Taxing Externalities under Financing Constraints

Hoffmann, Florian and Inderst, Roman and Moslener, Ulf

University of Frankfurt, University of Frankfurt and Imperial College London, Frankfurt School of Finance and Management, UNEP Center for Climate and Sustainable Energy 2013

August 2013

Taxing Externalities under Financing Constraints*

Florian Hoffmann† Roman Inderst‡ Ulf Moslener§

August 2013

Abstract

We consider an economy where production generates externalities, which can be reduced by additional firm level expenditures. This requires firms to raise outside financing, leading to deadweight loss due to a standard agency problem vis-à-vis outside investors. Policy is constrained as firms are privately informed about their marginal cost of avoiding externalities. We first derive the optimal linear pollution tax, which is strictly lower than the Pigouvian tax for two reasons: First, higher firm outside financing creates additional deadweight loss; second, through redistributing resources in the economy, a higher tax reduces average productive efficiency. We analyze various instruments that achieve a more efficient allocation, in particular, nonlinear pollution taxes, which can no longer be implemented through a tradable permit scheme alone, and grants tied to loans, which are frequently observed in practice.

*We thank seminar participants at Carlos III Madrid, Imperial College London, and the University of Frankfurt for helpful comments.

†University of Frankfurt. E-mail: fhoffmann@finance.uni-frankfurt.de.

‡University of Frankfurt and Imperial College London. E-mail: inderst@finance.uni-frankfurt.de.

§Frankfurt School of Finance and Management, UNEP Center for Climate and Sustainable Energy Finance. E-mail: u.moslener@fs.de.

1 Introduction

We consider an economy where financially constrained firms must invest to reduce externalities from production. The amount of external financing that firms raise interferes with productive efficiency due to a standard moral-hazard problem. We show how in such an environment, the optimal linear tax on externalities - or, likewise, the optimal amount of tradable pollution rights - differs from the Pigouvian tax and how there is scope for additional policies, such as loan-based grants. As we discuss below, such grants are frequently observed in practice, in particular in relation to sizable investments into CO_2 reduction, which should indeed be non-negligible in terms of firms' financing capability. Our contribution is, however, more general, as we explore optimal government intervention in a production economy with externalities, financing constraints, and private information about the individual costs of reducing these externalities.

Pigou (1920) showed that the optimal tax on a good that generates externalities should be equal to the marginal external effect that arises from the consumption and production of that good. We identify two reasons why, in the presence of financing constraints, the optimal linear tax is strictly below the respective Pigouvian tax. First, the optimal tax - or, likewise, the aggregate amount of permissible pollution rights - must take into account the inefficiencies caused when more outside finance must be raised to cover additional abatement costs. Second, when firms have different (marginal) abatement costs, even when tax proceeds are distributed back to agents in equilibrium, a higher tax leads to a redistribution that reduces average productive efficiency - again working through the underlying agency problem of firms vis-à-vis outside investors.

When the government is not constrained by, for instance, a supranational scheme of tradable pollution rights, which is akin to a linear tax, we show how it can improve efficiency through implementing, instead, a nonlinear tax. Under the optimal nonlinear tax, the tax-induced benefit from abatement is highest at relatively low levels and at relatively high levels. Though this thus dampens the impact on the "average polluter", compared to those firms that end up with relatively low or relatively high abatement, we show that due to the increased efficiency it can lead to higher abatement for all firms under the optimal nonlinear tax (compared to the optimal linear tax).

In particular when the use of nonlinear taxes is restricted by a system of tradable pollution rights, but also more generally, efficiency can be improved by introducing loan-based grants. Though these must still respect incentive compatibility, i.e., prevent opportunistic behavior, they allow to compensate more effectively for the redistribution of resources that is generated by a tax on externalities. Combining taxes on externalities with grants linked to loans - as is frequently observed in practice - thus dominates taxes or pollution permits, as it allows to achieve the same reduction in externalities while ensuring higher aggregate efficiency. In fact, even when tradable permits or emission taxes are used to regulate emissions, they are often applied in combination with other, finance-linked interventions. In Germany, for instance, a state-owned bank (KfW) provides on a large scale subsidized credit to businesses that apply energy-saving technologies or invest to reduce CO_2 output. The UK government, in turn, is setting up a "Green Investment Bank" which will provide investment subsidies and low-interest loans.

Our paper is somewhat related to the literature that analyzes the effect of liability on environmental care. In some of this literature (cf. the survey in Boyer et al. 2007), compensation for damages is restricted by agents' limited resources or the limited liability embedded in the financial structure that they use to finance production. Imposing an extended liability also on the providers of outside finance may then impact efficiency, in particular in the presence of financial frictions and imperfect financial markets (cf. Pitchford 1995, Boyer and Laffont 1997, Tirole 2010).[1] As noted above, our results deviate from the Pigou rule, which would prescribe to set the tax so as to fully internalize the marginal social damage from pollution. Our argument is different from that in Atkinson and Stern (1974), who find that the second-best provision of public goods under private information will be lower than the first-best provision, and also different from the general notion in the public finance literature that a tax on externalities may itself add distortions in production or consumption. Recent contributions in the public finance literature have restored the Pigou rule, most notably through the use of nonlinear income taxes that compensate for such tax-induced distortions on production and consumption in these models (cf. Kaplow 2006; Jacobs and de Moij 2011).[2] We solve for the optimal mechanism and show that in

[1] The interaction between private financial frictions and public policy has been addressed also in the literature on entrepreneurship that examines various rationales for policy intervention, in particular the possible spillover effects created by start-ups. Boadway and Tremblay (2005) offer a broad overview of the literature, which mainly focuses on tax considerations.

[2] As pointed out in Kaplow (2006), the key is to recognize that the environmental tax will induce, ceteris paribus, redistribution and that this affects the relationship between the Pigouvian tax increment

the presence of outside financing constraints there is still a wedge between the optimal marginal tax rate and the Pigouvian tax.[3]

The rest of this paper is organized as follows. Section 2 introduces the economy. Section 3 derives some preliminary results. In Sections 4 and 5 we solve for the optimal linear and nonlinear tax on the externality. In Section 6 we allow the government to use, as an additional instrument, a grant linked to the size of firms' loans. Section 7 summarizes our results. The Appendix collects proofs as well as additional technical material to which we refer to in the main text.

2 The Economy

Agents and Endowments. We consider an economy populated by a unit mass of agents indexed by $i \in I = [0, 1]$. There are two points of time: $t = 0$ and $t = 1$. Agents have access to the same production technology that pays out in the final period $t = 1$. Abstracting first from both the presence of a policymaker and the presence of externalities, starting production in $t = 0$ requires the investment of $I_0 \geq 0$ and generates in $t = 1$ either zero output or an output of $x > 0$.[4] The likelihood of a positive outcome depends on the non-observable, real-valued effort e that the respective agent exerts. For our purpose it will be convenient to make agents' utility separable in effort cost (as well as in the consumption of externalities, in what follows). The respective effort cost function is denoted by $c(e)$, while the likelihood of positive output x is given by $p(e)$. Here, $c(e)$ and $p(e)$ are both strictly increasing.

For what follows, we only need that the functions are continuously differentiable, so that a marginal change in contractual and policy parameters has indeed a marginal impact on efficiency through affecting the choice of effort. However, it is convenient to stipulate, in addition, that $c'' > 0$, $p'' \leq 0$, $c'(0) = 0$, $p'(0) > 0$, and that $c'(e)/p'(e)$ becomes sufficiently

and social marginal damages. Cf. earlier in a similar vein Diamond and Mirrlees (1971), who state that distributional concerns do not justify violating production efficiency if the government can optimally adjust taxes (on consumption).

[3] Rothschild and Scheuer (2013) also find a deviation of the optimal externality correction from the Pigouvian level in a model where agents can earn their income in traditional activities as well as through socially unproductive rent-seeking. When the government cannot observe the shares of individual agents' income earned in the two activities the corrective component of the optimal non-linear income tax scheme in their model deviates from the Pigouvian correction due to a "sectoral shift effect" in general equilibrium.

[4] In our model, the investment outlay and output are both measured in the same unit of "resources" or capital.

large as $p(e)$ approaches one. Taken together, these conditions ensure a unique, interior solution for the effort choice problem of the agent.

Agents have originally zero resources and thus have to raise capital from outside investors. To be specific, we may think of this as raising capital abroad. We stipulate that the agents' utility is linear in the resources that they consume and that they do not discount future consumption, which is why in our model the only rationale for borrowing (i.e., raising outside finance) is for production. In terms of contracting with outside investors, we stipulate that the output realization is verifiable and can thus be part of a financial contract.

Externalities. Production generates negative externalities, which can be reduced by additional investment. Precisely, we stipulate that when the production of agent i creates $y_i \geq 0$ units of these externalities, then this affects all other agents equally and, thereby, generates the social loss ρy_i with $\rho > 0$. Recall that we assume that utility is also additively separable in externalities. That is, when after $t = 1$ agent i is left with w_i resources for consumption and has exerted effort e_i, then his total utility is

$$u_i = w_i - c(e_i) - \rho \int_{j \in I} y_j dj. \tag{1}$$

Note that this implies that an agent's private incentives to reduce his own externality y_i are zero, given that the impact is distributed uniformly across all agents (with mass one). Externalities are verifiable.

Generating a level of externalities y, e.g., by using the respective technology mix or by operating production accordingly, is associated with a particular production cost. It is convenient, albeit this is without loss of generality, to stipulate that there is a given maximum level of externalities $\overline{y}$ (per firm). Consequently, the respective *avoided* externalities can be denoted by $a = \overline{y} - y$.

An agent's cost of avoiding externalities depends on his type. Define the strictly positive real-valued type by θ_i, where we assume that θ_i is, for all $i \in I$, independently and identically distributed according to the distribution function $F(\theta)$, permitting a density function $f(\theta) > 0$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$. We capture abatement costs by the twice continuously differentiable function $K(a, \theta)$ with $K(0, \theta) = 0$, $K_1(a, \theta) > 0$ for $a \in (0, \overline{y})$, $K_1(0, \theta) = 0$, $K_1(\overline{y}, \theta) = \infty$, and $K_{11}(a, \theta) > 0$, as well as $K_2(a, \theta) < 0$ for $a \in (0, \overline{y})$.[5] Further, we

[5] The subscripts indicate partial derivatives with respect to the respective argument, e.g., $K_1(a, \theta) :=$

stipulate that types θ are ordered such that

$$K_{12}(a,\theta) := \frac{\partial^2 K(a,\theta)}{\partial a \partial \theta} < 0. \tag{2}$$

That is: Higher types θ have everywhere strictly lower marginal costs of abatement. For instance, we could take $K(a,\theta) = k(a)/\theta$, where $k(a)$ inherits the properties of $K(a,\cdot)$. Note that the respective costs are incurred, together with the investment I_0, right when production starts in $t = 0$. Further, observe that the agent's economic success (output of x or zero) does not directly interact with the generation of externalities. However, as we show below, the likelihood of success interacts with the incurred abatement costs and thus the need to raise more external finance.[6]

Our chosen set-up, where the reduction in negative externalities is a function of investment, allows also for the following alternative interpretation. We could think of y or likewise $a = \overline{y} - y$ as a (continuous) verifiable technology choice, e.g., the "amount" of additionally installed fuel-efficient equipment or of energy-efficient building material that is used when setting up production. Though the purchase costs may be the same for all agents, agents' costs of installation or, more generally, their total opportunity costs may differ, given the buildings, equipment, and technologies that they already own. For higher θ, the associated costs are lower according to condition (2).[7]

Feasible Policies. We introduce a utilitarian policymaker, who maximizes the expected utility of all agents: $E\left[\int u_i di\right]$. For simplicity, we refer to the policymaker as the government and consider various policy instruments. Our benchmark is that of a linear tax on externalities, coupled with a transfer that is paid out of tax receipts. We characterize the optimal linear tax and show how with financial constraints this is strictly different from the Pigouvian tax. As we argue, the outcome can also be implemented through a market for pollution rights. There is, however, no scope for the government to raise finance on

$\partial K(a,\theta)/\partial a$.

[6] Our assumption that there is no direct interaction between externalities and the likelihood of high output allows to restrict contracts with external investors to repayments that are conditional only on realized output, as this is then a sufficient statistic for effort. However, we will later allow for the possibility that also output is taxed as a way to redistribute resources between agents of different type.

[7] When taking this interpretation, a given choice of technology could then be associated with some (possibly stochastic) generation of externalities (for which we would then need a different notation). Without loss of generality, any policy could then, however, target directly the adoption of the technology, a. Also, the agents' utility function (1) and thus also the policymaker's objective function could be rewritten accordingly, namely as a function of expected externalities, without changing results.

behalf of agents, unless it would use this to, at the same time, redistribute resources. This can, however, also be achieved through providing grants linked to the amount of outside finance that agents privately raise, which is a policy that is frequently observed in practice (cf. the Introduction). We characterize the optimal grant scheme. With such grants in place, we argue further that there is no additional role that taxes levied on *output* could play for the purpose of efficiently redistributing resources.[8]

A further improvement of efficiency can, however, be achieved when the linear tax - or, likewise, a market for pollution rights that induces such a linear tax - is replaced by a nonlinear tax scheme. Here, we use a mechanism design approach to solve for the optimal such nonlinear tax scheme and we illustrate the difference to the optimal linear tax with the help of numerical examples.

3 Preliminary Result: The Outside Financing Problem

Consider the problem of an agent who must raise capital L (a "loan") to start production. As he can only pay back in case output is positive, the contract with outside investors can be restricted to a single variable: The repayment R that is made in case the output equals x. Given some repayment R, note that the agent's payoff is

$$p(e)(x-R)-c(e),$$

so that the uniquely optimal effort level e^* is given by the first-order condition

$$p'(e^*)(x-R)-c'(e^*)=0. \tag{3}$$

This can be substituted to obtain the investors' break-even requirement

$$p(e^*)R=L. \tag{4}$$

While (3) and (4) together may have multiple solutions, we pick in what follows the pair (R,e^*) that has the lowest value R and, consequently, achieves the highest payoff for the agent. Clearly, this is the unique equilibrium in a game where either outside investors compete or the agent makes a take-it-or-leave-it offer. Further, it is immediate that e^* is then strictly decreasing in L while R is strictly increasing.

[8] As we discuss below, the right to these tax receipts could then be sold *ex-ante* so as to alleviate financial constraints (at least for some types).

The Surplus Function. Denote

$$\omega = p(e^*)(x - R) - c(e^*),$$

which given the binding break-even constraint (4) is the total expected surplus *net* of the funding expenditures. By the previous discussion we can write this as a function of L: $\omega(L)$.

With $L = 0$ and thus $R = 0$, the agent would choose a first-best value solving $p'(e_{FB})x - c'(e_{FB}) = 0$, thereby realizing a total surplus of $\omega(0)$. Note that this is gross of externalities and all possible transfers. Clearly, it holds that $e^* < e_{FB}$ whenever $L > 0$. When we thus compare the total *net* surplus at the benchmark with $L = 0$ and at any other choice $L > 0$, we have generally

$$\omega(L) < \omega(0) - L.$$

This difference *beyond* the change in funding requirements captures the crucial inefficiency that arises from the outside financing problem. It follows as the agent shirks when he no longer realizes the full benefits from putting in higher effort. Observe further that from the break-even constraint (4) the agent is the residual claimant, so that it follows immediately that $\omega(L)$ is strictly decreasing in L: As more of the output has to be promised away as repayment to outside investors, the agent's incentives to exert effort further decline, resulting in a further reduction of efficiency. This captures the key inefficiency that arises from the agency problem due to non-observable effort and the need to raise outside finance.

General Contracts. So far, we restricted attention to deterministic contracts with outside investors, as described by the required repayment R. Given risk neutrality of both agents and outside investors, without loss of generality the most general, stochastic contract is described as follows. Note first that $(L, \omega(L))$ describes investors' and the agent's expected payoffs once financing is sunk (net of effort costs, but gross of initial capital L). A contract with investors could now prescribe the following, next to the initial provision of capital L: A distribution over values L^n with $E[L^n] = L$ so that when a particular value L^n is drawn, the contract that is then implemented stipulates for the investor the expected repayment of L^n and, consequently, the expected payoff $\omega(L^n)$ for the agent. Clearly, by optimality the chosen lottery would maximize $E[\omega(L^n)]$ subject to $E[L^n] = L$. Denote the respective realized value by

$$\widehat{\omega}(L) = \max E[\omega(L^n)].$$

Importantly, while it is immediate that also $\widehat{\omega}(L)$ is strictly decreasing in L, it is also concave: $\widehat{\omega}''(L) \leq 0$ at points of differentiability. (It is continuous and differentiable almost everywhere.) The argument follows simply by contradiction: If this was not the case, then we could find from Jensen's inequality three values $L^1 < L < L^2$ and two probabilities β^1 and β^2 so that $L = L^1\beta^1 + L^2\beta^2$ while $\widehat{\omega}(L) < \widehat{\omega}(L^1)\beta^1 + \widehat{\omega}(L^2)\beta^2$, contradicting the asserted optimality of $\widehat{\omega}(\cdot)$ at L.

Lemma 1 *If the agent needs initial financing of L, then under the optimal contract that lets outside investors just break even he realizes the expected payoff (net of costs of effort) $\widehat{\omega}(L)$, which is continuous, strictly decreasing with $\widehat{\omega}'(L) < -1$, and concave.*

In what follows, we find it more convenient to suppose that already the payoff function $\omega(L)$, for which we did not use lotteries, is strictly concave. (Still, our subsequent results hold generally when we use, instead, $\widehat{\omega}(L)$.) For instance, we could stipulate that $p(e) = e$ and $c(e) = e^2/(2\gamma)$, where γ is then taken to be sufficiently large so as to ensure that $p(e) < 1$ holds in equilibrium. This example is also used further below for an illustration.

4 Linear Tax and Tradable Pollution Rights

4.1 Preliminary Results

In this Section, we consider the following problem that a government faces. The government can choose the parameters of a linear tax policy. That is, depending on the volume of produced externalities, y_i, each agent is taxed according to the function[9]

$$\tau(y) = \tau_0 + \tau_1 y. \tag{5}$$

Here, τ_1 is the per-unit tax on the externality, while the fixed component τ_0 takes into account the overall distribution that is achieved by (optimally) making the government's budget just balance:

$$\tau_0 + \tau_1 \int y_i di = 0. \tag{6}$$

[9] Strictly speaking, the tax schedule is an (affine) two-part tariff.

Our motivation for the restriction to such a linear tax is the following. First and foremost, as we argue in more detail below, such a scheme corresponds to the implementation of a system of tradable pollution rights. In that case, the government's choice parameter would be the aggregate volume of externalities. Taking as a benchmark the outcome where such a system of tradable pollution rights is in place, we later argue how this can be optimally complemented with additional policies, such as tax-subsidized loans. Further, the characterization of the optimal linear tax will make transparent how both the presence of financing constraints and agent heterogeneity generally affect the implications of taxing externalities and, thereby, the optimal level and form of government intervention.

Without loss of generality, we stipulate that the agent must "purchase" the respective pollution rights (or pay the tax) when starting production in $t = 0$. Consequently, without resources on his own, an agent must raise outside finance equal to

$$L(y, \theta) = \max \{I_0 + K(\overline{y} - y, \theta) + \tau(y), 0\}. \tag{7}$$

Note that, from $\omega'(L) < -1$, it is not optimal for the agent to raise outside finance for consumption while it is equally optimal to use all of his own resources to reduce the amount of external financing raised.

Optimal Abatement. Given the tax scheme $\tau(y) = \tau_0 + \tau_1 y$, we consider first the program of an individual agent. The agent chooses y_i and, consequently, has to raise $L(y_i, \theta_i)$, as given by (7). Dropping the subscript i, an agent of type θ thus chooses y to maximize $\omega(L(y, \theta))$ with $L(y, \theta)$ given by (7).

Lemma 2 *Suppose the government imposes a (budget-balancing) tax-cum-transfer $\tau(y) = \tau_0 + \tau_1 y$. Then, an agent of type θ chooses the optimal level of externalities $y^*(\theta)$ and thus a unique level of abatement $a^*(\theta) = \overline{y} - y^*(\theta)$ so that*

$$K_1(a^*(\theta), \theta) = \tau_1, \tag{8}$$

from which $a^*(\theta)$ is strictly increasing in both τ_1 and θ. Still, higher-type agents invest less in abatement and thus need to raise less outside finance:

$$\frac{dL(\cdot)}{d\theta} = K_2(a^*(\theta), \theta) < 0. \tag{9}$$

Proof. See Appendix.

Hence, with a linear tax on externalities, each agent chooses a level of abatement so that the marginal *financial* benefits that follow from a reduction in the incurred tax are equal to the marginal cost of abatement. Importantly, productive efficiency, as expressed through the slope ω', plays *no role* in this trade-off. Moreover, note that, under the agent's optimal choice, his need to raise outside finance is always strictly *decreasing* in his type. In fact, as the agent chooses his privately optimal level of abatement, depending on θ, this follows immediately from optimality, as otherwise higher-type agents could not enjoy a higher expected utility $\omega(L)$.

4.2 Optimal Tax

The objective function of a utilitarian government is to maximize the expected utility of all agents:

$$E[u_i] = \int [\omega(L_i) - \rho y_i] \, di.$$

(This uses that $\omega(L_i)$ already takes into account the investment costs, as these are funded by outside investors.) Given the agent's optimal decision, using Lemma 2, the government's program is then to maximize

$$E[u_i] = \int_\Theta [\omega(L(y^*(\theta), \theta)) - \rho y^*(\theta)] \, dF(\theta) \tag{10}$$

subject to the budget-balancing constraint (cf. (6))

$$\tau_0 + \tau_1 \int_\Theta y^*(\theta) dF(\theta) = 0. \tag{11}$$

Take for a moment the benchmark without financial constraints, so that everywhere $\omega'(\cdot) = -1$. Then, from substitution of (11) into (10) while using the agent's first-order condition (cf. Lemma 2) we would obtain the Pigou rule $\tau_1 = \rho$. This obviously implements the first-best outcome, despite agents' private information about their marginal cost of abatement. The following result characterizes, instead, the optimal linear tax when agents must raise outside finance and when this gives rise to a deadweight loss due to agency problems.

Proposition 1 *The optimal linear per-unit tax τ_1 satisfies*

$$\tau_1 \left(- \int_\Theta \omega'(L(y^*(\theta), \theta)) dF(\theta) \right) = \rho - \frac{\int_\Theta \omega'(L(y^*(\theta), \theta)) \left[y^*(\theta) - \int_\Theta y^*(\theta') dF(\theta') \right] dF(\theta)}{\int_\Theta \frac{dy^*(\theta)}{d\tau_1} dF(\theta)}, \tag{12}$$

which implies that τ_1 is strictly smaller than ρ.

Proof. See Appendix.

From (12) an optimal tax τ_1 is strictly lower than ρ. This is so for two (albeit related) reasons. The first reason is captured by the multiplier

$$\left(- \int_\Theta \omega'(L(y^*(\theta), \theta)) dF(\theta) \right) > 1 \tag{13}$$

on the left-hand side (cf. Lemma 1 for $\omega' < -1$); given $\int_\Theta \frac{dy^*(\theta)}{d\tau_1} dF(\theta) < 0$, the second reason is captured by the term

$$\left(- \int_\Theta \omega'(\cdot) \left[y^*(\theta) - \int_\Theta y^*(\theta') dF(\theta') \right] dF(\theta) \right) > 0, \tag{14}$$

which is subtracted on the right-hand side in (12). We discuss both terms in turn.

The term (13) captures the fact that, to reduce externalities, agents must raise outside finance. Due to the associated agency problem this involves an additional "shadow cost", namely in the form of lower efficiency as effort becomes inefficiently low. (Formally, $\omega' < -1$.)

Next, the term (14) captures the efficiency implications of the redistribution of resources that goes hand-in-hand with the applied taxation, namely from agents with higher marginal abatement costs to agents with lower marginal abatement costs. The impact of redistribution on aggregate productive efficiency is negative. This follows from the following two observations: First, with a linear tax high-type agents incur, under the optimal choice $y^*(\theta)$, strictly lower costs of abatement; second, ω is strictly concave. As the tax on externalities shifts resources to high-type agents, this reduces the agency problem of high-type agents but increases the agency problem of low-type agents. Thus, it makes the already more productive high-type agents (endogenously) still more productive, while further reducing productivity of low-type agents. This leads to a reduction in *aggregate* efficiency of production in the economy.

4.3 Pollution Permit Scheme

It is straightforward to see that the government could implement the outcome of Proposition 1 also as follows. The government could set a total maximum capacity for externalities Y and allocate this uniformly (and for free) across all agents. Thus, each agent receives the same capacity, which we may write as $Y_i = Y$, as there is the measure one of agents in the economy. These capacities (or "pollution rights") are then traded in the market.

When τ_1 is the resulting price, we obviously have that $K_1(a^*(\theta), \theta) = \tau_1$, as previously in (8), together with

$$\int a^*(\theta) dF(\theta) = \overline{y} - Y.$$

This uniquely links τ_1 to Y, and vice versa. The equivalence of the two policy instruments can then be seen immediately from substituting into the funding retirement (7)

$$\begin{aligned} L(y^*(\theta), \theta) &= I_0 + K(\overline{y} - y^*(\theta), \theta) + \tau(y^*(\theta) - Y) \\ &= I_0 + K(\overline{y} - y^*(\theta), \theta) + \tau(y^*(\theta) - \int_\Theta y^*(\theta) dF(\theta)). \end{aligned}$$

This is just the same as under the linear tax, after substituting the "break-even" constraint for taxes (11).

Corollary 1 *The optimal linear tax can be implemented through a pollution permit scheme, where each agent receives the same level of tradable pollution rights Y.*

Roadmap to Further Analysis. Proposition 1 isolated two (related) reasons for why the optimal tax under financing constraints is strictly below the optimal Pigouvian tax: the shadow cost of raising financing, which is due to the agency problem, and the associated redistribution, which in the aggregate exacerbates this agency problem. In the following, we discuss ways how policy makers could reduce these inefficiencies and, thereby, optimally induce a higher level of aggregate abatement at less inefficiency.

The derivation of the optimal nonlinear tax in Section 5 further clarifies the tension between reducing externalities and redistributing resources, which in our case increases agency costs in the aggregate. We show that the optimal such nonlinear tax would essentially "dampen" the impact of the tax for "average polluters", namely through imposing high marginal taxes for both relatively low and relatively high levels of pollution. As we show in a numerical example, however, the resulting efficiency gains may make it optimal to, thereby, induce a strictly higher level of abatement for all agents.

Such a scheme of nonlinear taxes may, however, not always be feasible, in particular when a supranational market of tradable pollution rights is in place. As we show, the government could then still increase efficiency through introducing grants that are linked to the amount of outside financing that each agent raises. As the agent's type, namely the true (opportunity) costs of higher abatement, is only privately known, these grants must still be incentive compatible to forestall opportunistic behavior. As we argue, grants that are linked directly to the amount of financing that is raised are also more efficient than redistributing resources through taxes on final output. Still, for completeness we show in the Appendix that without loan-based grants, taxes on output can be optimal.

5 Nonlinear Taxes

So far we have restricted our analysis to a linear tax on externalities and identified the inefficiencies such a scheme causes when firms are financially constrained. First, the tax-induced additional abatement costs increase the required amount of external financing, thus exacerbating the agency problem vis-à-vis outside investors. Second, as firms have different (marginal) abatement costs, a tax-cum-transfer scheme leads to a redistribution from high to low-cost types resulting in a further decline in average productive efficiency again working through the same agency problem. As we show below it is this second inefficiency, caused by the reallocation of resources, that can be mitigated by relaxing the restriction to a linear tax on externalities. One way to counteract this resource reallocation, which is studied in this section, is thus to allow the government to implement a general (nonlinear) tax on externalities.

We thus depart from the assumption of a linear tax scheme or an equivalent choice of pollution rights that are sold. Consider thus a general tax $\tau(y_i)$ as a function of the respective externalities y_i that agent i produces. Our approach is the following. Instead of solving directly for the optimal nonlinear tax, we set up a general mechanism-design problem. (This is then extended in the subsequent Sections to introduce other policy instruments). Such a mechanism maps agents' truthful revelation of their type θ into both a prescribed level of externalities $y(\theta)$ and a respective transfer $T(\theta)$, which can be positive or negative. Once we have derived the optimal mechanism, we obtain from this the respective optimal tax scheme $\tau(y_i)$.

5.1 Control Problem

The mechanism must ensure truthtelling and thus incentive compatibility for each agent type.[10] That is, for all types $\theta \in \Theta$ it must hold that

$$\omega\left(L(y(\theta), T(\theta), \theta)\right) \geq \omega\left(L(y(\widehat{\theta}), T(\widehat{\theta}), \theta)\right) \text{ for all } \widehat{\theta} \in \Theta, \tag{15}$$

where

$$\begin{aligned} L(y(\theta), T(\theta), \theta) &= I_0 + K(\overline{y} - y(\theta), \theta) + T(\theta), \\ L(y(\widehat{\theta}), T(\widehat{\theta}), \theta) &= I_0 + K(\overline{y} - y(\widehat{\theta}), \theta) + T(\widehat{\theta}). \end{aligned}$$

In words, type θ must not strictly prefer to pretend to be any other type $\widehat{\theta}$.

It is convenient to express the following optimization problem purely in terms of (permitted) externalities $y(\theta)$, rather than optimal abatement $a(\theta)$. As is standard, we will employ optimal control techniques in what follows, for which we restrict the mechanism $\{(y(\theta), T(\theta))\}$ to piecewise continuously differentiable functions.[11] The incentive constraint (15) holds locally if "truthtelling", i.e., $\widehat{\theta} = \theta$, solves the respective first-order condition:

$$\left.\frac{d\omega\left(L(y(\widehat{\theta}), T(\widehat{\theta}), \theta)\right)}{d\widehat{\theta}}\right|_{\widehat{\theta}=\theta} = \omega'(\cdot)\left(T'(\theta) - y'(\theta)K_1(\overline{y} - y(\theta), \theta)\right) = 0. \tag{16}$$

We presently assume that the "first-order approach" is valid, so that (16) is sufficient to ensure global incentive compatibility. As is immediate from the single-crossing property (2), note that this requires also that the characterized function $y(\theta)$ be nonincreasing.[12]

Define now (with some abuse of notation) the payoff function *under truthtelling*[13]

$$U(\theta) = \omega\left(L(y(\theta), T(\theta), \theta)\right).$$

[10] As is well-known, under the considered environment the restriction to direct, truthtelling mechanisms follows without loss of generality from the "revelation principle".

[11] Note that incentive compatibility alone requires that $y(\theta)$ has to be nonincreasing and, hence, differentiable almost everywhere.

[12] Compare Appendix B for a more general solution ("second-order approach") allowing for the possibility of "bunching".

[13] This still presumes that taxes and subsidies are fully "used" to increase or reduce the amount of funds that must be raised externally (instead of being immediately consumed or saved for consumption in the final period). As agents are not impatient and have risk neutral preferences, this restriction is without loss of generality.

We know that by incentive compatibility $U(\theta)$ is nondecreasing and continuous and thus a.e. continuously differentiable with

$$\frac{dU(\theta)}{d\theta} = \left.\frac{\partial\omega\left(L(y(\widehat{\theta}), T(\widehat{\theta}), \theta)\right)}{\partial\theta}\right|_{\widehat{\theta}=\theta} = \omega'(\cdot)K_2(\overline{y} - y(\theta), \theta) > 0 \tag{17}$$

when $y(\theta) > 0$.

To solve for the optimal menu we set up the government's optimal control problem. With some abuse of notation define the financing requirement under truthtelling

$$L(\theta) = I_0 + K(\overline{y} - y(\theta), \theta) + T(\theta),$$

which we take as the *state variable*. As thus $U(\theta) = \omega(L(\theta))$, we have from (17) that

$$\frac{dL(\theta)}{d\theta} = K_2(\overline{y} - y(\theta), \theta) < 0. \tag{18}$$

Further, from

$$T(\theta) = L(\theta) - [I_0 + K(\overline{y} - y(\theta), \theta)] \tag{19}$$

we can substitute pointwise for $T(\theta)$, once the state variable $L(\theta)$ as well as $y(\theta)$ have been determined. This leaves us with the single control variable $y(\theta)$.

Summing up, the government's objective is thus to maximize total utility over all agents

$$\int_\Theta [\omega(L(\theta)) - \rho y(\theta)]\, dF(\theta) \tag{20}$$

subject to the "law of motion" (18) and the budget balance condition

$$\int_\Theta [L(\theta) - K(\overline{y} - y(\theta), \theta) - I_0]\, dF(\theta) = 0, \tag{21}$$

where we have substituted from (19).

5.2 Characterization

We now relegate to the Appendix the formulation of the respective Hamiltonian and the solution of the control problem. There, we also translate the solution into the optimal tax schedule. This is obtained from the characterized menu through substituting $\tau(y) = T(\theta(y))$, where we use $\theta(y) = y^{-1}(y(\theta))$.[14] Denote now the lowest and highest realized level of externalities by

$$y_l = y(\overline{\theta}) < y_h = y(\underline{\theta}).$$

[14] Note for this that presently we assume that the optimal mechanism prescribes a *strictly* decreasing level of externalities, with $y'(\theta) < 0$. See Appendix B for the case with "bunching".

Proposition 2 *Under the optimal nonlinear tax $\tau(y)$ the marginal tax rate is strictly positive, but strictly less than ρ, and highest at the two extremes, y_l and y_h. This implies that when the generated externality is already low (and abatement thus high), then the marginal benefits from further reducing pollution are strictly increasing. Instead, when the externality generated by the agent is still high (and abatement thus low), then the reduction in taxes achieved by limiting pollution is strictly decreasing.*

Proof. See Appendix.

While the inefficiency caused by the need to raise outside financing implies that $\tau'(y) < \rho$ for all levels of externalities, it holds from Proposition 2 that the optimal nonlinear tax rewards a reduction of externalities in particular at very high and very low realizations, i.e., at the "first units" and the "last units". The intuition for this result is as follows: At the heart is the attempt to restrict the redistribution that is made to high-type agents as a consequence of the tax on externalities.[15] Recall that redistribution is higher when the marginal tax is higher. Obviously, at the lower boundary $\underline{\theta}$, there is no need to further distort the implemented choice of externalities, as there is "no one" contributing to redistribution below $\underline{\theta}$. At the other extreme, when $\theta = \overline{\theta}$, there is also clearly no longer a benefit from further reducing the marginal tax since there is nobody benefitting from redistribution above $\overline{\theta}$.[16] These observations explain the derived properties of the tax scheme $\tau(y)$ at low and high levels of the externality.

One immediate implication of Proposition 2 is that a constant marginal tax and, therefore, also a system of tradable pollution rights alone is not optimal, at least not in our model. The reason is that it implies too much redistribution of resources away from firms with high abatement costs, which in our model increases aggregated deadweight loss in the economy. As discussed in the Introduction, if a system of tradable pollution rights is in place, e.g., due to an international or supranational agreement, then Proposition 2 would

[15]The intuition behind this result is similar to findings in the optimal income taxation literature (cf. Mirrlees (1971), Seade (1977, 1982)), according to which the optimal marginal income tax is zero at the endpoints of the income scale (in the absence of bunching and with a bounded distribution of skills) and strictly positive elsewhere. There, the tradeoff is between a redistributive gain and a negative incentive effect of a non-zero marginal tax.

[16]Recall that clearly also for the lowest and highest types the implemented choice of externalities will be below the first-best benchmark (or similarly the marginal tax rate strictly below ρ) due to the inefficiency caused by the need to raise external financing. However, as argued above, there is no need to further distort the level of externalities in order to reduce redistribution.

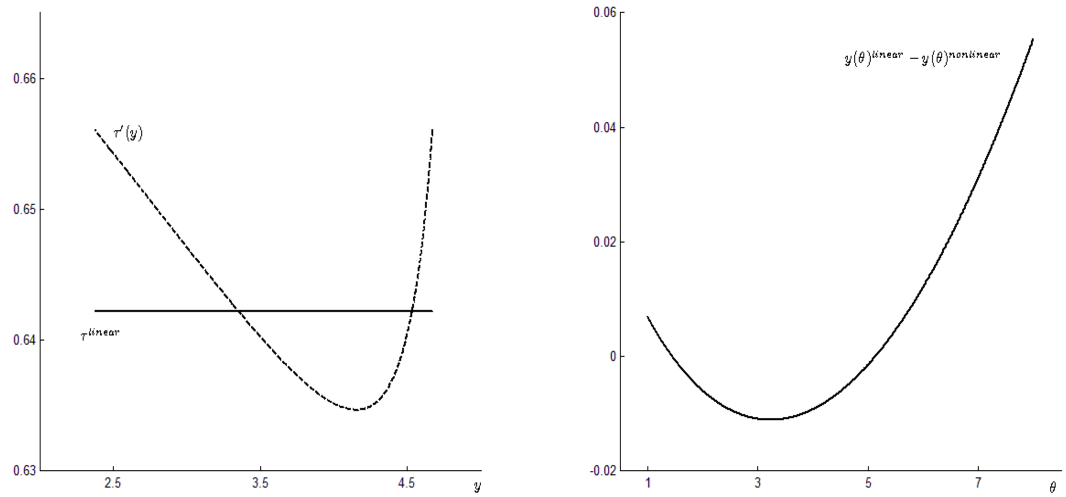

Figure 1: Example with high degree of heterogeneity

suggest that an additional, non-linear scheme of taxes and subsidies, depending on the realized externality, would improve efficiency. To achieve the benefits of the optimal non-linear tax, it should essentially dampen the impact on the "average polluter" but increase the impact on high-level and low-level polluters.

5.3 Illustration

In order to gain more intuition about the characterization of the optimal nonlinear tax scheme, suppose now that $p(e) = e$, $c(e) = \frac{e^2}{2\gamma}$, and $K(a) = \frac{1}{\alpha}\frac{a^2}{\theta}$, with θ distributed uniformly on $[\underline{\theta}, \overline{\theta}]$ ($\overline{\theta} > \underline{\theta} > 0$). It is straightforward to derive the agent's optimal effort $e^* = \gamma\,(x - R)$. We suppose that γ is always sufficiently small so as to ensure an interior solution $e^* = p(e^*) < 1$ for the probability of success. Together with the investor's break-even condition, after solving for the equilibrium repayment requirement R, we obtain for the "surplus function"

$$\omega(L) = \frac{\gamma}{8}\left(x + \left(x^2 - \frac{4}{\gamma}L\right)^{\frac{1}{2}}\right)^2,$$

which is, for $L > 0$, strictly decreasing with $\omega'(L) < -1$ and strictly concave. (The restriction that $x^2 - 4\frac{L}{\gamma} \geq 0$ ensures financial feasibility, for given L.) For this example, the respective solutions for the linear and nonlinear taxes are derived numerically.

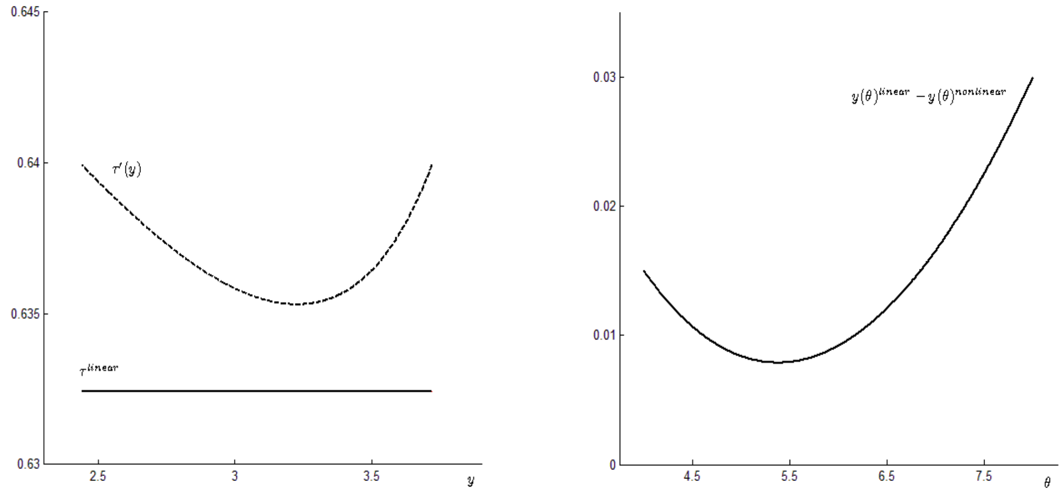

Figure 2: Example with low degree of heterogeneity

Figure 1 shows the marginal tax rate under the nonlinear scheme, together with the optimal linear tax rate (left panel) as well as the difference in generated externalities under the optimal linear tax and the optimal nonlinear tax scheme (right panel) for a particular specification: $\overline{\theta} = 8$, $\underline{\theta} = 1$, $x = 10$, $\alpha = 1$, $\gamma = 0.1$, $I_0 = 0.5$, and $\rho = 0.75$. We refer to this as a case with a (relatively) high degree of heterogeneity in types, as the difference $\overline{\theta} - \underline{\theta}$ is large compared to the case that we characterize further below.

The optimal nonlinear tax rate is U-shaped and maximized at the lowest and highest realized level of externalities, for which it is also strictly above the optimal linear rate. In this example, the optimal marginal tax rate at intermediate levels of externalities is, however, strictly below the optimal linear tax. As the second panel in Figure 1 shows, for intermediate types the resulting level of externalities is also strictly lower under the optimal linear tax than under the optimal nonlinear tax. This contrasts with our second example (Figure 2), where the marginal tax rate is strictly higher for all levels of externalities and where all types generate lower externalities compared to the case with the optimal linear tax. Compared to Figure 1, for Figure 2 we only change the low boundary of types from $\underline{\theta} = 1$ to $\underline{\theta} = 4$, thereby reducing the heterogeneity between agents.

6 Loan-Based Grants

So far, policy intervention was restricted to a tax on externalities. As discussed in the Introduction, governments frequently use also grants linked to loans as a way to steer firm behavior, in particular in the context of environmental policies.

We consider the following extension to our present model. We now allow government policy to be made contingent on the amount of financing that is raised by each agent. Importantly, note that the government can not directly verify the real cost of abatement $K(\cdot)$. As noted above, even when a loan-based grant is tied to specific expenditures, e.g., for particular equipment, the overall costs, including opportunity costs, could still substantially differ between firms. In essence, what we use in what follows is the restriction that while the actual loan that a firm raises is verifiable, its true financial needs are still the firm's private information. Still, loan-based grants will prove effective in improving efficiency.

Suppose thus that agents are taxed on their externality according to a constant marginal rate τ_1 or that, likewise, τ_1 is the prevailing price for "pollution rights". As noted above, our present restriction to such a linear tax is warranted in circumstances where nonlinear taxes are not feasible due to the existence of a supranational system of tradable pollution rights. Still, in Appendix C we also solve for the optimal mechanism when both nonlinear taxes and loan-based grants can be used.

With a linear tax or a market for tradable pollution rights with fixed capacity, the optimal level of abatement for each agent satisfies the first-order condition

$$K_1(\overline{y} - y^*(\theta), \theta) = \tau_1. \tag{22}$$

That is, irrespective of the firm's overall financial needs, including other taxes or grants, the firm simply chooses $y^*(\theta)$ to minimize expenditures. Taking for now τ_1 as given, an additional instrument thus serves the purpose of reducing the redistribution of resources that is generated by $\tau_1 > 0$. A loan-based grant (or tax) stipulates a positive (or negative) payment $G(L_i)$, given an agent's loan volume L_i. As previously, it proves, however, convenient to first set up the problem in the language of mechanism design.

Truthful Mechanism Such a mechanism now specifies for each agent a loan level $L(\theta)$ together with a payment *made* by each agent $t(\theta)$. It is immediate that under the optimal mechanism, $L(\theta)$ will *just* be sufficient to cover the agent's true expenditures. However, when an agent deviates and mimics another type $\widehat{\theta}$ by raising a higher-than-necessary loan, then his payoff becomes[17]

$$U(\theta,\widehat{\theta}) = \omega(L(\widehat{\theta})) + \left[L(\widehat{\theta}) - I_0 - K(\overline{y} - y^*(\theta),\theta) - \tau_1 y^*(\theta) - t(\widehat{\theta})\right]. \tag{23}$$

Here, the term in rectangular brackets captures the amount of financing that is raised above the true financial needs, which are unobservable to the government.[18] The mechanism is locally incentive compatible if $U(\theta,\widehat{\theta})$ is maximized at $\widehat{\theta} = \theta$, so that the additional term in (23) is zero: $U(\theta) = \omega(L(\theta))$.

When $t(\theta)$ can not condition on θ, as was the case in Section 4, then we have that

$$\frac{dU(\theta)}{d\theta} = \omega'(L(\theta))L'(\theta) = \omega'(L(\theta))K_2(\overline{y} - y^*(\theta),\theta). \tag{24}$$

This describes the slope of agents' utility and thus the extent to which a pure tax on externalities leads to a redistribution of resources in the economy. (Clearly, when $y^*(\theta) = \overline{y}$ holds for all agents, then $U(\theta) = \omega(I_0)$ is constant.) We show now how by linking $t(\theta)$ to $L(\theta)$, this slope can be reduced, so that resources are more evenly distributed across agents.

We know from incentive compatibility that $U(\theta)$ must *always* be non-decreasing (in fact, it is strictly increasing when $y(\theta) > 0$). More precisely, this holds for any choice of $t(\theta)$ that is still incentive compatible. Therefore, the government wants to make the derivative $t'(\theta) > 0$ for all θ as steep as possible. (Recall that we frame $t(\theta)$ as a payment from agents, in analogy to our previous approach.) The upper boundary on $t'(\theta) > 0$ is obtained from the first-order condition of the agent's reporting problem.[19] Making use of (22), this yields for the slope of agents' payoff

$$\frac{dU(\theta)}{d\theta} = \omega'(L(\theta))L'(\theta) = t'(\theta) - L'(\theta) = -K_2(\overline{y} - y^*(\theta),\theta), \tag{25}$$

which, for a given schedule $y^*(\theta)$ is clearly smaller than the slope in (24), showing the benefits of linking transfers to the amount of financing raised. Similarly, one obtains for

[17] Note that we have dropped any fixed part τ_0 for the tax on externalities, as this can be subsumed into $t(\theta)$.

[18] Note again, that we assume here that, for a firm of type θ, $y^*(\theta)$ is fixed from (22) independent of the report $\widehat{\theta}$. Still, the results of this Section continue to hold if we, instead, assumed that a firm of type θ, reporting $\widehat{\theta}$, would have to implement $y^*(\widehat{\theta})$.

[19] More precisely, this only considers a *feasible* deviation to a type $\widehat{\theta} < \theta$.

the slope of the transfer $t(\theta)$,[20]

$$t'(\theta) = L'(\theta)\left[1 + \omega'(L(\theta))\right]. \tag{26}$$

Clearly, when $\omega' = -1$, as in the case without an agency problem of external financing, then $t'(\theta) = 0$. Otherwise, we have that $t'(\theta) > 0$. In fact, the strictly larger is $\omega'(L(\theta)) < -1$ in absolute terms, the steeper can the loan-based grant be made, for the moment as a function of θ. (We translate this back into a function $G(L)$ below.) Such a transfer *back* to low-type agents, now linked to a loan, is made possible precisely as raising outside finance generates a deadweight loss. This holds also when more outside finance is raised than actually needed or, in a more general context, when a firm would take out a subsidized loan to fund abatement activities even though this would not be necessary, given the firm's resources. The deadweight loss of raising (too much) external finance prevents agents with lower costs of abatement to claim a higher grant that is intended for agents with higher marginal costs of abatement.

Optimal Loan-Based Grant Scheme. Under a grant scheme $G(L)$, we must have that

$$G(L(\theta)) = -t(\theta). \tag{27}$$

Differentiating (27) and substituting from (26), we have that

$$G'(L) = -\left[1 + \omega'(L(\theta))\right] > 0, \tag{28}$$

which tells us how under the incentive compatible grant scheme the absolute subsidy $G(L)$ varies with the loan size L.

Proposition 3 *Suppose that a constant marginal tax τ_1 on externalities is in place. When the loan size L_i is verifiable, then the government can strictly increase efficiency by introducing, in addition, a loan-based grant. The optimal grant $G(L)$ is strictly increasing in loan size with $G'(L)$ given by (28).*

Proof. See Appendix.

Efficiency can thus be improved by linking transfers to the outside financing that agents raise to cover their abatement costs. By combining a tax on externalities with grants that

[20] An alternative way of expressing this is to substitute for $L' = K_2 + t'$, so that $t' = -K_2 \frac{1+\omega'}{\omega'}$.

are linked to loans, the redistribution that is generated by the tax on externalities can be mitigated. Agents with lower abatement costs find it less profitable to mimic those with higher abatement costs, so as to claim additional grants. This follows as raising more-than-necessary outside finance is costly as it exacerbates the agency problem vis-à-vis outside investors. As a consequence, combining taxes on externalities with grants linked to loans - as is frequently observed in practice - dominates, in our setting, pure linear taxes or pollution permits.

In practice, even though economic instruments such as tradable permits or emission taxes are used to regulate emissions, they are often applied in combination with other interventions. Frequently used additional instruments of support are subsidies to environmentally friendly investment. As noted in the Introduction, for instance, in Germany a state-owned bank (KfW) provides on a large scale subsidized credit to businesses that apply energy-saving technologies or invest to reduce CO_2 output. The UK government, in turn, is just about to set up a "Green Investment Bank" which will provide investment subsidies and low-interest loans to accelerate private-sector investment in environmentally friendly infrastructure.

Discussion: Taxes on Output. As noted above, we have throughout restricted potential redistribution of resources to the initial stage. As we argue now, however, this is without loss of generality once we allow for loan-based grants. Consider thus taxes on output. The right to these taxes could then be sold ex-ante so as to alleviate (at least for some types) the external financing constraint.[21] At first, it may seem that this gives the government an additional lever for redistribution, as agents with lower abatement costs and thus lower financing needs end up with a higher equilibrium probability of success. However, the link from the agent's type to the likelihood of realizing high vs. low output is only indirect, namely through the agency problem and, therefore, through the amount of outside financing that is raised by the agent. A loan-based grant thus provides such a redistribution more directly through linking transfers to financing.[22] We show in the

[21] We discuss these options explicitly in Appendix D, where we consider the case of taxes on output *without* a loan-based grant.

[22] In this sense, in the presence of loan-based grants, output does not provide an additional "tag" that would be optimally used for transfers. Originally, the term "tagging" was coined by Akerlof (1978) to describe the use of taxes contingent on personal characteristics in order to improve on a purely income-based tax scheme (cf. also Mankiw and Weinzierl 2010 for a recent application).

Appendix how a tax on output could, however, increase efficiency when we assume that loan-based grants are not feasible.

7 Concluding Remarks

This paper analyzes the optimal policy towards externalities in the light of two constraints. First, agents who generate such (negative) externalities must raise outside finance when they want to increase their abatement. This generates inefficiencies due to agency costs vis-à-vis outside investors. Second, marginal abatement costs are private information, so that policies must be incentive compatible. We generate three sets of results in a simple, highly stylized model with these features. Our first result is that the optimal linear tax is strictly smaller than the benchmark Pigou tax, which would be equal to the marginal benefits from lowering externalities. In fact, we isolate two reasons for why this is the case: first, raising the necessary finance generates productive inefficiencies; second, with heterogeneous agents a higher tax generates aggregate productive inefficiencies as it leads to a redistribution of resources, thereby exacerbating aggregate financial frictions.

Our second result is that a nonlinear tax on externalities enhances efficiency, as it allows to achieve a given aggregate level of abatement more efficiently. As we show, this is the case as the nonlinear tax allows to limit redistribution of resources to higher-type agents with lower marginal abatement costs. We further show that under the optimal nonlinear tax the marginal benefits of abatement are highest for low and high levels of abatement (the "first units" and the "last units"). Importantly, this can not be implemented by a scheme of tradable "pollution rights".

Our third result is that the government can further improve efficiency by linking transfers to the outside finance that agents raise so as to (purportedly) reduce externalities. In contrast to nonlinear taxes, it can also be used when there is a (supranational) system of tradable pollution rights in place, which essentially implements a linear tax on externalities. Though agents with lower abatement costs can still mimic those with higher abatement costs, when additional grants are linked to credit, this becomes more costly, simply as raising more-than-necessary outside finance exacerbates the agency problem vis-à-vis outside investors. As a consequence, using jointly taxes on externalities and grants linked to loans - as is frequently observed in practice - may be an efficient instrument, as it allows to improve aggregate productive efficiency.

8 References

Akerlof , G., 1978. The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning. American Economic Review 68, 8-19.

Atkinson, A.B., Stern, N.H.,1974. Pigou, Taxation and Public Goods. Review of Economic Studies 41, 119-128.

Boadway, R., Tremblay, J.F., 2005. Public Economics and Start-up Entrepreneurship, in: Kanniainen, V., Keuschnigg, C. (eds.), Venture Capital, Entrepreneurship and Public Policy, CESifo Seminar Series, Cambridge: MIT Press, 198-219.

Boyer, M., Hiriart, Y., Martimort, D., 2007. Frontiers in the Economics of Environmental Regulation and Liability. Ashgate.

Boyer, M., Laffont, J.-J., 1997. Environmental Risks and Bank Liability. European Economic Review 41, 1427-1459.

Diamond, P.A., Mirrlees, J.A., 1971. Optimal Taxation and Public Production II: Tax Rules. American Economic Review 61, 261-278.

Guesnerie, R., Laffont, J.-J., 1984. A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm. Journal of Public Economics 25, 329-369.

Ebert, U., 1992. A Reexamination of the Optimal Nonlinear Income Tax. Journal of Public Economics 49, 47-73.

Jacobs, B., de Mooij, R.A., 2011. Pigou Meets Mirrlees: On the Irrelevance of Tax Distortions for the Second-Best Pigouvian Tax. CES Ifo Working Paper 3342.

Kaplow, L., 2006. Optimal Control of Externalities in the Presence of Income Taxation. NBER Working Paper 12339.

Mankiw, N.G., Weinzierl, M., 2010. The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution. American Economic Journal: Economic Policy 2, 155-176.

Mirrlees, J.A., 1971. An Exploration in the Theory of Optimum Income Taxation. Review of Economic Studies 43, 175-208.

Pigou, A., 1920. The Economics of Welfare. London: Macmillan.

Pitchford, R., 1995. How Liable Should a Lender Be? The Case of Judgement-Proof Firms and Environmental Risk. American Economic Review 85, 1171-1186.

Rothschild, C., Scheuer, F., 2013. Optimal Taxation with Rent-Seeking. Working Paper Stanford University.

Seade, J.K., 1977. On the Shape of Optimal Tax Schedules. Journal of Public Economics 7, 203-235.

Seade, J.K., 1982. On the Sign of the Optimum Marginal Income Tax. Review of Economic Studies 49, 637-643.

Tirole, J., 2010. From Pigou to Extended Liability: On the Optimal Tax of Externalities Under Imperfect Financial Markets. Review of Economic Studies 77, 697-729.

Appendix

A: Omitted Proofs

Proof of Lemma 2. We have

$$\frac{d\omega}{dy} = (\tau_1 - K_1(\overline{y} - y, \theta))\,\omega'\left(L(y,\theta)\right).$$

As the problem is strictly quasiconcave, this yields the optimality condition (8). From implicit differentiation and using (2) we have further that

$$\begin{aligned}
\frac{da^*(\theta)}{d\theta} &= -\frac{K_{12}(a^*(\theta),\theta)}{K_{11}(a^*(\theta),\theta)} > 0, \\
\frac{da^*(\theta)}{\tau_1} &= \frac{1}{K_{11}(a^*(\theta),\theta)} > 0.
\end{aligned}$$

Finally, expression (9) follows from substituting the first-order condition (8) into the total derivative of $L(\cdot)$. **Q.E.D.**

Proof of Proposition 1. We can substitute from (11) to obtain for each type the financing requirement

$$L(y^*(\theta),\theta) = I_0 + K(\overline{y} - y^*(\theta),\theta) + \tau_1 y^*(\theta) - \tau_1 \int_\Theta y^*(\theta')dF(\theta'),$$

so that $\frac{dE[u_i]}{d\tau_1}$ equals

$$\int_\Theta \left[\begin{array}{c} \left(\left(y^*(\theta) - \int_\Theta y^*(\theta')dF(\theta')\right) + \left(\frac{dy^*(\theta)}{d\tau_1}(\tau_1 - K_1(\cdot)) - \tau_1 \int_\Theta \frac{dy^*(\theta')}{d\tau_1}dF(\theta')\right)\right)\cdot\omega'(\cdot) \\ -\rho\frac{dy^*(\theta)}{d\tau_1} \end{array} \right] dF(\theta).$$

Substituting the first-order condition (8) for $y^*(\theta)$, $\tau_1 = K_1$, this gives rise to the first-order condition (12). From

$$\int_\Theta \frac{dy^*(\theta)}{d\tau_1}dF(\theta) < 0,$$

given that $y^*(\theta)$ is strictly decreasing, it remains to prove that

$$\int_\Theta \omega'(\cdot)\left[y^*(\theta) - \int_\Theta y^*(\theta')dF(\theta')\right]dF(\theta) < 0. \tag{29}$$

To see this, note first that, next to $\frac{dy^*(\theta)}{d\tau_1} < 0$, we have from expression (9) and Lemma 1 that

$$\frac{d}{d\theta}\left(\omega'\left(L(y,\theta)\right)\right) = \frac{dL(\cdot)}{d\theta}\omega'' > 0. \tag{30}$$

Define now the unique type $\widehat{\theta}$ satisfying

$$y^*(\widehat{\theta}) = E[y^*] = \int_\Theta y^*(\theta')dF(\theta'),$$

while $y^*(\theta) > E[y^*]$ holds for $\theta < \widehat{\theta}$ and $y^*(\theta) < E[y^*(\theta)]$ holds for $\theta > \widehat{\theta}$. We can now rewrite the left-hand side of (29) as

$$LS = \int_{\theta<\widehat{\theta}} \omega'(\cdot) \left[y^*(\theta) - E[y^*]\right] dF(\theta) + \int_{\theta>\widehat{\theta}} \omega'(\cdot) \left[y^*(\theta) - E[y^*]\right] dF(\theta). \qquad (31)$$

There, the terms in the first integral are all strictly negative and the terms in the second integral are all strictly positive. Given strict monotonicity of $\omega'(\cdot)$, we can thus derive the upper bound

$$\begin{aligned} LS &< \int_{\theta<\widehat{\theta}} \omega'(L(y,\widehat{\theta})) \left[y^*(\theta) - E[y^*]\right] dF(\theta) + \int_{\theta<\widehat{\theta}} \omega'(L(y,\widehat{\theta})) \left[y^*(\theta) - E[y^*]\right] dF(\theta) \\ &= \omega'(L(y,\widehat{\theta})) \left[\int_\Theta \left[y^*(\theta) - \int_\Theta y^*(\theta')dF(\theta')\right] dF(\theta)\right] = 0. \end{aligned}$$

This determines thus that the right-hand side of (12) is strictly smaller than ρ, so that together with the preceding argument we have indeed that $\tau_1 < \rho$. **Q.E.D.**

Proof of Proposition 2. The Hamiltonian is given by

$$H = \left[\omega(L(\theta)) - \rho y(\theta)\right] f(\theta) + \eta \left[L(\theta) - K(\overline{y} - y(\theta), \theta) - I_0\right] f(\theta) + \lambda(\theta) K_2(\overline{y} - y(\theta), \theta),$$

an optimal solution must satisfy the first-order condition for $y(\theta)$

$$f(\theta) \left[-\rho + \eta K_1(\overline{y} - y(\theta), \theta)\right] - \lambda(\theta) K_{12}(\overline{y} - y(\theta), \theta) = 0 \qquad (32)$$

and for the costate variable

$$\frac{\partial H}{\partial L} = -\lambda'(\theta) \;\Leftrightarrow\; f(\theta) \left[\omega'(L(\theta)) + \eta\right] + \lambda'(\theta) = 0. \qquad (33)$$

There are no terminal conditions, and the transversality conditions are given by

$$\lambda(\overline{\theta}) = 0, \qquad (34)$$

$$\lambda(\underline{\theta}) = 0. \qquad (35)$$

Using (34) and (35), we thus obtain from integrating (33)

$$\lambda(\theta) = \int_\theta^{\overline{\theta}} \left(\omega'(L(\vartheta)) + \eta\right) dF(\vartheta) = -\int_{\underline{\theta}}^\theta \left(\omega'(L(\vartheta)) + \eta\right) dF(\vartheta)$$

and

$$\eta = -\int_{\underline{\theta}}^{\overline{\theta}} \omega'(L(\vartheta))dF(\vartheta) > 1. \tag{36}$$

Here, η expresses the *marginal benefits* when the economy's resource constraint was marginally relaxed (e.g., by some initial endowment that could be allocated by the government).

From (18) and the concavity of $\omega(L(\theta))$, we have that $\omega'(L(\theta))$ is increasing in θ. Thus, making use of (36),

$$\lambda'(\theta) = f(\theta) \int_{\underline{\theta}}^{\overline{\theta}} \left(\omega'(L(\vartheta)) - \omega'(L(\theta))\right) dF(\vartheta)$$

is first positive and then negative, i.e., $\lambda(\theta)$ is first increasing and then decreasing. Clearly the transversality conditions (34) and (35) then imply that $\lambda(\theta) \geq 0$ holds everywhere, which we use in what follows.

Rearranging now the first-order condition for the control $y(\theta)$ in (32), we have

$$\eta K_1(\overline{y} - y(\theta), \theta) = \rho + \frac{\lambda(\theta)}{f(\theta)} K_{12}(\overline{y} - y(\theta), \theta), \tag{37}$$

which, using $K_{12}(\cdot) < 0$, $\lambda(\cdot) \geq 0$ and $\eta > 1$, implies first that

$$K_1(\overline{y} - y(\theta), \theta) < \rho.$$

That is, also with nonlinear taxes, externalities are for *all* types higher than under the Pigou tax. Moreover, note that it holds only at the boundaries $\underline{\theta}$ and $\overline{\theta}$ (when they are finite) that $\lambda(\theta) = 0$ and thus

$$\eta K_1(\overline{y} - y(\underline{\theta}), \underline{\theta}) = \rho \text{ and } \eta K_1(\overline{y} - y(\overline{\theta}), \overline{\theta}) = \rho. \tag{38}$$

Instead, for all other types we have the following:

$$\eta K_1(\overline{y} - y(\theta), \theta) < \rho \text{ for all } \theta \in (\underline{\theta}, \overline{\theta}). \tag{39}$$

Hence, under the optimal mechanism the marginal abatement costs are highest at the lowest and at the highest type, when evaluated at the respective choice $y(\theta)$. We will finally analyze how this translates into the respective nonlinear tax scheme $\tau(y_i)$. For this the following observation is also useful. We obtain

$$\begin{aligned} T'(\theta) &= L'(\theta) - K_2(\cdot) + y'(\theta)K_1(\cdot) \\ &= y'(\theta)K_1(\cdot) \leq 0. \end{aligned}$$

That is, the nonlinear tax on the externality still involves a transfer from low-type agents to high-type agents, given that $T'(\theta) \leq 0$ (and strictly so where $y'(\theta) < 0$).[23]

We now substitute $\tau(y) = T(\theta(y))$, where we use $\theta(y) = y^{-1}(y(\theta))$. Note for this that presently we assume that the optimal mechanism prescribes a *strictly* decreasing level of externalities, with $y'(\theta) < 0$. That is, while $y(\theta)$ must be non-increasing from incentive compatibility, there is also no "bunching" (cf. also Appendix B). From substituting the obtained characterization, we then have that

$$\begin{aligned} \tau'(y) &= T'(\theta)\frac{d\theta}{dy} = \frac{T'(\theta)}{y'(\theta)} = K_1(\overline{y} - y, \theta(y)) \\ &= \frac{1}{\eta}\left[\rho + \frac{\lambda(\theta(y))}{f(\theta(y))}K_{12}(\overline{y} - y, \theta(y))\right], \end{aligned} \tag{40}$$

such that the marginal tax is always (strictly) positive, but also strictly smaller than the Pigouvian tax.

Here, as discussed above, the term $\eta > 1$ (cf. expression (36)) applies to all types and creates a first wedge between the "Pigou tax" and the marginal tax with outside financing and agency costs. Turn now to the second term in rectangular brackets. As $K_{12} < 0$ (cf. the key "sorting condition" (2)) and as we obtained that $\lambda(\theta) \geq 0$, this term is negative and now captures the second rationale for why the optimal marginal tax is strictly lower, namely the inefficient redistribution of resources that goes along with the tax on externalities in our model. Note, however, that $\lambda(\underline{\theta}) = \lambda(\overline{\theta}) = 0$ holds at the boundaries, where under the optimal non-linear tax this effect no longer plays a role (cf. (38)). However, for all $\theta \in (\underline{\theta}, \overline{\theta})$ we have that $\lambda(\theta) > 0$.

From further differentiating, we obtain next

$$\eta\tau''(y) = -\frac{\lambda(\theta)}{f(\theta)}K_{112} + \frac{d\theta}{dy}\left[\frac{\lambda(\theta)}{f(\theta)}K_{122} + K_{12}\frac{d}{d\theta}\left[\frac{\lambda(\theta)}{f(\theta)}\right]\right]. \tag{41}$$

Expression (41) describes how the marginal tax changes. Recall once more from the transversality conditions (34)-(35) that at the boundaries $\underline{\theta}$ and $\overline{\theta}$ (when they are finite) we have $\lambda(\theta) = 0$. Recall that in the main text we have defined the lowest and highest realized level of externalities by $y_l = y(\overline{\theta}) < y_h = y(\underline{\theta})$. Further, recall that $\lambda' > 0$ for low

[23] In fact, incentive compatibility implies that in both cases, i.e., with linear and nonlinear taxes, the marginal tax w.r.t. the agent's type is given by $y'(\theta)K_1(\cdot)$. (For the linear tax we can use that $T'(\theta) = \tau_1 y'(\theta)$ and that $K_1(\cdot) = \tau_1$.)

and $\lambda' < 0$ for high θ, while $\frac{d\theta}{dy} < 0$ (when there is no bunching). Using then

$$\frac{d}{d\theta}\left[\frac{\lambda(\theta)}{f(\theta)}\right] = \frac{1}{f^2(\theta)}\left[f(\theta)\lambda'(\theta) - f'(\theta)\lambda(\theta)\right]$$

we have at the "endpoints"

$$\begin{aligned} \eta\tau''(y_l) &= \frac{d\theta}{dy}K_{12}\frac{\lambda'(\overline{\theta})}{f(\overline{\theta})} < 0, \\ \eta\tau''(y_h) &= \frac{d\theta}{dy}K_{12}\frac{\lambda'(\underline{\theta})}{f(\underline{\theta})} > 0. \end{aligned}$$

In words, under the optimal nonlinear tax $\tau(y)$, at very high levels of y the marginal tax $\tau'(y) > 0$ becomes strictly decreasing, while at very low levels of y the marginal tax is strictly increasing. This completes the characterization of the solution and the proof. **Q.E.D.**

Proof of Proposition 3. What remains to be shown is that the characterized mechanism is globally incentive compatible. Hence, we need to show that, for all θ and $\widehat{\theta}$ with $\theta > \widehat{\theta}$, it holds that

$$U(\theta) \geq U(\theta, \widehat{\theta}).$$

Using (23) we can rewrite this inequality to obtain

$$\omega(L(\theta)) - \omega(L(\widehat{\theta})) \geq K(\overline{y} - y^*(\widehat{\theta}), \widehat{\theta}) - K(\overline{y} - y^*(\theta), \theta) + \tau_1 y^*(\widehat{\theta}) - \tau_1 y^*(\theta).$$

From $U(\theta) = \omega(L(\theta))$ together with (25), we then can write equivalently

$$\begin{aligned} \int_{\widehat{\theta}}^{\theta} \frac{dU(\varphi)}{d\varphi} d\varphi &= \int_{\widehat{\theta}}^{\theta} -K_2(\overline{y} - y^*(\varphi), \varphi) d\varphi \\ &\geq -\int_{\widehat{\theta}}^{\theta} \frac{dK(\overline{y} - y^*(\varphi), \varphi)}{d\varphi} d\varphi - \tau_1 \int_{\widehat{\theta}}^{\theta} \frac{dy^*(\varphi)}{d\varphi} d\varphi, \end{aligned}$$

which can be simplified, using (22), to obtain

$$\begin{aligned} & -\int_{\widehat{\theta}}^{\theta} K_2(\overline{y} - y^*(\varphi), \varphi) d\varphi \\ \geq & -\int_{\widehat{\theta}}^{\theta} K_2(\overline{y} - y^*(\varphi), \varphi) d\varphi + \int_{\widehat{\theta}}^{\theta} \underbrace{(K_1(\overline{y} - y^*(\varphi), \varphi) - \tau_1)}_{=0} \frac{dy^*(\varphi)}{\varphi} d\varphi, \end{aligned}$$

establishing global incentive compatibility. **Q.E.D.**

B: Nonlinear Tax with Bunching

In the main text, we assume that the first-order approach applies and that $y(\theta)$, as characterized by (37), is nonincreasing. Differentiating (37) w.r.t. θ we obtain

$$\begin{aligned} & \eta\left(K_{12}(\overline{y}-y(\theta),\theta)-y'(\theta)K_1(\overline{y}-y(\theta),\theta)\right) \\ = \; & \frac{d}{d\theta}\left(\frac{\lambda(\theta)}{f(\theta)}\right)K_{12}(\overline{y}-y(\theta),\theta)+\frac{\lambda(\theta)}{f(\theta)}\left(K_{122}(\overline{y}-y(\theta),\theta)-y'(\theta)K_{112}(\overline{y}-y(\theta),\theta)\right) \end{aligned}$$

and thus

$$\begin{aligned} y'(\theta) \; = \; & \frac{1}{\left(\frac{\lambda(\theta)}{f(\theta)}K_{112}(\cdot)-\eta K_1(\cdot)\right)} \\ & \cdot\left[\frac{d}{d\theta}\left(\frac{\lambda(\theta)}{f(\theta)}\right)K_{12}(\cdot)+\frac{\lambda(\theta)}{f(\theta)}K_{122}(\cdot)-\eta K_{12}(\cdot)\right]. \end{aligned} \tag{42}$$

If we have for all $\theta \in \Theta$ that $y'(\theta)$, as characterized by (42), is non-positive, then the first-order approach characterizes the optimal nonlinear tax on externalities. If, however, $y'(\theta) > 0$ for some $\theta \in \Theta$, then the first-order approach is not valid and the monotonicity constraint on externalities ($y'(\theta) \leq 0$) has to be taken into account explicitly in the optimization program ("second-order approach").[24] This approach is shortly outlined in this Appendix. The resulting optimal menu will feature "bunching" for some subsets of Θ.

The problem is to maximize (20), subject to (21), (18), and

$$w(\theta) := \frac{dy(\theta)}{d\theta} \leq 0. \tag{43}$$

In the optimal control problem, we now take $w(\theta)$ as control and $y(\theta)$ as (additional) state variable. The Hamiltonian then reads

$$\begin{aligned} H \; = \; & \left[\omega(L(\theta))-\rho y(\theta)\right]f(\theta)+\eta\left[L(\theta)-K(\overline{y}-y(\theta),\theta)-I_0\right]f(\theta) \\ & +\lambda(\theta)K_2(\overline{y}-y(\theta),\theta)+\nu(\theta)w(\theta)-\kappa(\theta)w(\theta), \end{aligned}$$

and an optimal solution must satisfy the first-order condition

$$\frac{\partial H}{\partial w} = 0 = \nu(\theta)-\kappa(\theta), \tag{44}$$

as well as the costate equations

$$\frac{\partial H}{\partial L} = -\lambda'(\theta) = f(\theta)\left[\omega'(L(\theta))+\eta\right], \tag{45}$$

$$\frac{\partial H}{\partial y} = -\nu'(\theta) = f(\theta)\left[-\rho+\eta K_1(\overline{y}-y(\theta),\theta)\right]-\lambda(\theta)K_{12}(\overline{y}-y(\theta),\theta). \tag{46}$$

[24] Cf. Guesnerie and Laffont (1984) or, in the context of optimal nonlinear income taxation, Ebert (1992).

The transversality conditions are given by

$$\lambda(\overline{\theta}) = \lambda(\underline{\theta}) = 0, \tag{47}$$

$$\nu(\overline{\theta}) = \nu(\underline{\theta}) = 0, \tag{48}$$

and we have the complementary slackness condition

$$\kappa(\theta) \geq 0 \ (= 0,\ if\ w(\theta) < 0). \tag{49}$$

Note first that (45) is exactly the same as (33), so that we can conclude that, again, $\eta > 1$, and $\lambda(\theta) \geq 0$, with $\lambda'(\theta)$ first positive and then negative.[25] Further, if the monotonicity constraint is not binding on some interval, $w(\theta) < 0$, then we must have $\nu(\theta) = 0$ on this interval, and, thus, $y(\theta)$ is determined from

$$\eta K_1(\overline{y} - y(\theta), \theta) = \rho + \frac{\lambda(\theta)}{f(\theta)} K_{12}(\overline{y} - y(\theta), \theta), \tag{50}$$

which coincides with (37).

Now consider an interval $[\theta_0, \theta_1]$ where the monotonicity constraint (43) is binding.[26] Then, $y(\theta)$ is constant on this interval. Denote this value by $\widetilde{y}$. Integrating (46) and noting that $\nu(\theta_0) = \nu(\theta_1) = 0$ by continuity, we obtain

$$\int_{\theta_0}^{\theta_1} \left[\left[-\rho + \eta K_1(\overline{y} - \widetilde{y}, \theta) \right] f(\theta) - \lambda(\theta) K_{12}(\overline{y} - \widetilde{y}, \theta) \right] d\theta = 0.$$

This equation, together with the continuity requirement $y(\theta_0) = y(\theta_1) = \widetilde{y}$ jointly determines θ_0, θ_1, and $\widetilde{y}$. Finally, observe that the optimal nonlinear tax function $\tau(y)$ will then have a kink at $\widetilde{y}$. Else, it is still determined from (40); and Proposition 2 remains valid as long as there is no bunching at $\underline{\theta}$ or $\overline{\theta}$.

[25] Note, however, that the exact values of η and $\lambda(\theta)$ will in general differ from those in the case of no bunching.

[26] The extension to several possible intervals of bunching is straightforward.

C: Non-linear Tax with Loan-Based Grant

As in Section 6, government policy can be made contingent on both the level of externalities as well as the amount of financing that is raised by each agent. We stipulate now that the government is no longer restricted to a constant marginal tax on externalities. Thus, the mechanism prescribes, in addition to the transfer $T(\theta)$, also (a general) $y(\theta)$.[27] Denote by $U(\theta, \widehat{\theta})$ the payoff of a firm of type θ reporting $\widehat{\theta} \leq \theta$, which is given by

$$U(\theta, \widehat{\theta}) = \omega\left(L(\widehat{\theta})\right) + K(\overline{y} - y(\widehat{\theta}), \widehat{\theta}) - K(\overline{y} - y(\widehat{\theta}), \theta). \tag{51}$$

Again, we only consider downwards incentive compatibility constraints for this relaxed problem. (In equilibrium, deviating "upwards" by pretending to be a higher type will indeed not be feasible, given the prescribed $L(\theta)$.) There are no incentives to deviate locally when[28]

$$\frac{dU(\theta)}{d\theta} \geq \left.\frac{\partial U(\theta, \widehat{\theta})}{\partial \theta}\right|_{\widehat{\theta}=\theta} = -K_2(\overline{y} - y(\theta), \theta) > 0 \text{ when } y(\theta) > 0. \tag{52}$$

As the government will find it optimal to redistribute as much as possible to lower-type agents, by optimality this will hold with equality.[29] Comparing (52) with (17) clearly shows the benefits from linking transfers to the amount of financing raised by agents, in that, for any given schedule $y(\theta)$, the function $U(\theta)$ can be made flatter, thus, reducing inefficient redistribution to higher types.

Again, it is convenient to take

$$L(\theta) = I_0 + K(\overline{y} - y(\theta), \theta) + T(\theta)$$

as the state variable. Then, from $\frac{dU}{d\theta} = -K_2(\cdot)$ and $U'(\theta) = \omega'(L(\theta))L'(\theta)$ we have the law of motion

$$L'(\theta) = -\frac{K_2(\overline{y} - y(\theta), \theta)}{\omega'(L(\theta))}.$$

[27] Then, from $L(\theta) = I_0 + K(\overline{y} - y(\theta), \theta) + T(\theta)$, also the size of the loan is pinned down for any (truthfull) report θ. Note that, in contrast to Section 5, this also holds off equilibrium, i.e., when reporting to be of type $\widehat{\theta}$, the loan that can be raised is the same, $L(\theta, \widehat{\theta}) = L(\widehat{\theta})$, independent of the true type θ, as $L(\theta)$ is observable and can be prescribed by the government.

[28] We further require the optimal schedule $y(\theta)$ to be nonincreasing, which is assumed throughout this Appendix. We thus take a first-order approach.

[29] Formally, this can be shown by introducing an additional control that accounts for a possible slack in (52).

The Hamiltonian is given by

$$H = [\omega(L(\theta)) - \rho y(\theta)]\, f(\theta) + \eta\, [L(\theta) - K(\overline{y} - y(\theta), \theta) - I_0]\, f(\theta) - \lambda(\theta) \frac{K_2(\overline{y} - y(\theta), \theta)}{\omega'(L(\theta))},$$

so that the optimality condition for $y(\theta)$ is, in analogy to condition (37),

$$\eta K_1(\overline{y} - y(\theta), \theta) = \rho - \frac{\lambda(\theta)}{f(\theta)} \frac{K_{12}(\overline{y} - y(\theta), \theta)}{\omega'(L(\theta))}. \tag{53}$$

A comparison of (53) with (37) again shows the difference when transfers can be based on the (minimum) loan that agents raise. Then, for a given schedule $y(\theta)$ it is incentive compatible to make the strictly decreasing schedule of loan volume $L(\theta)$ flatter, namely through an equally flatter $T(\theta)$, thereby reducing redistribution to high-type agents.

To complete the characterization, we have

$$[\omega'(L(\theta)) + \eta]\, f(\theta) + \lambda(\theta) \frac{K_2(\overline{y} - y(\theta), \theta)\omega''(L(\theta))}{(\omega'(L(\theta)))^2} = -\lambda'(\theta)$$

next to the transversality conditions

$$\lambda(\overline{\theta}) = \lambda(\underline{\theta}) = 0.$$

Taken together, the resulting differential equation in λ can be integrated to obtain

$$\begin{aligned} \lambda(\theta) &= \int_{\theta}^{\overline{\theta}} \left[\exp\left(\int_{\theta}^{\vartheta} \frac{K_2(\overline{y} - y(\phi), \phi)\omega''(L(\phi))}{(\omega'(L(\phi)))^2} d\phi \right) [\omega'(L(\vartheta)) + \eta]\, f(\vartheta) \right] d\vartheta \\ &= -\int_{\underline{\theta}}^{\theta} \left[\exp\left(\int_{\theta}^{\vartheta} \frac{K_2(\overline{y} - y(\phi), \phi)\omega''(L(\phi))}{(\omega'(L(\phi)))^2} d\phi \right) [\omega'(L(\vartheta)) + \eta]\, f(\vartheta) \right] d\vartheta, \end{aligned}$$

which implies

$$\int_{\underline{\theta}}^{\overline{\theta}} \left[\exp\left(\int_{\underline{\theta}}^{\vartheta} \frac{K_2(\overline{y} - y(\phi), \phi)\omega''(L(\phi))}{(\omega'(L(\phi)))^2} d\phi \right) [\omega'(L(\vartheta)) + \eta]\, f(\vartheta) \right] d\vartheta = 0,$$

and, hence, together with $\omega'(\cdot) < -1$ that $\eta > 1$. Further $\lambda(\theta)$ is again hump-shaped.

D: Taxes on Output

As we argued in the main text, when loan-based grants are feasible there is no role for taxes on output. As we show now, however, in the absence of loan-based grants taxes on output together with taxes on externalities can improve efficiency by compensating for the redistribution that the latter tax induces in the first place. This reduces aggregate inefficiency that arises from external financing due to the underlying agency problem.

Suppose thus that the government specifies, in addition to a linear tax on externalities, also a tax $\overline{z} \leq x$ on (high) output as well as a (potentially negative) tax $\underline{z}$ in case output is low.[30] More generally, a contract with an investor now stipulates repayments (one potentially negative) in case of low or high output, $\underline{R}$ and $\overline{R}$, where $\overline{R} \leq x - \overline{z}$ and $\underline{R} \leq -\underline{z}$. The agent then chooses effort e to maximize

$$p(e)\left(x - \overline{z} - \overline{R}\right) - (1 - p(e))\left(\underline{z} + \underline{R}\right) - c(e),$$

giving rise to the first-order condition

$$p'(e^*)\left(x - (\overline{z} - \underline{z}) - \left(\overline{R} - \underline{R}\right)\right) - c'(e^*) = 0.$$

Our argument below will be restricted to show that there will always be some redistribution through an output tax. Consequently, we only consider (arbitrarily) small taxes. Then, as the total payoff from the enterprise is still almost completely distributed between the agent and the outside investor, it follows from standard arguments that optimally $\underline{R} = -\underline{z}$. Hence, with $\overline{R}$ as remaining contractual variable, an investor's break-even requirement is given by

$$p(e^*)\overline{R} - (1 - p(e^*))\,\underline{z} = L,$$

and we have the agent's expected surplus given by

$$\omega = p(e^*)\left(x - (\overline{z} - \underline{z})\right) - L - \underline{z} - c(e^*).$$

For notational simplicity define $z := (\overline{z} - \underline{z})$ and note that $\underline{z}$, which is used to reduce the financing requirement L, can be subsumed into the constant part of the tax on externalities τ_0. The aggregate budget balancing transfer τ_0 thus solves

$$\int_{\Theta} \left[\tau_0 + \tau_1 y^*(\theta) + zp(e^*(\theta)))\right] dF(\theta) = 0, \tag{54}$$

with $y^*(\theta)$ still given by (8).

[30] Importantly, this can not condition on the subsequently agreed repayment R with the respective investors, which would proxy for making the mechanism contingent on the size of external finance (cf., however, Section 6).

Optimal Output Taxation. Substituting for τ_0 from (54), the government solves the program

$$\max_{\tau_1, z} \int_\Theta \left[p(e^*(\theta))\,(x-z) - L(\theta) - c(e^*(\theta)) - \rho y^*(\theta)\right] dF(\theta)$$

with

$$L(\theta) = I_0 + K(\overline{y} - y^*(\theta), \theta) + \tau_1 y^*(\theta) - \int_\Theta \left[\tau_1 y^*(\theta) + z p(e^*(\theta))\right] dF(\theta).$$

The optimal tax rate on the externality and on output are then determined from

$$\tau_1 \int_\Theta \left[-\mu(\theta)\right] dF(\theta) = \rho - \frac{1}{\int_\Theta \left[\frac{\partial y^*(\theta)}{\partial \tau_1}\right] dF(\theta)} \int_\Theta \left[\mu(\theta)\left(y^*(\theta) - \int_\Phi y^*(\phi) dF(\phi)\right)\right] dF(\theta) \tag{55}$$

and

$$\int_\Theta \left[\mu(\theta)\left(p(e^*(\theta)) - \int_\Phi p(e^*(\phi)) dF(\phi)\right)\right] dF(\theta) = 0, \tag{56}$$

respectively, where we have used

$$\mu(\theta) := -\left(1 - \frac{\frac{p'(e^*(\theta))}{p(e^*(\theta))}\left(p'(e^*(\theta))x - c'(e^*(\theta))\right)}{p''(e^*(\theta))\left(x - \frac{L(\theta)}{p(e^*(\theta))}\right) - c''(e^*(\theta)) + \left(\frac{p'(e^*(\theta))}{p(e^*(\theta))}\right)^2 (L(\theta) - p(e^*(\theta))z)}\right) < -1.$$

We focus now on the optimal output tax. Further, as noted above, we restrict ourselves to showing that - in the absence of loan-based grants - there will always be such redistribution through output taxes.

Proposition. *In the absence of loan-based grants, the optimal output tax, z, will always be strictly positive.*

Proof. Note first that, for $z = 0$, we have $\mu(\theta) = \omega'(L(\theta))$. Hence, (55) simplifies to (12) for $z = 0$. To prove the Proposition, we show the following result: The derivative of utilitarian welfare with respect to z, as given by the left-hand-side of (56), is strictly positive at $z = 0$.

To see this note first that $\frac{d}{d\theta}\omega'(L(\theta)) = \omega''(L(\theta))L'(\theta) > 0$ while $e^*(\theta)$ is increasing in θ. Hence, there exists a type $\widetilde{\theta}$ such that

$$\begin{aligned}
&\int_\Theta \left[\omega'(L(\theta))\left(p(e^*(\theta)) - \int_\Phi p(e^*(\phi)) dF(\phi)\right)\right] dF(\theta) \\
= &\int_{\theta < \widetilde{\theta}} \left[\omega'(L(\theta))\left(p(e^*(\theta)) - E\left[p(e^*(\theta))\right]\right)\right] dF(\theta) \\
&+ \int_{\theta > \widetilde{\theta}} \left[\omega'(L(\theta))\left(p(e^*(\theta)) - E\left[p(e^*(\theta))\right]\right)\right] dF(\theta),
\end{aligned}$$

where the terms in the first integral are all strictly positive and the terms in the second integral are all strictly negative. Therefore, we have finally that indeed

$$\begin{aligned}
& \int_{\theta<\widetilde{\theta}} \left[\omega'(L(\theta))\left(p(e^*(\theta)) - E\left[p(e^*(\theta))\right]\right)\right] dF(\theta) \\
& \quad + \int_{\theta>\widetilde{\theta}} \left[\omega'(L(\theta))\left(p(e^*(\theta)) - E\left[p(e^*(\theta))\right]\right)\right] dF(\theta) \\
> \; & \int_{\theta<\widetilde{\theta}} \left[\omega'(L(\widetilde{\theta}))\left(p(e^*(\theta)) - E\left[p(e^*(\theta))\right]\right)\right] dF(\theta) \\
& \quad + \int_{\theta>\widetilde{\theta}} \left[\omega'(L(\widetilde{\theta}))\left(p(e^*(\theta)) - E\left[p(e^*(\theta))\right]\right)\right] dF(\theta) \\
= \; & 0.
\end{aligned}$$

Q.E.D.