\delta}{\lambda} = \frac{(\rho_1 N_1 + \rho_2 N_2)\delta}{\lambda} \tag{43}$$

Then, since $t_E = \overline{N}\delta / \lambda$ by (42), we obtain $t_D = 0$. Thus, unlike in the linear tax case, the Pigouvian tax consists only of a tax on emissions: no tax on *D* is required. Moreover, the interpretation of the Pigouvian tax is the same as above. More weight is put on marginal damages to the low-wage than the high-wage persons.

We can summarize these results of Pigouvian taxation in the presence of nonlinear income taxation as follows.

Proposition 5: *When households differ in wage rates but have the same quasilinear-in-leisure preferences, and the government can use a nonlinear income tax:*

1. *the Pigouvian component of the second-best commodity tax system is a tax on good D which differs from social marginal damages by putting relatively more weight on damages to low-wage persons and less weight on damages to high-wage persons,*
2. *the structure of the optimal nonlinear income tax is the same as in the absence of pollution, and*
3. *when a costly abatement technology is available, the Pigouvian tax applies only to emissions and not to the output of the dirty good, and differs from social marginal damages in the same way as 1. above.*

8. Concluding Remarks

Our purpose in this paper has been to study how Pigouvian taxes differ from marginal damages to households when the taxes are levied alongside distortionary taxes. Because taxes used for corrective purposes and taxes used for revenue-raising, which we refer to as Ramsey taxes, are interdependent, it is generally not possible to disentangle one from the other. We have specified a model in which Pigouvian taxes can readily be disaggregated from Ramsey taxes and used that to determine how the former deviates from marginal social damages, and how that deviation varies with the amount of revenue that must be raised. The formulation we use involves household preferences that are quasilinear in leisure so that demands for goods depend only on own prices. In this setting, qualitatively clear results emerge.

When the government relies only on commodity taxes, the Pigouvian component of the tax is the deviation from uniformity of commodity taxes on the clean and dirty goods. The Pigouvian component will exceed marginal damages when goods' demands are inelastic, and vice versa. Not surprisingly, the same result applies when, instead of

commodity taxes, a labor income tax is used along with a tax on the dirty good. When an abatement technology is introduced into the basic setting, the Pigouvian component of taxation follows the same pattern as in the basic model. However, while damage is now a result of emissions, the Pigouvian component includes both a tax on emissions and a tax on the sales of the dirty good, even though in a first-best world, only the former would be used. Finally, when households are heterogeneous and the tax system is used not just for revenue-raising but also for redistribution, matters change. In the case where the government uses all the instruments at its disposal and levies a nonlinear income tax, there is no longer a presumption that the Pigouvian tax will be either higher or lower than marginal social damages. Instead, the tax is set equal to a sum of damages to the low- and high-wage persons in the economy but with relatively more weight put on marginal damages to the former and less to the latter compared with social damages. The addition of abatement possibilities changes little in this case. Unlike with linear taxation, only a tax on emissions should be used and not a tax on total use of the dirty good.

Our analysis relies on some important simplifying assumptions, particularly the quasilinearity of preferences. As mentioned, these preferences were used to facilitate the clean separation of corrective from revenue-raising components of taxation. In a more complicated setting, analytical solutions would generally be hard to come by, so one may have to resort to computational methods.

References

Atkinson, A.B. and J.E. Stiglitz, 1976, "The Design of Tax Structure: Direct vs. Indirect Taxation", *Journal of Public Economics*, 6, 55–75.

Atkinson, A.B. and J.E. Stiglitz, 1980, *Lectures on Public Economics* (New York: McGraw-Hill).

Auerbach, A.J. and J.R. Hines, Jr., 2002, "Taxation and Economic Efficiency", in: A.J. Auerbach and M.S. Feldstein, eds., *Handbook of Public Economics*, Vol. 3 (Amsterdam: Elsevier), Chapter 21.

Bovenberg, A.L. and R.A. de Mooij, 1994, "Environmental Levies and Distortionary Taxation", *American Economic Review*, 94, 1085–1089.

Cremer, H. and F. Gahvari, 2001, "Second-Best Taxation of Emissions and Polluting Goods", *Journal of Public Economics*, 80, 169–197.

Cremer, H., F. Gahvari and N. Ladoux, 1998, "Externalities and Optimal Taxation", *Journal of Public Economics*, 70, 343–364.

Cremer, H., F. Gahvari and N. Ladoux, 2001, "Second-Best Pollution Taxes and the Structure of Preferences", *Southern Economic Journal*, 68, 258-280.

Edwards, J., M. Keen and M. Tuomala, 1994, "Income Tax, Commodity Taxes and Public Good Provision: A Brief Guide", *Finanzarchiv*, 51, 472–497.

Feldstein, M.S., 1972, "Distributional Equity and the Optimal Structure of Public Prices", *American Economic Review*, 62, 32–36.

Jaeger, W.K., 2002, "Carbon Taxation when Climate Affects Productivity", *Land Economics*, *78*, 354–67.

Nava, M., F. Schroyen and M. Marchand, 1996, "Optimal Fiscal and Public Expenditure Policy in a Two-Class Economy", *Journal of Public Economics*, 61, 119–137.

Sandmo, A., 1975, "Optimal Taxation in the Presence of Externalities", *Swedish Journal of Economics*, *77*, 86–98.

Sandmo, A., 2006, "Global Public Economics: Public Goods and Externalities", Norwegian School of Economics and Business Administration, mimeo. http://nhh.no/sam/res-publ/2006/32.pdf

Stiglitz, J.E., 1987, "The Theory of Pareto-Efficient and Optimal Redistributive Taxation", in: A.J. Auerbach and M.S. Feldstein, eds., *Handbook of Public Economics*, Vol. 2 (Amsterdam: Elsevier), Chapter 15.