

Aigner, Rafael

Working Paper

Environmental taxation and redistribution concerns

Preprints of the Max Planck Institute for Research on Collective Goods, No. 2011,17

Provided in Cooperation with:

Max Planck Institute for Research on Collective Goods

Suggested Citation: Aigner, Rafael (2011) : Environmental taxation and redistribution concerns, Preprints of the Max Planck Institute for Research on Collective Goods, No. 2011,17, Max Planck Institute for Research on Collective Goods, Bonn

This Version is available at:

<https://hdl.handle.net/10419/57482>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

**Preprints of the
Max Planck Institute for
Research on Collective Goods
Bonn 2011/17**



**Environmental Taxation
and Redistribution
Concerns**

Rafael Aigner



MAX PLANCK SOCIETY



Environmental Taxation and Redistribution Concerns

Rafael Aigner

July 2011

Environmental Taxation and Redistribution Concerns

Rafael Aigner*

22nd July 2011

This paper shows how the optimal level of Pigouvian taxation is influenced by distributive concerns. With second-best instruments, a higher level of income redistribution calls for a lower level of Pigouvian taxation. More redistribution implies that tax collection via the income tax creates higher distortions, which in turn makes revenues from Pigouvian taxation more valuable. Contrary to naive intuition, this reduces the optimal level of Pigouvian taxation. The social planner trades off environmental tax revenues against the marginal social damage and accepts a lower tax if the welfare created per dollar is higher. The paper also shows that the relation between levels of redistribution and Pigouvian taxation is reversed in first-best. It thus highlights that second-best Pigouvian taxes are very different from their first-best counterpart – despite apparently identical first order conditions.

Keywords: Pigouvian taxation, comparative statics, externalities, optimal income taxation, second-best

JEL classification: H21, H23, D62

*University of Bonn and Max Plank Institute for Research on Collective Goods. I thank Felix Bierbrauer, Pierre Boyer, Brian Cooper, Christoph Engel, Martin Hellwig, Marco Kleine, Simon Lange, Johannes Spinnewijn, and Christian Traxler for very helpful comments and discussions, as well as participants of the BGSE Micro Workshop, the LSE work in progress seminar and the MMM Workshop. The usual disclaimer applies. Contact: Kurt-Schumacher-Str. 10, D-53113 Bonn, Germany. E-mail: aigner@uni-bonn.de

1. Introduction

In June, 2008 *The Economist* published an article discussing the pros and cons of the oil price boost over the preceding years. On the one hand, the rise might be considered as “a gigantic carbon tax” that helped fighting global warming. On the other hand, it particularly hurt the poor who spent a considerably higher proportion of their income on fuel than the rich. Financial compensations for the core demand for energy could help to solve the issue. However “it seems odd to try to prevent energy use with higher taxes . . . and then to subsidise it” (The Economist, 2008). The article thereby raised the question: how to design green taxes optimally while accounting for distributive concerns? Rising awareness for global environmental problems under persisting inequality has increased the salience of that question. In this paper, I give an answer to it. I focus on the optimal level of environmental taxation and how it changes with the level of redistribution.

To tackle the problem formally, I employ a simple Mirrlees (1971) income taxation framework that is extended by a consumption externality as in Cremer et al. (1998). Within this framework, a welfare-optimising government uses non-linear income taxes to redistribute and Pigouvian taxation to reduce negative externalities as proposed by Pigou (1932). I show that the two tax design problems are interconnected. In particular, the higher the level of redistribution, the lower the optimal level of environmental taxation. The tax level has two determinants. First, the marginal social damage caused by the externality. Second, the cost of public funds. By those I mean a measure of direct welfare losses associated with marginal income tax collection. If the government puts more weight on redistribution, it will have to accept higher cost of public funds. Marginal revenues from the environmental tax are then more valuable from the government’s point of view. Contrary to naive intuition, this calls for a lower environmental tax rate. The reason is that the tax rate is at its efficient level if the marginal revenues exactly compensate society for the marginal externality that is associated with the harmful activity. The more valuable the marginal revenues are, the less one needs to compensate for the marginal externality. To put it another way consider the Pigouvian tax as a bribe that consumers pay the authorities in order to get allowance for pollution. The government is willing to accept a lower bribe if its utility per dollar is higher. Exactly this is the case if the costs of public funds are higher.

In the analysis, the level of redistribution is measured by a parameter that corresponds to the weight of less productive agents in a social-welfare function. As explained, Pigouvian taxation needs to decrease if the parameter increases. The result is reversed, though, if first-best instruments are available. Without distortions, the costs of public funds actually decrease in the parameter, as the disutility of the hard-working high-productive agents receives less weight in the welfare function. Hence the first-best level of Pigouvian taxation increases if the willingness to redistribute increases.

The main contribution with respect to the existing literature is to draw attention to the *level* of Pigouvian taxation. Most of the literature focuses on tax *rules* and concludes that the distortions caused by second-best instruments do not alter the tax rules compared to first-best. I show that, despite the first-best shape of these rules, the second-best level of Pigouvian taxation in fact depends on the distortions and the available income tax instruments. Also, Pigouvian taxation is lower in second- than in first-best.

As a second contribution, I analyse a laissez-faire regime that introduces a green tax. To leave the income distribution untouched, green tax revenues are given back in a lump-sum fashion, such that every agent's net tax payment is zero. The lack of monetary transfers between agents might appeal to policy makers for various reasons. Yet, I show that such a reform is not Pareto-optimal and that one always needs a welfare judgement to determine an optimal environmental tax.

Related literature

Sandmo (1975) integrates externality taxation into a Ramsey framework with linear taxation and finds that a Pigouvian term is just added to the standard Ramsey term ("additivity property"). The linear model is also used to explore the double-dividend hypothesis, according to which revenues from Pigouvian taxes may be recycled to reduce pre-existing, distortio-nary taxes to evoke a second dividend. Surveys include Goulder (1995), Bovenberg (1999), and Schöb (1997).

This paper contributes to a strand of literature in which Pigouvian taxation meets non-linear income taxes under asymmetric information (Mirrlees, 1971). Cremer et al. (1998) show that, under the separability assumptions from Atkinson and Stiglitz (1976), the optimal Pigouvian tax rate is uniform, i.e., it does not discriminate between agents. Gauthier and Laroque (2009) generalise the insight: a certain part of the second-best problem can be separated such that first-best rules apply for that part of the problem. Examples include Pigouvian taxation and the Samuelson Rule. Hellwig (2010) presents a similar result.

Kopczuk (2003), Pirttilä and Tuomala (1997), and Kaplow (2006) explicitly analyse externality taxation. Kaplow (2006) summarises that "simple first-best rules – unmodified for labour supply distortion or distribution – are correct in a natural, basic formulation of the problem." My analysis highlights that distribution and distortions have a significant influence on environmental policy with respect to tax *levels*, though. For a similar model, Jacobs and de Mooij (2011, p. 2) find that the "optimal second-best tax on an externality-generating good should not be corrected for the marginal cost of public funds". The different conclusions are due to different definitions of the cost of public funds.

Last but not least, my analysis relates to the literature on comparative static properties of non-linear taxation with and without a public good (Weymark, 1987; Brett and Weymark, 2008; Bierbrauer and Boyer, 2010) and to more applied analyses of the question how to overcome negative distributional effects of environmental taxes (like Metcalf, 1999; West, 2005; Ekins and Dresner, 2004). For a recent study on the U.S. economy, see Rausch et al. (2011).

The paper is organised as follows: Section 2 outlines the model. Section 3 states the rule for optimal internalisation. Section 4 examines properties of tax systems that implement the optimal allocations and provides the main result on the comparative statics of Pigouvian taxation. Section 5 describes optimal internalisation without redistributive transfers. Section 6 concludes. The Appendix holds proofs and formal results. It also provides a characterisation of optimal allocations.

2. Model

Production

The model is based on Cremer et al. (1998). It considers three different goods. First, an intermediate good that is referred to as output and is denoted by Y . It serves as the numeraire and may also be interpreted as money. Second, a *clean*, completely private consumption good, C , and third, a *dirty* consumption good, D . The intermediate good can be transformed into the consumption goods at a fixed rate of transformation equal to p_C and p_D , respectively. Parameters p_C and p_D may be interpreted as the producer prices of C and D . The intermediate good itself can be produced with a linear technology using labour as the single input good (but labour is not modelled explicitly). The rate of transformation between labour and the intermediate good mirrors productivity and is denoted by w . It may be interpreted as the wage rate.

Households and allocations

There is a continuum of measure one of agents that differ in exactly one dimension, namely their productivity, which can be either low or high. An agent's type is denoted by $\theta \in \{L, H\}$. The fraction of low-type agents is denoted by $\gamma \in (0, 1)$. An *allocation* A specifies levels of (C, D, Y) for both generic types, i.e., $A \equiv (C_L, D_L, Y_L, C_H, D_H, Y_H)$. For a given allocation the utility of an agent of type θ is

$$U_\theta(A) = u(C_\theta, D_\theta) - \frac{Y_\theta}{w_\theta} - (\gamma D_L + (1 - \gamma)D_H)e. \quad (1)$$

Function u is continuously differentiable three times, strictly increasing, strictly concave, has nonnegative cross derivatives and satisfies the Inada conditions.¹ It represents private consumption utility. In order to produce Y_θ units of output, an agent has to provide Y_θ/w_θ units of labour. This provision is associated with a linear disutility. The last term in the utility function reflects the externality. Independently of his type, every agent suffers from the overall consumption of the dirty good $\gamma D_L + (1 - \gamma)D_H$. The social harm is proportional to total dirty good consumption, and $e > 0$. From an agent's point of view, the own consumption has no negative effect on the own utility as their single contribution is negligible in comparison to the larger contribution of others. Individual contributions are in fact zero due to the assumption of a continuum of agents.²

Social planner

This paper takes a normative perspective by examining what a social planner (SP) would do in order to maximise the social welfare function W , defined as

$$W(A) = \alpha U_L(A) + (1 - \alpha)U_H(A), \quad \alpha \in (0, 1),$$

¹Formally, $u_{CD} \geq 0$ as well as $u_K \rightarrow \infty$ as $K \rightarrow 0$, and $u_K \rightarrow 0$ as $K \rightarrow \infty$, for $K \in \{C, D\}$. The Inada conditions are imposed in order to guarantee strictly positive optimal consumption levels. Strict concavity guarantees unique solutions.

²Externalities of this type were termed "atmospheric" by Meade (1952). A different way to interpret the mechanism is a public good that is provided by nature (like "fresh air" or "nice atmosphere") and that is reduced by the consumption of the dirty good in such a way that only total consumption matters. In the model presented here the initial amount of this public good would be normalised to zero.

where A is the allocation. The welfare function is a weighted sum of the generic types' utilities. The parameter α measures the weight SP puts on the generic low-type agent. If $\alpha = \gamma$, then W is the utilitarian welfare function. For $\alpha = 1$, W would be the Rawlsian welfare function.

The social planner has to take account of some constraints. Overall, the economy cannot consume more than it produces in terms of output. Furthermore, an exogenous revenue requirement r has to be met. The resource constraint is

$$\gamma(Y_L - p_C C_L - p_D D_L) + (1 - \gamma)(Y_H - p_C C_H - p_D D_H) - r \geq 0. \quad (2)$$

If (2) holds and $A \geq 0$, then A is *feasible*. An allocation that maximises W among all feasible allocations is a *first-best allocation*.

If the social planner does not observe an agent's type, not all feasible allocations are implementable. If, for instance, an allocation disadvantages the high-type agents, they might have an incentive to pretend to be low-types, making it impossible to implement this allocation.³ Thus, under asymmetric information, SP has to ensure that agents do not want to misrepresent their type. This is the case if the following *incentive-compatibility* constraints hold.

$$u(C_L, D_L) - \frac{Y_L}{w_L} \geq u(C_H, D_H) - \frac{Y_H}{w_L}, \quad (3)$$

$$u(C_H, D_H) - \frac{Y_H}{w_H} \geq u(C_L, D_L) - \frac{Y_L}{w_H}. \quad (4)$$

An allocation that maximises welfare among all feasible, incentive-compatible allocations is a *second-best allocation*.

I restrict the analysis to the case where SP likes to redistribute from high-type agents to low-type agents.

Assumption 1. Parameters are such that $\alpha(1 - \gamma)w_H > (1 - \alpha)\gamma w_L$.

The assumption holds if SP puts a sufficiently high welfare weight on low-type agents. The lower w_L is relative to w_H , the lower α may be, because a large difference in productivity provides an efficiency argument for making high-types work more than low-types. A low population share γ of low-types makes redistribution in their favour very cheap, hence it also allows for a low α .

Given the shape of u , it is always optimal to produce strictly positive amounts of the consumption goods rather than abstain from economic activity. Hence the non-negativity constraint is an issue only with respect to output requirements. It may be the case, though, that indeed only one type of agents produces output and that the non-negativity constraint of the other type is binding. In fact, at the first-best allocation under Assumption 1 only high-type agents work. For second-best, it is ambiguous whether low-type agents work. I analyse the cases in which they do so. In these, the incentive constraint of high-types is always binding; low-types' IC is always slack.

³For a precise notion of implementation and its relation to incentive-compatible allocations see Section 4 and Appendix B.

3. Optimal Internalisation in First- and Second-Best

This section provides a general property of Pareto-optimal allocations with respect to the externality. At first sight, the presented rule is identical for first- and second-best allocations. This is the reason why redistribution or distortions are sometimes considered to have no important influence on Pigouvian taxation. In the next step, however, I show in what way the first- and second-best rules are in fact different. To shorten exposition, I use the following notation for $J, K \in \{C, D\}$. $u^L := u(C_L, D_L)$, $u_J^L := \partial u(C_L, D_L)/\partial J_L$, $u_{KJ}^L := \partial^2 u(C_L, D_L)/(\partial K_L \partial J_L)$. Analogous definitions apply to $u^H := u(C_H, D_H)$. The Lagrangian multiplier of the resource constraint is denoted by λ . All results in the current section are derived in Appendix A.

A rule for optimal internalisation

Both first- and second-best allocation feature the property that the marginal rates of substitution (MRS) between the two consumption goods are the same for both types of agents. Rather than being equal to the rate of transformation (namely, producer-price ratio), as would be the case in an unregulated market, the MRS is equal to

$$MRS = \frac{u_D^L}{u_C^L} = \frac{u_D^H}{u_C^H} = \frac{p_D}{p_C} + \frac{e}{\lambda p_C}. \quad (5)$$

This is a standard result in the literature. It follows, for instance, from the more general analysis by Hellwig (2010). It is driven by the separability feature of the utility functions. Cremer et al. (1998) point out the relation to the famous result in Atkinson and Stiglitz (1976), namely that under the given assumptions commodity prices should not be distorted, and all redistribution can be done within the labour market. The intuition is as follows. By assumption, all agents have the same consumption pattern.⁴ Therefore the commodity demand cannot be used to screen types and commodity taxation cannot contribute to relax the equity-efficiency trade-off. Hence there is no point in distorting them.

The intuition is expected to carry over to the case where an externality is introduced. In fact, as agents are equal in terms of their consumption preferences and their exposure to the externality, there is no point in treating them differently in this respect. Yet, it is no longer true that redistribution only affects the labour market. Optimal consumption now depends on multiplier λ , which measures the marginal loss in welfare given a marginal increase in the revenue requirement r . The multiplier is crucially related to redistribution. Also, while at first sight the above formula is the same for both first- and second-best allocation, λ is different in first- and second-best, and this has significant consequences for the relation between the degree of redistribution and the degree of intervention in the commodity market.

The quasi-linearity in labour allows for closed-form solutions for the multiplier λ and makes the dissimilarity between first- and second-best evident at once. It also reveals the dependency on the underlying parameters α and γ . The multiplier at an interior second-best allocation is

$$\lambda^* = \frac{\alpha}{w_L} + \frac{1 - \alpha}{w_H}. \quad (6)$$

⁴More precisely, for a given amount of total consumption spending, all agents consume the same commodity bundle.

To grasp the intuition, note that agents do not benefit from r , so an increase is pure burden. A way to finance the additional requirement is to increase output. As the incentive constraint for the high-type agents is binding, their output may only be increased if the low-type's output is increased as well. The weighted welfare loss of such an increase is equal to α/w_L for the generic low-type and $(1 - \alpha)/w_H$ for the generic high-type. Note that the multiplier does not depend on the population shares. The reason is that a higher revenue requirement has to be produced by all agents (and independently of their type) in order to sustain incentive compatibility.

The multiplier at the first-best allocation is

$$\lambda^F = \frac{1 - \alpha}{w_H(1 - \gamma)}. \quad (7)$$

Because only high-types work in first-best, only parameters related to them matter for the cost of public funds. If SP needs an additional unit of revenue, he will make high-type agents work more. As there are only $1 - \gamma$ high-type agents, the generic high type has to provide $1/(1 - \gamma)$ (marginal) units of output and needs to work $1/(w_H(1 - \gamma))$ additional hours. The incurred marginal disutility is weighted by $1 - \alpha$.

The multipliers are not only different in size, but also with respect to their directions of change in the parameters α and γ . The welfare weight thereby has an impact on the optimal tax design with respect to the externality. If interpreted naively, the rule itself hides this fact.

4. Taxation

The current section adapts the interpretation of output being money. In this interpretation, Y denotes gross income, w corresponds to the wage rate, and p_C, p_D are producer prices. Now it is possible to introduce taxes and to find tax systems that implement particular allocations.

A tax system $\tau = (t_C, t_D, T(\cdot))$ consists of an income tax function T and specific commodity taxes $t_C, t_D \in \mathbb{R}$. Consumer prices are defined as $q_K := p_K + t_K$, $K \in \{C, D\}$. With $T : (w, Y) \mapsto T(w, Y)$, it is possible to find a system τ , such that individual maximisation will result in the first-best allocation. With $T : Y \mapsto T(Y)$, it is possible to find a system τ , such that individual maximisation will result in the second-best allocation. This insight allows to restrict attention to the chosen tax structure. (Formal arguments are provided in appendix B.)

For any type θ , let $(C_\theta(\tau), D_\theta(\tau), Y_\theta(\tau))$ be the maximisers of individual utility:

$$(C_\theta(\tau), D_\theta(\tau), Y_\theta(\tau)) \in \underset{(C, D, Y)}{\operatorname{argmax}} \left(u(C, D) - \frac{Y}{w_\theta} - (\gamma D_L + (1 - \gamma) D_H) e \right. \\ \left. \text{s. t. } q_C C + q_D D \leq Y - T(\cdot) \right). \quad (8)$$

As before, the externality is not relevant for an agent's decision. Let $V_\theta(\tau)$ denote the highest utility an individual agent can achieve given a tax system. In order to internalise and redistribute optimally, the social planner then needs to find a tax system τ that maximises

$$\alpha V_L(\tau) + (1 - \alpha)V_H(\tau)$$

subject to the fiscal budget constraint

$$\gamma \left(T[Y_L(\tau), \cdot] + t_C C_L(\tau) + t_D D_L(\tau) \right) + (1 - \gamma) \left(T[Y_H(\tau), \cdot] + t_C C_H(\tau) + t_D D_H(\tau) \right) \geq r,$$

which is equivalent to the resource constraint (2). In first-best, the income tax may be contingent on the agent's type, while in second-best it can only be contingent on observed gross income Y . $T(\cdot)$ may be negative, in which case it is a transfer to the agent. Also, consumption may be subsidised through negative commodity taxes.

4.1. Pigouvian Taxation

This section investigates how the need to internalise the externality shapes the optimal tax system. At first sight, the tax on the dirty good t_D should be the object of interest. However, the taxation problem does not have a unique solution. It would be possible to normalise one of the commodity taxes to zero. Yet, this approach suffers from the fact that the normalisation choice might influence the conclusions. Schöb (2003, 1997), for instance, demonstrates this in the context of the double dividend hypothesis. Therefore I do not use a normalisation but propose a different concept. I define an object g that is uniquely determined by the implemented allocation and that mirrors the tax system's inherent incentives to reduce pollution. As it turns out, g coincides with t_D if t_C is set to zero. Hence, after all, normalising t_C to zero and examining t_D is equivalent to my analysis. Still, my approach exposes the full range of possible commodity tax combinations that yield optimal internalisation.

A tax system's greenness

Consider an agent who faces some tax system τ and decides to purchase an additional unit of D , while keeping her total spendings constant by reducing her consumption of C . The reallocation has consequences for the social planner's tax revenue. This hypothetical change in revenues is the *greenness* g of the tax system.

Definition 1. The *greenness* of tax system $\tau = (t_C, t_D, T)$ is denoted by g and is defined as

$$g := t_D - t_C \frac{q_D}{q_C}.$$

The greenness is constructed in such a way that individual utility maximisation implies $MRS = (p_D + g)/p_C$. From the agent's point of view g not only includes what they have to pay in taxes for additional consumption of D , but also what they save in taxes when consuming less C . The Pigouvian rationale is to provide agents with the incentive to shift consumption from D to C . The greenness is an accurate measure of how intense this incentive is, and, unlike a particular tax rate, it is independent of the normalisation of the tax system.

As a further advantage, the concept of greenness provides the full range of possible internalisation schemes (given tax structure τ). If, for technical or political reasons, one of the commodity taxes is not available or restricted, the definition of g gives an immediate answer on how to set the other commodity tax in order to reach efficiency.

Still, there is a clear relation between actual taxes and g . In particular, if t_C is normalised to zero, g is simply equal to t_D . If t_D is normalised to zero then, instead of punishing consumption of the dirty good, the tax system subsidises consumption of the clean good. Note that for *any* t_D , one can find a t_C such that g equals some desired value.

Rather than speaking of a (single) Pigouvian tax, I use the term *Pigouvian taxation* when referring to a tax system's feature of incentivising agents to reduce an externality. More precisely, a tax system involves Pigouvian taxation if and only if $g > 0$.⁵

Efficient Pigouvian taxation

How high should g be? An optimum is characterised by the fact that a marginal reallocation does not change welfare. In particular, keeping private spending constant, a marginal change in consumption levels must not change welfare. Consider a marginal shift from C to D (for all agents, taking account of differences in prices). This has three effects: (1) Consumption utility u is unchanged as agents are at their individual optimum. (2) External harm increases at rate e . (3) Tax revenues increase at rate g and relax the budget constraint of the social planner. The Lagrangian multiplier, λ , tells how welfare is affected, because relaxing the budget constraint is equivalent to reducing the revenue requirement r . Thus the marginal effect of tax revenues on welfare amounts to $g\lambda$. The overall marginal change in welfare is $-e + g\lambda$. For the change to be zero, g needs to be

$$g = \frac{e}{\lambda}. \quad (9)$$

Obviously g is increasing in e . If the externality is a more severe problem, incentives to lower dirty good consumption should be higher. It is less obvious, though, that g is inversely proportional to λ , to which I refer to as the *cost of public funds*. To grasp the intuition behind the inverse relation consider the purpose of Pigouvian taxation: its only goal is to restore the efficient level of dirty good consumption. From a welfare perspective, a unit of the dirty-good should be consumed if and only if consumption is not only individually optimal, but private benefits also outweigh social harm. Consequently, dirty-good consumption is at its socially optimal level if marginal private (net) benefits exactly equal marginal social harm. To measure and compare these two objects, it is useful to quantify them in terms of money.

(1) The optimising agent is willing to pay g units of additional taxes for the right to consume her last unit of D rather than spending the respective money on C . So g is a good measure of (net) private benefits of the marginal unit of dirty-good consumption.

(2) Now consider the social planner. If D increases by one unit, welfare decreases by e . If SP receives exactly e/λ units of money to relax the budget constraint, welfare increases by $(e/\lambda)\lambda = e$. Thus the marginal social harm measured in money is equal to e/λ . It is exactly the amount of money that society needs as a compensation for additional dirty-good consumption. The amount is lower if the received money is more useful in the sense that

⁵If $g < 0$, a Pigouvian subsidy is present to correct for a positive externality, i.e., for $e < 0$

cost of public funds are higher. Putting together (1) and (2) shows that if $g = e/\lambda$ then individual maximisation leads to an allocation in which, at the margin, private (net) benefits equal social harm.

The inverse relation of g and λ drives the comparative statics of Pigouvian taxation. While naive intuition suggests that higher cost of public funds (associated with marginal tax revenue being more valuable) should lead to higher Pigouvian tax rates, the above analysis reveals that the opposite is true. The more valuable the marginal tax revenues, the less is needed to compensate for the marginal externality, and – because the one and only purpose of Pigouvian taxation is to induce alignment of private benefits and social harm at the margin – a lower Pigouvian tax rate is asked for.

A note on the double-dividend hypothesis

The double-dividend hypothesis states that revenues from a Pigouvian tax may be used to reduce other, distortionary taxes and may thereby entail a positive effect that exceeds the environmental benefits. Consequently, Pigouvian tax revenues are more useful if other taxes are more distortionary, and – taking the revenues into account – one might want the Pigouvian tax to be higher in that case. The above discussion reveals, though, why this reasoning is misleading. First, the optimal level of Pigouvian taxation is based on marginal considerations only, total revenues are irrelevant. Second, if marginal revenues are more useful, then, contrary to the double dividend rationale the tax level should be lower.

The strong form of the double-dividend hypothesis states that a revenue-neutral introduction of green taxes is desirable even if environmental benefits are not taken into account. Empirical investigations by Goulder (1995) tend to reject the hypothesis. In the model that I present, it fails clearly: Pigouvian taxation, namely $g > 0$, is optimal only if an externality is present, i.e., if $e > 0$. Bovenberg (1999) gives the same argument, albeit for a model with linear taxation.

A note on regressive taxes

Environmental taxes are regressive if tax payments in proportion to total consumption spending decrease in total consumption spending. A tax designer should take this into account with respect to equity concerns. Applied work on the impact of green tax reforms on low-income households include Metcalf (1999), West (2005), Ekins and Dresner (2004), and Rausch et al. (2011). From a theoretical perspective, the optimal level of Pigouvian taxation is not affected by these considerations – at least under the separability assumptions employed in this paper. While the next section finds a dependency of optimal tax levels on the desire to redistribute, the relationship has nothing to do with regressivity.

4.2. Comparative Statics of Pigouvian Taxation

A tax system τ is said to *implement* allocation A if A is the result of individually optimal behaviour given τ . The formal requirement is $A = (C_L(\tau), D_L(\tau), Y_L(\tau), C_H(\tau), D_H(\tau), Y_H(\tau))$, where the right-hand side is defined by (8).

This subsection examines the comparative statics properties of the greenness of tax systems which implements the first- and second-best allocation, respectively. The greenness g

depends only on e and λ . While the former is an exogenous constant, the latter is endogenously determined and drives the results.

Proposition 1 (First-best Pigouvian taxation). *Let τ^F be a tax system that implements the first-best allocation A^F . Then the greenness g^F of the tax system is uniquely determined and features the following properties.*

$$\frac{\partial g^F}{\partial \alpha} > 0, \quad \frac{\partial g^F}{\partial \gamma} < 0.$$

Recall from (7) that the cost of public funds at a first-best allocation are $\lambda^F = (1 - \alpha)/(w_H(1 - \gamma))$ and are determined by a welfare-weighted output increase of high-types. If α increases, SP cares less about high-types working more and tax revenues generated by g are less valuable per unit so that more (marginal) revenues need to be collected at the optimum. If γ is increased, the generic high-type has to work more for an higher overall output requirement and cost of public funds increase. Marginal revenues generated by g now have higher value per unit and less is needed to satisfy optimality condition (9).

Proposition 2 (Second-best Pigouvian taxation). *Let τ^* be a tax system that implements an interior first-best allocation A^* . Then the greenness g^* of the tax system is uniquely determined and features the following properties.*

$$\frac{\partial g^*}{\partial \alpha} < 0, \quad \frac{\partial g^*}{\partial \gamma} = 0.$$

From (6) the costs of public funds at a second-best allocation are $\lambda^* = \alpha/w_L + (1 - \alpha)/w_H$. As λ does not depend on γ , neither does g . The higher the welfare weight of low-type agents, the more redistribution is asked for and the more distortions are accepted. Higher distortions imply higher excess burden of taxation and thereby higher cost of public funds. Marginal revenues from Pigouvian taxation are then more valuable and less marginal revenue is needed to satisfy optimality condition (9).

Proposition 3. *Let g^* be the greenness of a tax system that implements an interior second-best allocation. Analogously, let g^F be the respective element for the first-best allocation. Then (at fixed parameters)*

$$g^* < g^F.$$

Proof. (7) and (6) imply $\lambda^* - \lambda^F > 0$ through Assumption 1. $g^* < g^F$ follows from (9). \square

Pigouvian taxation is higher in first-best because cost of public funds are lower.

This section has highlighted a link between Pigouvian taxation and the degree of redistribution measured by the welfare weight α . Welfare optimising societies with different opinions about equity need to have different levels of Pigouvian taxation, even if first-best instruments are feasible. This is not entirely obvious because in basic partial equilibrium models, the level of Pigouvian taxation is pinned down solely by Pareto efficiency. While Pareto efficiency is a widely accepted goal, there is no consensus about the “right” level of equity,

in particular if equity-efficiency trade-offs are present. This gives rise to the question whether a society that does not agree on a particular α or has ideological objections against any redistribution of income can nevertheless agree on some Pigouvian taxation to deal with an externality. Put differently, what is the optimal policy that avoids transfers between agents? The next section gives an answer to this question.

5. A Transfer-Free Allocation

The major exercise of this paper is to find out the relation between optimal internalisation and distributive concerns. This section tries to focus solely on efficiency by dropping the possibility of redistributing income. Consider a social planner who is aware of the externality and wants to internalise it with Pigouvian taxation, but objects to (monetary) transfers between agents. Such a situation will be defined by requiring an allocation to be *transfer-free*. To avoid the question who should pay the cost of running government, this section presumes that these costs are zero, i.e., $r = 0$.

Definition 2. An allocation A is called *transfer-free* if

$$Y_L = p_C C_L + p_D D_L \quad \text{and} \quad Y_H = p_C C_H + p_D D_H.$$

From a political-economy perspective, such a requirement makes sense if agents have veto rights and care predominantly about the monetary implications of a proposed reform rather than their environmental benefits. The media often provide figures about who wins or loses from a particular proposal in terms of monetary consequences. These “hard” facts might be more relevant for voters than predicted long-term or environmental effects which are hard to measure. Also, reforms involving monetary transfers may be opposed just for ideological reasons. From a theoretical perspective, a transfer-free allocation is interesting because it is as close as possible to the laissez-faire allocation while still internalising the externality.

Overall, the requirement tries to rule out distributive issues in the design of an environmental policy. The analysis, however, shows that the optimal transfer-free allocation is not pinned down by Pareto Efficiency. It thereby reinforces the claim that environmental policy cannot be separated from welfare judgments.

Lemma 1. *If A^{tf} is an allocation that maximises the welfare function $W(A) = \alpha U_L(A) + (1 - \alpha)U_H(A)$ among all transfer-free allocations, then it satisfies*

$$\frac{u_D^L}{u_C^L} = \frac{p_D + ew_L \gamma / \alpha}{p_C}, \quad \frac{u_D^H}{u_C^H} = \frac{p_D + ew_H (1 - \gamma) / (1 - \alpha)}{p_C}.$$

Proof. Straightforward Lagrangian optimisation. □

The main message from the Lemma is that the optimal transfer-free allocation depends on welfare weight α as well as on the population share γ . Furthermore, the marginal rates of substitution between consumption goods differ for the two types.

At first glance, the results look surprising. By construction, the transfer-free optimum is expected to be independent of welfare judgements. Also, all previously analysed optima featured identical marginal rates of substitutions. To outline the intuition behind the results, consider the simple case of $p_C = p_D = 1$. Consumption of D_L creates a net benefit for the group of low-type agents, taking into account the *internal externality* it imposes on the group itself. In addition it imposes an *external externality* on the group of high-type agents. The trade-off between the benefits of the group L and the external externality on group H requires a welfare judgement. Group H dislikes group L 's consumption of D , but is not able to pay them for reducing D_L . To be more concrete, consider the marginal effects of increasing D_L (and accordingly Y_L):

Internal-external effect on a generic low-type: $u_D^L - 1/w_L - e\gamma$.

External-external effect on a generic high-type: $-e\gamma$.

The marginal effect on welfare is $\alpha(u_D^L - 1/w_L - e\gamma) + (1 - \alpha)(-e\gamma)$ and must equal zero at a solution. Rearranging terms gives

$$u_D^L = \frac{1}{w_L} + \frac{e\gamma}{\alpha}$$

If α is higher, then the external-external effect is less important and group L should be allowed to consume more of the dirty good. If γ is higher, the bigger group size of group L implies a higher external-external effect and group L should consume less of the dirty good. To be even clearer about what is going on, suppose α goes to one. Then D_H should be very low because the consumption utility it creates is weighted by only $1 - \alpha \approx 0$, while the external-external effect weighs close to one. On the other hand, D_L may be high because the external-external effect can be neglected. With no monetary transfer allowed, regulation of D_L, D_H is the only way to distribute utility among the two type agents. It is, however, an inefficient way as the next Lemma reveals.

Lemma 2. *Let $\alpha w_H(1 - \gamma) \neq (1 - \alpha)w_L\gamma$. Then the optimal transfer-free allocation A^{tf} is not Pareto-efficient.*

The statement can be proofed explicitly, but the result also becomes clear if one appreciates the fact that the constraint of A being transfer-free is more restrictive than the ordinary resource constraint (2) with $r = 0$, and that the optimal transfer-free allocation does not satisfy the efficiency condition (5) for optimal internalisation.⁶

The discussion shows that requiring allocations to be transfer-free is not a very good idea in the first place. Not only does it fail in providing an internalisation policy that is free from welfare judgements, but it is also Pareto-dominated by an allocation that allows for monetary transfers. From a political or ideological perspective, a transfer-free allocation might be appealing, but from a welfare perspective it is not.

⁶For a formal proof that works for all parameter ranges, consider a small perturbation around A^{tf} that is feasible and holds one type's utility constant. Then show that the other type's utility can be increased.

6. Conclusion

This paper looks at the interdependency of distributive and environmental policies from a normative perspective. It is obvious that if agents have different tastes for environmental protection, the optimal policy depends on distributional considerations. To reduce these kinds of effects, all agents are modelled to be as homogeneous as possible. This is hence not meant to be a simplification, but rather is necessary to isolate the effects that actually are of interest.

As a consequence, the structure of the model parallels those of previous contributions, which find that the rules for environmental taxes do not discriminate between agents and look the same for first- and second-best settings. Normally, first-best worlds do not suffer from a conflict between equity and efficiency, and one might be tempted to conclude that in these settings there is no real interdependency between environmental and distributive goals.

This paper, however, shows that the optimal level of Pigouvian taxation indeed depends on distributive goals and that this dependency crucially is down to whether first- or second-best instruments are available. It thus argues that it is not possible to separate the two task of redistribution and environmental protection. It strengthens the point by showing that a society that does not want to intervene in the income distribution will nonetheless need to make a welfare judgement. Moreover, Pareto efficiency requires monetary transfers between agents.

What can be learned in terms of policy implications? First, the view that the two goals of redistribution and environmental protection can be addressed independently by means of two different instruments (income tax and Pigouvian taxation) needs to be reconsidered. In particular, the designer of environmental taxes has to account for the value in terms of welfare that is created by the tax revenues. To determine the optimal level of taxation (and pollution), the designer has to trade external harm off against useful tax revenues. Yet, and this is the second point, the optimum is pinned down by looking at the revenues from the last unit of pollution, i.e., the marginal revenues. Total revenues in contrast are irrelevant for the optimal level of Pigouvian taxation. Environmental taxation does not have a general function to shift the tax burden from labour income to externalities. The designer of the income tax system just takes total revenues from environmental taxes as given and adjust the total revenue requirements accordingly.

APPENDIX

A. Properties of First- and Second-Best Allocation

Lemma 3. *If A^F is a solution to the first-best problem under Assumption 1, then $Y_L^F = 0$ and $Y_H^F > 0$.*

Proof. Suppose $Y_L^F > 0$. If the total output of all low-type agents is lowered by $\Delta \in (0, \gamma Y_L^F)$, every low-type individual may reduce his own output by Δ/γ . The immediate

welfare gain is $\alpha\Delta/(\gamma w_L)$. To finance the output reduction high types have to increase their total output by Δ resp. their individual output by $\Delta/(1 - \gamma)$. The immediate welfare loss is $(1 - \alpha)\Delta/((1 - \gamma)w_H)$. The net effect of the alteration is strictly positive given Assumption 1, a contradiction.

Hence $Y_L^F = 0$ and $Y_H^F > 0$ need to hold given the Inada-conditions on u . \square

Lemma 4. *If A is a solution to the second-best problem, then*

1. *At most one incentive compatibility constraint is binding.*
2. *$Y_H > Y_L$ and $u^H > u^L$.*

Proof. 1. Suppose the contrary. Summation of both ICs yields $Y_L = Y_H$ and $u(C_L, D_L) = u(C_H, D_H)$. Due to the shape of u , this can be optimal only if $(C_L, D_L) = (C_H, D_H)$. To complete the argument, it suffices to show that such a bunching allocation is dominated by a *constrained laissez-faire* allocation. Fix any feasible bunching allocation $A^b = (C^b, D^b, Y^b; C^b, D^b, Y^b)$ and define for any type θ , $(C_\theta^{lf}, Y_\theta^{lf}) := \operatorname{argmax}_{C, Y} \{u(C, D^b) - Y/w_\theta \text{ s.t. } Y \geq p_C C + p_D D^b + r\}$. Then, in particular, $u(C_\theta^{lf}, D^b) - Y_\theta^{lf}/w_\theta \geq u(C^b, D^b) - Y^b/w_\theta$. Furthermore, maximisers are unique and $(C_L^{lf}, Y_L^{lf}) \neq (C_H^{lf}, Y_H^{lf})$. Thus there exist θ such that $u(C_\theta^{lf}, D^b) - Y_\theta^{lf}/w_\theta > u(C^b, D^b) - Y^b/w_\theta$. The constraint laissez-faire allocation A^{lf} thereby Pareto-dominates the bunching allocation A^b . A^{lf} is also incentive compatible and feasible. Hence, A^b cannot be a solution to the second-best problem and the contradiction is completed. (The argument builds on Bierbrauer and Boyer, 2010, Lemma 1)

2. Add both ICs to obtain $Y_H \geq Y_L$. Equality would imply a bunching allocation which is not optimal as shown above. Hence $Y_H > Y_L$. IC_H then implies $u^H > u^L$. \square

Lemma 5. *If A^* is an interior solution of the second best problem under Assumption 1, then high types' incentive constraint (4) is binding, low types' incentive constraint (3) is slack.*

Proof. Suppose by contradiction that (4) was slack, i.e., $u^H - Y_H/w_H > u^L - Y_L/w_L$. Then there exists an $\epsilon > 0$ such that also $u_H - (Y_H + \epsilon)/w_H > u_L - (Y_L - \epsilon(1 - \gamma)/\gamma)/w_L$. The ϵ -perturbed allocation is constructed in a way to keep total output constant. Incentive compatibility is sustained, too. The welfare effect of the perturbation is

$$dW = \alpha \frac{1 - \gamma}{\gamma w_L} \epsilon - \frac{(1 - \alpha)\epsilon}{w_H} > 0 \Leftrightarrow \frac{\alpha}{\gamma w_L} > \frac{(1 - \alpha)}{(1 - \gamma)w_H}$$

dW is strictly positive precisely under Assumption 1, hence a contradiction.

If IC_H is binding then IC_L must be slack by Lemma 4. \square

First-Order Conditions

Considering the lemmas, an appropriate Lagrangian is as follows.

$$\begin{aligned} \mathcal{L} = & \alpha [u(C_L, D_L) - Y_L/w_L - (\gamma D_L + (1 - \gamma)D_H)e] \\ & + (1 - \alpha) [u(C_H, D_H) - Y_H/w_H - (\gamma D_L + (1 - \gamma)D_H)e] \\ & + \lambda (\gamma(Y_L - p_C C_L - p_D D_L) + (1 - \gamma)(Y_H - p_C C_H - p_D D_H) - r) \\ & + \mu (u(C_H, D_H) - Y_H/w_H - u(C_L, D_L) + Y_L/w_H) \\ & + \delta Y_L \end{aligned} \quad (10)$$

Next, set the partial derivatives to zero.

$$\alpha u_C^L - \gamma \lambda p_C - \mu u_C^L = 0 \Leftrightarrow (\alpha - \mu)u_C^L = \lambda \gamma p_C \quad (11)$$

$$\alpha u_D^L - \gamma \lambda p_D - \mu u_D^L - \gamma e = 0 \Leftrightarrow (\alpha - \mu)u_D^L = \lambda \gamma p_D + \gamma e \quad (12)$$

$$-\alpha/w_L + \gamma \lambda + \mu/w_H + \delta = 0 \Leftrightarrow \gamma \lambda = \alpha/w_L - \mu/w_H - \delta \quad (13)$$

$$(1 - \alpha)u_C^H - (1 - \gamma)\lambda p_C + \mu u_C^H = 0 \Leftrightarrow (1 - \alpha + \mu)u_C^H = (1 - \gamma)\lambda p_C \quad (14)$$

$$(1 - \alpha)u_D^H - (1 - \gamma)\lambda p_D + \mu u_D^H - (1 - \gamma)e = 0 \Leftrightarrow (1 - \alpha + \mu)u_D^H = (1 - \gamma)(\lambda p_D + e) \quad (15)$$

$$-(1 - \alpha)/w_H + (1 - \gamma)\lambda - \mu/w_H = 0 \Leftrightarrow (1 - \alpha + \mu)/w_H = (1 - \gamma)\lambda \quad (16)$$

It follows that

$$\mu = w_H \left(\frac{(1 - \gamma)\alpha}{w_L} - \frac{\gamma(1 - \alpha)}{w_H} \right) - w_H(1 - \gamma)\delta \quad (17)$$

$$\lambda = \frac{\alpha}{w_L} + \frac{1 - \alpha}{w_H} - \delta \quad (18)$$

For a first best optimum, set $\mu = 0$, for an interior second best optimum, set $\delta = 0$.

Proposition 4 (First-best allocation). *An allocation A^F is a first-best solution given Assumption 1 if and only if it satisfies the following system of equations.*

$$u_C^L = \frac{p_C}{w_H} \frac{1 - \alpha}{\alpha} \frac{\gamma}{1 - \gamma}, \quad u_D^L = \frac{p_D + e/\lambda^F}{w_H} \frac{1 - \alpha}{\alpha} \frac{\gamma}{1 - \gamma}, \quad (19a)$$

$$u_C^H = \frac{p_C}{w_H}, \quad u_D^H = \frac{p_D + e/\lambda^F}{w_H}, \quad (19b)$$

$$Y_L^F = 0, \quad Y_H^F = \frac{\gamma}{1 - \gamma} (p_C C_L^F + p_D D_L^F) + p_C C_H^F + p_D D_H^F + \frac{r}{1 - \gamma}, \quad (20)$$

$$\lambda^F = \frac{1 - \alpha}{w_H(1 - \gamma)}. \quad (21)$$

Low-type agents do not work at all. Due to linear disutility from working, Assumption 1 implies that any given amount of output requirement fosters lower aggregated disutility if it is provided solely by high types rather than low types. If Y_L could be negative, welfare would be unbounded.

For a moment, ignore the Lagrangian multiplier of the resource constraint λ^F . Then consumption of high types is independent of the welfare weight and the population shares, and is just determined by efficiency considerations. It departs from standard results only through a corrective element that takes care of the external effects of dirty-good consumption. The consumption levels of the low-type agents, though, heavily depend on welfare weights as well as the population shares. The underlying trade-off lies between consumption utility of low-types and disutility of high types, who have to work for the provision of low-type consumption. Low-type productivity w_L is irrelevant for the allocation given that they do not work.

Proof of proposition 4. With $\mu = 0$, the Lagrange function (10) is concave and the first order conditions are necessary and sufficient for a solution.

Take the first-order conditions (13) and (16) and set $\mu = 0$. Then

$$\lambda^F = \frac{1 - \alpha}{w_H(1 - \gamma)}, \quad \delta^F = \frac{\alpha}{w_L} - \gamma\lambda > 0 \Rightarrow Y_L^F = 0.$$

Notice that the inequality is satisfied if and only if Assumption 1 holds. The statements of the Proposition now follow from conditions (11), (12), (14), (15), and the binding resource constraint (2). \square

Lemma 6. *Let $u(C, D)$ be strictly concave and continuously differentiable and let k^C, k^D be two constants such that the system $u_C(C, D) = k^C, u_D(C, D) = k^D$ has a solution. Then the solution is unique.*

Proof. Consider the three-dimensional space. Let $s = (s^C, s^D)$ be a solution. The tangential plane at $S = (s^C, s^D, u(s))$ is spanned by the directions of the two partial derivatives at S . As u is strictly concave, the whole range of u – except $u(s)$ – lies below that plane. Now consider a point s' that also solves the above system but is different from s . The tangential plane at s' is parallel to the one at s , yet one of the plane is higher than the other. But than it is no longer possible that the whole range of u lies below the lower plane. This creates a contradiction. \square

Proposition 5 (Interior second-best allocation). *If A^* is an interior solution of the second-best problem, then it is unique and solves the following system of equations, where $d = \mu^*(1 - w_L/w_H)/(\alpha - \mu^*)$.*

$$u_C^L = \frac{p_C}{w_L}(1 + d) \quad u_D^L = \frac{p_D + e/\lambda^*}{w_L}(1 + d) \quad (22)$$

$$u_C^H = \frac{p_C}{w_H} \quad u_D^H = \frac{p_D + e/\lambda^*}{w_H} \quad (23)$$

$$Y_L^* = r + \gamma(p_C C_L^* + p_D D_L^*) + (1 - \gamma)(p_C C_H^* + p_D D_H^*) - w_H(u^H - u^L)(1 - \gamma)$$

$$Y_H^* = r + \gamma(p_C C_L^* + p_D D_L^*) + (1 - \gamma)(p_C C_H^* + p_D D_H^*) + w_H(u^H - u^L)\gamma$$

$$\lambda^* = \frac{\alpha}{w_L} + \frac{1 - \alpha}{w_H}, \quad \mu^* = \alpha(1 - \gamma)\frac{w_H}{w_L} - (1 - \alpha)\gamma. \quad (24)$$

The conditions for high-type consumption levels are almost identical to the corresponding first-best conditions (19). The subtle but important difference lies in the Lagrangian multiplier λ , which is different in first- and second-best and, most importantly, features different comparative statics properties.

The consumption levels of low types are distorted downwards, i.e., the labour-consumption choice is distorted in favour of leisure. The distortion is captured by d and is higher if μ^* is higher or the difference in productivities is larger.

Proof of proposition 5. If A^* is an interior second-best solution, then it satisfies conditions (13) and (16) with δ set to zero. Then λ^* and μ^* are uniquely determined and strictly positive. For given values of λ^* and μ^* , (11), (14), (12), and (15) uniquely determine the consumption levels (uniqueness is established by Lemma 6). Output requirements follow from the binding resource constraint (2) combined with the binding incentive constraint (4). \square

B. Taxation and Implementation

Definition 3. A tax system $\tau = (t_C, t_D, T(\cdot))$ specifies constant specific tax rates for the two consumption goods and an income tax function T . Resulting consumer prices are denoted by $q_C := p_C + t_C$ and $q_D := p_D + t_D$.

There is no need to relax the linearity of commodity taxes. First- and second-best allocations can be implemented with a system that features linear consumption taxes. In the first-best case, the income tax may be conditioned on the type of an agent, whereas under information constraints, the income tax does depend only on Y .

Definition 4. A tax system $\tau = (t_C, t_D, T(\cdot))$ is said to *implement* allocation A if

1. For any θ , $(C_\theta, D_\theta, Y_\theta) \in \operatorname{argmax}_{(C,D,Y)} \{u(C, D) - Y/w_\theta \text{ s.t. } q_C C + q_D D \leq Y - T(\cdot)\}$
2. $\gamma(Y_L - p_C C_L - p_D D_L) + (1 - \gamma)(Y_H - p_C C_H - p_D D_H) - r \geq 0$.

A tax system with linear commodity taxes can only implement allocations in which the marginal rates of substitution between the two consumption good are equal across all agents.

Lemma 7. If tax system τ implements allocation A , then

$$\frac{q_D}{q_C} = MRS = \frac{u_D(C_L, D_L)}{u_C(C_L, D_L)} = \frac{u_D(C_H, D_H)}{u_C(C_H, D_H)}$$

Proof. The individual choices $(C_\theta(\tau), D_\theta(\tau), Y_\theta(\tau))$ as defined in (8) satisfy the first-order conditions $u_C^\theta = q_C/w_\theta$, and $u_D^\theta = q_D/w_\theta$. It follows that $u_D^\theta/u_C^\theta = q_D/q_C$. \square

Lemma 8. There exists a tax system $\tau = (t_C, t_D, T)$ with $T : (w, Y) \mapsto T(w, Y)$ that implements the first-best allocation $A^F = (C_L^F, D_L^F, Y_L^F, C_H^F, D_H^F, Y_H^F)$.

Proof. Fix some t_C, t_D such that $q_D/q_C = (p_D + e/\lambda^F)/p_C$. Define the income tax function for $w \in \{w_L, w_H\}$ to be

$$T(w, Y) = \begin{cases} Y & w = w_L, \text{ and } Y \neq 0 \\ -(q_C C_L^F + q_D D_L^F), & w = w_L \text{ and } Y = 0 \\ Y_H^F - (q_C C_H^F + q_D D_H^F) + Y, & w = w_H \text{ and } Y \neq Y_H^F \\ Y_H^F - (q_C C_H^F + q_D D_H^F), & w = w_H \text{ and } Y = Y_H^F \end{cases}$$

We show that an agent who solves the individual maximisation problem (8) chooses exactly the bundle that is intended for his or her type.

(a) Low types. For low-type agents it is never optimal to choose $Y \neq 0$. So their problem reduces to $\max_{C,D} u(C, D)$ s.t. $q_C C + q_D D \leq q_C C_L^F + q_D D_L^F$. Necessary and sufficient conditions for a solution are $u_D/u_C = q_D/q_C$ and $q_C C + q_D D = q_C C_L^F + q_D D_L^F$. By construction, the first-best consumption levels (C_L^F, D_L^F) satisfy these conditions.

(b) High types. For high-type agents, it is never optimal to choose $Y \neq Y_H^F$. So their problem reduces to $\max_{C,D} u(C, D)$ s.t. $q_C C + q_D D \leq q_C C_H^F + q_D D_H^F$. Necessary and sufficient conditions for a solution are $u_D/u_C = q_D/q_C$ and $q_C C + q_D D = q_C C_H^F + q_D D_H^F$. By construction the first best consumption levels (C_H^F, D_H^F) satisfy these conditions. \square

Lemma 9. *There exists a tax system with $T : Y \mapsto T(Y)$ that implements the second-best allocation $A^* = (C_L^*, D_L^*, Y_L^*, C_H^*, D_H^*, Y_H^*)$.*

Proof. Fix some t_C, t_D such that $q_D/q_C = (p_D + e/\lambda^*)/p_C$. Define the income tax function $T(Y)$ to be

$$T(Y) = \begin{cases} T_L, & Y = Y_L^* \\ T_H, & Y = Y_H^* \\ Y & \text{otherwise} \end{cases}$$

where $T_\theta = Y_\theta - q_C C_\theta^* - q_D D_\theta^*$.

An agent who faces the individual maximisation problem defined in (8) will never choose $Y \notin \{Y_L^*, Y_H^*\}$. Thus the agents' problem can be decomposed into two steps: (1) $\hat{u}(B) := \max_{C,D} u(C, D)$ s.t. $q_C C + q_D D = B$. (2) $\max_Y \hat{u}(B) - Y/w$ s.t. $[B = (Y_L^* - T_L)$ if $Y = Y_L^*$, $B = (Y_H^* - T_H)$ if $Y = Y_H^*$, and $B = 0$ otherwise].

Necessary and sufficient conditions for a solution of step (1) are $u_D/u_C = q_D/q_C$ and $q_C C + q_D D = B$. By construction the conditions are met by (C_L^*, D_L^*) if $B = Y_L^* - T_L$ and by (C_H^*, D_H^*) if $B = Y_H^* - T_H$. Consequently, step (2) is equivalent to choosing between (C_L^*, D_L^*, Y_L^*) and (C_H^*, D_H^*, Y_H^*) . As the second-best allocation satisfies the incentive compatibility constraints (4) and (3), each agent does indeed choose the bundle intended for his or her type. \square

Proof Propositions 1 and 2

Proof of Proposition 1. From Proposition 4 and Lemma 7 $MRS = p_D/p_C + e/(\lambda^F p_C) = q_D/q_C$. It follows that $e/\lambda^F = t_C - t_C q_D/q_C = g$. From (7) $\lambda^F = (1 - \alpha)/(w_H(1 - \gamma))$ and the results follow simply by taking derivatives. \square

Proof of Proposition 2. From Proposition 5 and Lemma 7 $MRS = p_D/p_C + e/(\lambda^* p_C) = q_D/q_C$. It follows that $e/\lambda^* = t_C - t_C q_D/q_C = g$. From (24) $\lambda^* = \alpha/w_L + (1 - \alpha)/w_H$ and the results follow simply by taking derivatives. \square

References

ATKINSON, A. B. AND J. E. STIGLITZ (1976): "The design of tax structure: Direct versus indirect taxation," *Journal of Public Economics*, 6, 55–75.

- BIERBRAUER, F. AND P. BOYER (2010): “The Pareto-Frontier in a simple Mirrleesian model of income taxation,” *Preprints of the Max Planck Institute for Research on Collective Goods*, Bonn 2010/16.
- BOVENBERG, A. (1999): “Green tax reforms and the double dividend: an updated reader’s guide,” *International Tax and Public Finance*, 6, 421–443.
- BRETT, C. AND J. A. WEYMARK (2008): “Public Good Provision And The Comparative Statics Of Optimal Nonlinear Income Taxation,” *International Economic Review*, 49, 255–290.
- CREMER, H., F. GAHVARI, AND N. LADOUX (1998): “Externalities and optimal taxation,” *Journal of Public Economics*, 70, 343 – 364.
- EKINS, P. AND S. DRESNER (2004): *Green taxes and charges: reducing their impact on low-income households*, Joseph Rowntree Foundation.
- GAUTHIER, S. AND G. LAROQUE (2009): “Separability and public finance,” *Journal of Public Economics*, 93, 1168–1174.
- GOULDER, L. (1995): “Environmental taxation and the double dividend: a reader’s guide,” *International Tax and Public Finance*, 2, 157–183.
- HELLWIG, M. (2010): “A generalization of the Atkinson-Stiglitz (1976) theorem on the undesirability of nonuniform excise taxation,” *Economics Letters*.
- JACOBS, B. AND R. DE MOOIJ (2011): “Pigou Meets Mirrlees: On the Irrelevance of Tax Distortions for the Second-Best Pigouvian Tax,” *CESifo Working Paper Series*, 3342.
- KAPLOW, L. (2006): “Optimal control of externalities in the presence of income taxation,” *NBER working paper*, 12339.
- KOPCZUK, W. (2003): “A note on optimal taxation in the presence of externalities,” *Economics Letters*, 80, 81–86.
- MEADE, J. (1952): “External economies and diseconomies in a competitive situation,” *The economic journal*, 62, 54–67.
- METCALF, G. (1999): “A distributional analysis of green tax reforms,” *National Tax Journal*, 52, 655–681.
- MIRRLYES, J. (1971): “An exploration in the theory of optimum income taxation,” *The Review of Economic Studies*, 175–208.
- PIGOU, A. (1932): *economics of welfare.*, Macmillan. London.
- PIRTTILÄ, J. AND M. TUOMALA (1997): “Income tax, commodity tax and environmental policy,” *International Tax and Public Finance*, 4, 379–393.

- RAUSCH, S., G. E. METCALF, AND J. M. REILLY (2011): "Distributional Impacts of Carbon Pricing: A General Equilibrium Approach with Micro-Data for Households," *National Bureau of Economic Research Working Paper Series*, No. 17087.
- SANDMO, A. (1975): "Optimal taxation in the presence of externalities," *The Swedish Journal of Economics*, 77, 86–98.
- SCHÖB, R. (1997): "Environmental taxes and pre-existing distortions: The normalization trap," *International Tax and Public Finance*, 4, 167–176.
- (2003): "The Double Dividend Hypothesis of Environmental Taxes: A Survey," *CESifo Working Paper*, 946.
- THE ECONOMIST (2008): "Let them heat coke – How green taxes hurt the poor," *The Economist*, June 12th 2008.
- WEST, S. (2005): "Equity implications of vehicle emissions taxes," *Journal of Transport Economics and Policy*, 39, 1–24.
- WEYMARK, J. (1987): "Comparative static properties of optimal nonlinear income taxes," *Econometrica: Journal of the Econometric Society*, 55, 1165–1185.