Pigouvian Taxation with Costly Administration and Multiple Externalities

Daniel Jaqua
Albion College

Daniel Schaffa*
University of Michigan

August 27, 2016

Abstract

This paper generalizes corrective taxation to settings with costly administration and many externalities. If administrative cost varies only with the pollution generating activity, the optimal tax is equal to the externality added to the marginal administrative cost, and the private market fully internalizes the externality and administrative cost. If, due to the nature of enforcement, administrative cost varies with tax rates, then optimal policy leaves some portion of externalities uncorrected. As a result, using taxes to modify complement and substitute activities will be welfare increasing. Optimal policy may include subsidizing some harmful activities in order to reduce levels of even more harmful substitutes. Similarly, if the optimal activity mix changes with the scale of production and if higher scales of production use less harmful activity mixes, subsidizing the output of a harmful production process may be optimal.

Keywords: Pigouvian tax, externality, optimal tax, administrative cost, optimal tax systems, second-best
JEL Codes: H23, H21

---

*corresponding author: dschaffa@umich.edu. The authors give special thanks to Jim Hines for his advice and support. The authors thank Steve Bond, Paul Courant, Adam Dearing, Marina Epelman, Morris Hamilton, Louis Kaplow, Natalia Lazzati, JJ Prescott, Daniel Reck, Nate Seegert, Joel Slemrod, Ugo Troiano, Mike Zabek, the members of the University of Michigan Public Finance community (in particular our seminar attendees) and the faculty of the Centre for Business Taxation at Oxford University for invaluable comments and feedback. Schaffa gratefully acknowledges support from the NIA training grant to the Population Studies Center at the University of Michigan (T32 AG000221). Any errors are our own.

1 Introduction

Taxes used to modify externality generating behaviors are named for Pigou (1920), who first described many of their features. Pigouvian taxes improve welfare by aligning private incentives to a notion of public wellbeing. A broad class of policies that influence behavior, including carbon taxes, gasoline taxes, and toll roads, can be understood in a Pigouvian tax framework. The large budgets afforded government revenue collection agencies and the large number of different polluting activities[1] suggest that a model augmented to address costly administration and multiple externalities is worth studying and will increase understanding of the optimal regulation of externalities.

Administrative cost includes expenditures on measurement, enforcement, collections, legislation, and litigation. Some of these costs appear in the Internal Revenue Service budget, which was 11.5 billion dollars for fiscal year 2015.[2] Variable administrative cost is likely a function of activity levels (or income, in the IRS case) and tax rates. We call the case when administrative cost increases with activity levels *measurement costs*. Measurement costs arise because it is costly to determine the level of pollution generating activities, regardless of who is making these measurements and even if all parties are behaving honestly. Specific examples of these costs include the monitoring devices and the scientists who design and operate them. We call the case when administrative cost increases with tax rates *enforcement costs*. Enforcement costs arise if evasion increases with tax rates and the government pours more resources into tax enforcement, including more auditors, more lawyers, and more evasion detection software. We call the case when administrative cost increases with tax revenue collected *bureaucracy costs*. Bureaucracy costs are akin to the Flypaper effect—if revenue is larger, bureacracies are larger—for whatever reason, the money sticks.

As noted in Ballard and Fullerton (1992), most of the Pigouvian tax literature assumes tax systems can be administered costlessly. Polinsky and Shavell (1982) is an exception that considers Pigouvian taxes with one externality and administrative cost that depends on the quantity of the harmful activity.[3] We confirm their result, that the optimal Pigouvian tax increases by the marginal administrative cost, but also show that the optimal Pigouvian tax is smaller if administrative cost varies with the tax rate. Both these effects must be included, for example, when administrative cost increases with tax revenue. As Fullerton and Wolverton (2005) notes, administrative cost also

[1]For example, the EPA requires factories to report on 650 chemicals, which are produced by several different productive activities, in a Toxic Release Inventory report.

[2]See IRS (2016) for additional information on the IRS budget.

[3]See also Polinsky and Shavell (1992), Kaplow (1990a), and Kaplow (1990b), which study the similar topic of criminal sanctions that are costly to administer and enforce.

raises the possibility that indirect taxation, if administratively cheaper, might be preferable to direct taxation.[4]

Administrative cost adds further complexity if there is more than one externality causing activity. As in Yitzhaki (1979), if there are multiple activities, the social planner must choose the base (which activities to tax) and also what rates to apply to the taxed activities.[5] When administrative cost makes it optimal to leave an externality partially uncorrected, the optimal policy may use a tax on one activity to induce changes in a different activity level, as suggested by Lipsey and Lancaster (1956)'s theory of the second best. Using one tax to influence another activity level contrasts with the *principle of targeting* in Dixit et al. (1985) and the related *additivity property* in Sandmo (1975).[6] Both of these principles imply that the optimal tax does not depend on the complementarity or substitutability of the activities being taxed. We show that this *complementarity irrelevance* is a special case but that Sandmo's additivity result is preserved in a form which does include the complementarity and substitutability of activities.[7]

In the spirit of Mayshar (1991) and Slemrod (1990),[8] we find a general rule of Pigouvian taxation: at the optimal tax, the marginal social benefit of the tax must equal the marginal social cost of the tax. Our general rule, however, takes different forms depending on whether (1) there is administrative cost, (2) all or only some of the harmful activities are taxed, and (3) the tax is levied on the activities or output. *Set the tax equal to the externality* and *Set the marginal social benefit of the polluting activity equal to the marginal social harm of that activity* are special cases of this general rule.

Our first results shed light on how the optimal tax depends on administrative cost. We allow

[4]This relates more generally to the literature comparing direct and indirect taxation. Within the externalities literature, see Green and Sheshinski (1976) and Cremer and Gahvari (2001). Outside the externalities context, see Atkinson and Stiglitz (1976), which discusses the optimal income tax and public good provision in terms of direct and indirect taxation.

[5]Bovenberg and van der Ploeg (1994), Pirttilä and Tuomala (1997), Kaplow (1996), and Goulder et al. (1997) consider similar issues.

[6]The principle of targeting suggests that it is better to tax the activity you intend to influence rather than tax a complement or substitute. Sandmo's additivity property states that the marginal social harm does not affect the optimal tax on other activities, regardless of complementarity—he calls it additivity because the marginal social harm enters the tax formula additively.

[7]The literature reinforcing complementarity irrelevance is large. See Bovenberg and De Mooij (1994), Bovenberg and van der Ploeg (1994), Bovenberg and Goulder (1996), Fullerton (1997), Pirttilä and Tuomala (1997), Ng (1980), Kopczuk (2003), and Kaplow (2012). A notable exception is Cremer et al. (1998), where complements and substitutes matter in a model that includes heterogeneous consumers and nonlinear external harm. Complementarity relevance does appear in the optimal tax literature outside of Pigouvian taxation. See notably Corlett and Hague (1953), where a good, e.g. leisure, cannot be taxed; and Cremer and Gahvari (1993), where there are multiple commodities, evasion, and costly audits.

[8]Stiglitz and Dasgupta (1971), Atkinson and Stern (1974), and Ballard and Fullerton (1992) analyze the relationship between Pigouvian taxes and the marginal cost of funds. An overview of the literature on tax administration and administrative cost can be found in Slemrod and Yitzhaki (2002), although the main focus is on tax evasion.

the variable administrative cost function to admit pollution generating activity levels and tax rates as arguments and also consider fixed costs. If the administrative cost varies only with the pollution generating activities, the optimal tax is equal to the externality added to the marginal administrative cost. The tax is greater compared to the case with no administrative cost, and the administrative cost can be thought of as an additional externality caused by the activity, which is fully corrected by the tax. If the tax rate enters the variable administrative cost function, higher tax rates decrease pollution levels but increase administrative cost. The optimal tax is lower than the case with no administrative cost, leaving the externality partially uncorrected.

We next extend the model to study multiple pollution generating activities. We call the case where not all the activities are taxed *incomplete taxation*. Incomplete taxation leads to uncorrected externalities but is optimal if adding activities to the tax base has large fixed costs. If there are uncorrected externalities, the optimal tax on each activity is determined in part by its ability to induce changes in other activity levels. Subsidizing a harmful activity may be optimal if that subsidy leads to less of a different harmful activity.

Lastly we describe the difference between output taxes and taxes on polluting activities. When there are multiple production activities, we show that an output tax is the same as setting a tax on each activity equal to the output tax rate times that activity's marginal product. The output tax is a blunter instrument because it cannot generally induce the private market to substitute between activities. An output tax can only lead to substitution between production activities if the optimal activity mix changes with the scale of production. If higher production levels use less harmful activity mixes, subsidizing a harmful production process may be optimal. We quantify the welfare difference between the optimal output tax and the optimal activity tax, which determines when a policymaker should use an output tax and when the policymaker should use production activity taxes.

2 Model with one activity

This section describes how administrative cost changes the optimal tax when there is only one activity and that activity generates an externality. The single activity model previews many of the results from the multiple externality model without the additional complexity of complementarity.

There is a single choice variable, $x \in \mathbb{R}_+$, which may be interpreted as production input quantity, production activity level, or other variables.[9] Following Ramsey (1927), the economy

[9] We find activity level the most natural interpretation. x could be a specific type of pollution, such as carbon dioxide, which has a well defined externality but a less well defined private benefit. x could be a consumption good, such as watts of electricity, with a clearly defined private benefit but a less well defined externality harm. x could be an

operates according to a net benefit function, b, which maps activity level to private net benefit. At low activity levels, the marginal benefit of production exceeds the marginal cost and increasing the activity leads to higher net benefit. At high activity levels, the marginal cost of production exceeds the marginal benefit of production and increasing the activity leads to lower net benefit.

Intuitively, consider a supply and demand diagram with activity level on the x-axis: $b(x)$ is simply the area below the activity demand curve and above the activity supply curve on the set $[0, x]$. No single individual or firm chooses x—rather it is the result of a market equilibrium. Formally b is twice continuously differentiable, strictly concave, and achieves its maximum somewhere in the interior of $\mathbb{R}_+$.[10]

Both total external harm and tax liability are linear functions of the activity level.[11] Let e be the externality and t be the tax. The discussion assumes a negative externality but the analysis is the same for a positive externality. Tax revenues are assumed to be lump-sum redistributed.

The social planner must construct an administrative apparatus to collect and enforce taxes. Let c map the activity level and tax rate to administrative cost. Formally c is continuously differentiable and weakly convex, and $\arg\min(c) = (0, 0)$. These assumptions are made for tractability, but they are consistent with a planner who employs the most effective tax collecting and enforcement resources first. Administrative cost and marginal administrative cost (with respect to x) are both increasing in activity level. A subsidy should also be costly to administer, so administrative cost and marginal administrative cost (with respect to t) are both increasing in the tax rate above 0 and are both decreasing in the tax rate below 0. The private market[12] solves the following problem:

$$\max_x b(x) - tx$$

which leads to the first order condition $b'(x) - t = 0$. Because b is strictly concave, b' is invertible. Therefore, by the implicit mapping theorem, there exists a continuously differentiable function, $x(t)$, such that $b'(x(t)) - t = 0$. $x(t)$ is the private market's best response function to the tax. Note that $b''(x(t))x'(t) = 1$, so $x'(t)$ is also invertible and $x'(t)^{-1} = t'(x)$. The social planner solves the

input, such as coal, but the amount of externality is difficult to define since it depends on, for example, whether the coal is an input in a modern 'clean' coal power plant or an older coal fired power plant. The activity *producing electricity in a clean coal power plant* has a clearly defined externality and a clearly defined benefit. Sandmo (1978) discusses this issue for taxes on commodities that can be consumed in several ways but only some consumption processes generate externalities.

[10] One way to ensure this outcome is to assume that $\lim_{x\to 0} \frac{\partial b}{\partial x} = \infty$ and $\lim_{x\to k} \frac{\partial b}{\partial x} = -\infty$ for some $k > 0$.

[11] It would not be difficult to allow external harm to be a nonlinear weakly convex function of x.

[12] Throughout the paper, for brevity, we speak of the private market as an agent. More precisely, individuals and firms make choices in the private market that result in an aggregate quantity of activity. We represent their behavior as a maximization problem. The solution to the maximization problem is the market equilibrium quantity of activity.

following problem:

$$\max_t b(x(t)) - ex(t) - c(x(t), t)$$

which leads to the first order condition $b'(x(t^*))\frac{\partial x}{\partial t} - e\frac{\partial x}{\partial t} - c_1\frac{\partial x}{\partial t} - c_2 = 0$ where c_i denotes the partial derivative of c with respect to its i^{th} argument. Note the general rule: the marginal social benefit of the tax (reduced externality and decreased administrative cost) is equal to the marginal social cost of the tax (reduced private benefit and increased administrative cost). Substituting $b'(x(t^*)) = t^*$, dividing by $\frac{\partial x}{\partial t}$, and rearranging yields:

$$t^* = e + c_1 + c_2\frac{\partial t}{\partial x} \tag{1}$$

The following table presents the optimal tax associated with several possible functions of administrative cost. Note that for all cases with administrative cost the optimal tax expressions are implicit formulas.

Table 1: single activity optimal taxes

Case	Cost function	Optimal Pigouvian tax
No administrative cost	$c = 0$	$t^* = e$
Measuring costs	$c = c(x)$	$t^* = e + \frac{\partial c}{\partial x}$
Enforcement costs	$c = c(t)$	$t^* = e + \frac{\partial c}{\partial t}\frac{\partial t}{\partial x}$
Bureaucracy costs	$c = c(xt) = c(R)$	$t^* = e + \frac{\partial c}{\partial R}\left(t^* + x(t^*)\frac{\partial t}{\partial x}\right)$
Arbitrary costs	$c = c(x, t)$	$t^* = e + c_1 + c_2\frac{\partial t}{\partial x}$

The case with no administrative cost is the classical Pigou (1920) result: *the optimal tax is equal to the externality*.[13] The market maximizes private net benefit when private marginal benefit is equal to private marginal cost. The social planner uses a tax to set social marginal cost equal to social marginal benefit by reducing private marginal benefit by the amount of the externality.

We call $c(x)$ the measuring costs case. Activity levels would affect administrative cost if, for example, higher activity levels implied more firms or factories that required more pollution measuring devices, more scientists, and so on. This is the functional form that Polinsky and Shavell (1982) posit.[14] Note that the optimal tax is not equal to the externality. If administrative cost is a function of x, the social planner should raise the tax until *the marginal social benefit of the activity*

[13] In the appendix we show if there is no administrative cost, t^* exists and is the unique global maximizer.

[14] It is possible that $b(x(0)) - ex(0) > b(x(t^*)) - ex(t^*) - c(x(t^*))$ in which case the optimal tax is $t = 0$.

equals the marginal social cost of the activity. Because administrative cost increases only with x the administrative cost can be interpreted as an additional externality. Thus the optimal tax induces the private market to fully internalize the externality and the administrative cost. The tax rate is always higher than the case with no administrative cost.[15] This shows up in the model because $\frac{\partial c}{\partial x} > 0$.

We call $c(t)$ the enforcement costs case. Tax rates would affect administrative cost if higher rates increased the incentive to evade taxes and thus increased the need to audit polluting firms. Note that the tax is not equal to the externality and that at the optimum the marginal social benefit of the activity is smaller than the marginal social cost of the activity. The social planner should raise the tax until *the marginal social benefit of the tax is equal to the tax's marginal social harm*. If administrative cost is a function of tax rates, higher taxes reduce the external harm but increase administrative cost. The marginal benefit of a higher tax is lower external harm, and the marginal cost of a higher tax is lower private net benefit and higher administrative cost. Thus the tax rate is always lower than the case with no administrative cost,[16] which leaves an uncorrected externality. This shows up in the model because $\frac{\partial c}{\partial t}/\frac{\partial x}{\partial t} < 0$.[17]

We call $c(xt)$ the bureaucracy costs case. Revenue would increase administrative cost if more revenue led to a larger bureaucracy. We also include $c(x, t)$, the arbitrary costs case which covers all the cases described above in addition to other possible administrative cost functions.

Tax administration could have fixed costs if a building of at least a certain size were required to administer the tax or if the tax involved drafting, disseminating, and litigating new statutes. If there are fixed costs of f, the social planner compares $b(x(0)) - ex(0)$ and $b(x(t^*)) - ex(t^*) - c(x(t^*), t^*) - f$ to determine whether the tax should be implemented.

The following two results, about revenue raising taxes and Pigouvian taxes on output, are presented in one dimension to preview the results in the model with multiple activities. Both results are simple in the one dimensional case but are more complicated and interesting when there are multiple activities.

Tax systems are designed to generate revenue in addition to aligning incentives. If there is only one activity and the optimal Pigouvian tax generates sufficient revenue to satisfy the revenue constraint, the optimal tax is the optimal Pigouvian tax and the excess revenue should be lump-sum redistributed. If the optimal Pigouvian tax generates insufficient revenue to satisfy the revenue constraint, there are two possible cases. If at the optimal Pigouvian tax, a higher tax rate generates

[15] However, the optimal subsidy is always smaller than the case with no administrative cost.

[16] Similarly, if there is a positive externality the subsidy rate is always lower than the case with no administrative cost.

[17] For positive externalities $\frac{\partial c}{\partial t}/\frac{\partial x}{\partial t} > 0$. In either case the optimal tax will never be 0 because the marginal administrative cost is 0 at $t = 0$.

more revenue, then the optimal tax is higher than the optimal Pigouvian tax. If at the optimal Pigouvian tax, a higher tax generates less revenue, then the optimal tax is lower than the optimal Pigouvian tax. The revenue requirement may be unobtainable.

Thus far we have not explicitly discussed output. If there is only one production activity and one output, the social planner can achieve the same welfare with either a tax on the activity or a tax on output. Assume that the activity generates output according to the strictly increasing and concave function q. With output tax τ, the private market solves:

$$\max_x b(x) - \tau q(x)$$

which leads to the first order condition $b'(x) - \tau q'(x) = 0$, which defines a best response function $x(\tau)$.[18] The social planner solves the following problem:

$$\max_\tau b(x(\tau)) - ex(\tau)$$

which leads to the first order condition $b'(x(\tau^*))x'(\tau^*) - ex'(\tau^*) = 0$, which sets the marginal benefit of the tax equal to the marginal cost of the tax —where the benefit is reduced external harm and the cost is reduced private net benefit. Substituting the private market's first order condition, $b'(x(\tau^*)) = \tau^* q'(x(\tau^*))$, dividing by $x'(\tau^*)$, and rearranging yields:

$$\tau^* q'(x(\tau^*)) = e \implies \tau^* = e/q'(x(\tau^*)) \tag{2}$$

The equation on the left shows that at the optimum the marginal tax burden is equal to the marginal external harm. The marginal product of the activity scales the tax to properly align private incentives. Accordingly, when there is a single activity, the optimal activity tax and the optimal output tax yield the same welfare. With only fixed administrative costs the social planner should choose whichever tax is cheaper to administer.

3 Model with multiple activities

If there are multiple activities much of the intuition from the previous section applies, but there are two important differences. First, it is possible that some but not all activities are taxed. We call this *incomplete* taxation. Second, increasing the tax on one activity could increase (substitutes) or decrease (complements) the level of another activity. We call the change in activity A in response

[18] $x(\tau)$ is well defined if $b(x) - \tau q(x)$ is concave and is increasing near $x = 0$ and decreasing for some finite x.

to a tax on B a *cross tax effect*. If cross tax effects appear in the formula for optimal taxes we say that *complementarity* is relevant to the optimal taxes. If taxes are complete and costless, cross tax effects do not enter into the optimal tax formula and we call this *complementarity irrelevance*. When the optimal tax leaves some externalities partially uncorrected (i.e. taxes are incomplete or administrative cost is a function of the tax rates), cross tax effects enter the optimal tax formula.

Let x now be an n-dimensional vector of activities. Activities are used to generate goods and thus indirectly increase utility. Activities have private costs and can only be performed in non-negative quantities. Let $X \subseteq \mathbb{R}^n_+$ be the set of allowable activities.[19] The net benefit function $b : X \to \mathbb{R}$ maps the activity levels to private net benefit.[20] Again we assume b is twice continuously differentiable, strictly concave, and achieves its maximum somewhere in the interior of X.[21] Intuitively $b(x)$ is the area between many activity supply and demand curves and a change in one activity could shift either the demand or supply curves for all the other activities. If x consists of all productive activities, then this is a general equilibrium model. Both externalities and taxes are linear functions of activities. Let e be the n-dimensional vector of activity externalities[22] and t be the n-dimensional vector of activity taxes. Tax revenues are assumed to be lump-sum redistributed.

3.1 Costless administration

We begin by generalizing the classical Pigouvian tax case.

Proposition 1. *If activity taxes are complete and tax administration is costless, the optimal tax vector is equal to the externality vector.*

Proof. The private market solves the following problem:

$$\max_x b(x) - t^\top x$$

which leads to the first order condition $b'(x) - t^\top = 0$.[23] Because b is strictly concave it has an invertible Hessian. Therefore, by the implicit mapping theorem, there exists a continuously differentiable function, $x(t) : \mathbb{R}^n \to \mathbb{R}^n$, such that $b'(x(t)) - t^\top = 0$. $x(t)$ is the private market's best response function to the tax vector. Note that $b''(x(t))x'(t) = I$, the identity matrix, so $x'(t)$

[19] We assume this set is convex and open, has finite measure, and lies entirely in the positive orthant.

[20] The function is similar to the one used in Ramsey (1927). Our model generalizes Sandmo (1975). To see this substitute the linear ppf into the strictly concave utility function. Since a strictly concave function of an affine transformation is strictly concave this yields the b function.

[21] One way to ensure this outcome is to assume that $\lim_{x_i \searrow \partial X} \frac{\partial b}{\partial x_i} = \infty$ and $\lim_{x_i \nearrow \partial X} \frac{\partial b}{\partial x_i} = -\infty$.

[22] It would not be difficult to allow harm to be a nonlinear weakly convex function of activities.

[23] Throughout the multidimensional sections of the paper we use matrix calculus and the associated notation.

is also invertible and $x'(t)^{-1} = t'(x)$. The social planner solves the following problem:

$$\max_t b(x(t)) - e^\top x(t)$$

which leads to the first order condition $b'(x(t^*))x'(t^*) - e^\top x'(t^*) = 0$. Substituting $b'(x(t^*)) = t^{*\top}$ yields $(t^* - e)^\top x'(t^*) = 0$. $t^* = e$ is clearly a solution, and the invertibility of $x'(t^*)$ ensures that it is the unique solution.[24] $\square$

This is Pigou (1920)'s remarkable result generalized to arbitrary dimensions: *the optimal tax is equal to the externality*. At $t = e$, the private market fully internalizes every externality, and the policymaker does not need to know or use information other than the externality vector. Any other tax policy is strictly worse. Using the tax on activity A to induce changes in activity B is not welfare improving because there is no benefit to changing the activity level in a market that already internalizes the externality.

3.2 Incomplete taxation

Incomplete taxation is optimal if the social planner is unable to measure certain activities or if taxing some activities incurs high fixed costs. Leaving even one of the relevant activities untaxed alters the analysis because the tax vector no longer sets the marginal social benefit of each activity equal to the marginal social cost of each activity.

Example 1. Let x_1 be operating a natural gas plant and x_2 be operating a coal plant to produce electricity. Assume there is only one tax, t_1 on x_1. The private market maximizes $b(x_1, x_2) - t_1 x_1$, leading to the first order conditions $\frac{\partial b}{\partial x_1} = t_1$ and $\frac{\partial b}{\partial x_2} = 0$. The social planner's problem is:

$$\max_{t_1} b(x_1(t_1), x_2(t_1)) - e_1 x_1(t_1) - e_2 x_2(t_1)$$

with first order condition $\frac{\partial b}{\partial x_1}\frac{\partial x_1}{\partial t_1} + \frac{\partial b}{\partial x_2}\frac{\partial x_2}{\partial t_1} - e_1\frac{\partial x_1}{\partial t_1} - e_2\frac{\partial x_2}{\partial t_1} = 0$. Substituting the private market first order condition leads to $(t_1^* - e_1)\frac{\partial x_1}{\partial t_1} - e_2\frac{\partial x_2}{\partial t_1} = 0$. Since $\frac{\partial x_1}{\partial t_1} < 0$ the optimal tax is:

$$t_1^* = e_1 + e_2\frac{\partial x_2}{\partial t_1}\frac{\partial t_1}{\partial x_1}$$

The sign of $\frac{\partial x_2}{\partial t_1}$ is ambiguous and depends on the complementarity of x_1 and x_2. If the two activities are substitutes then $\frac{\partial x_2}{\partial t_1} > 0$; if they are complements then $\frac{\partial x_2}{\partial t_1} < 0$. A subsidy will be optimal if

[24] In the appendix we show that if there is no administrative cost t^* exists and is the unique global maximizer. Depending on the exact functional form that c takes it is possible that there exists no finite t^* or that t^* is not unique.

the taxed activity has the smaller externality and the two activities are very substitutable. In our example, if coal has a much larger external harm, natural gas is a substitute, and the social planner cannot administer a tax on coal, then a subsidy on natural gas will be optimal, even though using natural gas to generate electricity has an external harm.

In the above example, the optimal tax depends on complementarity because the marginal private benefit of each untaxed activity will be 0 regardless of that activity's external harm. Because of this uncorrected externality the marginal social benefit of an untaxed activity will not equal that activity's marginal social cost. Using taxes to change the levels of untaxed, externally harmful activities improves welfare. At the optimum, the marginal social benefit of each tax is equal to that tax's marginal social cost. The marginal social benefit of a tax is the reduction in external harm of the taxed activity and all untaxed activities, and the marginal social cost of the tax is the lost private benefit. The following proposition generalizes the example to arbitrary dimensions.

Proposition 2. *If tax administration is costless, the optimal tax on each taxed activity is equal to the externality generated by that activity plus the externalities of all untaxed activities weighted by the responsiveness of the untaxed activity to changes in the taxed activity.*

Proof. Let Θ be the power set of $\{1, ..., n\}$, θ an arbitrary element of Θ, m the dimension of θ, x_θ the m-dimensional vector of taxed activities, $\bar{x}_\theta$ the $n-m$-dimensional vector of untaxed activities, t_θ the m-dimensional vector of taxes on x_θ, e_θ the m-dimensional vector of externalities generated by x_θ, and $\bar{e}_\theta$ the $n-m$-dimensional vector of externalities generated by $\bar{x}_\theta$. The private market solves the following problem:

$$\max_x b(x) - t_\theta^\top x_\theta$$

with solution $\frac{\partial b}{\partial x_j} = 0$ for $j \notin \theta$ and $\frac{\partial b}{\partial x_i} = t_i$ for $i \in \theta$. As before, the concavity of b ensures that a continuously differentiable best response function, $x(t_\theta)$ exists. The social planner solves the following problem:

$$\max_{t_\theta} b(x(t_\theta)) - e_\theta^\top x_\theta(t_\theta) - \bar{e}_\theta^\top \bar{x}_\theta(t_\theta)$$

which leads to the first order condition $b'(x(t_\theta^*))x'(t_\theta^*) - e_\theta^\top x'_\theta(t_\theta^*) - \bar{e}_\theta^\top \bar{x}'_\theta(t_\theta^*) = 0$. Note that the marginal harm of the tax is equal to the marginal benefit of the tax. Substituting the private market first order condition yields $(t_\theta^* - e_\theta)^\top x'_\theta(t_\theta^*) - \bar{e}_\theta^\top \bar{x}'_\theta(t_\theta^*) = 0$. Rearranging and applying the invertibilty of $x'_\theta(t^*_\theta)$ yields $t^{*\top}_\theta = e^\top_\theta + \bar{e}^\top_\theta \bar{x}'_\theta(t^*_\theta) x'_\theta(t^*_\theta)^{-1}$. For each individual tax:

$$t^*_i = e_i + \sum_{k \in \theta} \sum_{j \notin \theta} e_j \tfrac{\partial x_j}{\partial t_k} \tfrac{\partial t_k}{\partial x_i} \tag{3}$$

□

Note the special case, $m = n$ in which $x'(t^*_\theta) = x'_\theta(t^*_\theta)$ gives $t^*_\theta = e$ as before. If $m < n$ each activity tax may be above or below the associated externality; it may even have the opposite sign as the associated externality.

3.3 Variable administrative costs

As in the one-dimensional model, when administrative cost is a function of activity levels, taxes should equal the externality plus the marginal administrative cost, and taxes will be higher compared to the costless case. At the optimal tax the private market fully internalizes the external harm and administrative cost. Thus the optimal tax expression does not include cross tax effects.

When administrative cost is a function of tax rates, taxes cannot cause the private market to fully internalize the externality. The planner should use taxes with low marginal administrative cost to affect other activity levels. Thus the optimal tax expression does include cross tax effects. The social planner should set the marginal benefit of each tax equal to the marginal cost of that tax.

The cost function $c(x, t)$, $c : X \times \mathbb{R}^n \to \mathbb{R}$, is the multidimensional analogue of the single activity case. We assume that $c(x, t)$ is continuously differentiable, weakly convex, and has $\arg\min(c) = (0, 0)$.

Example 2. Let x_1 be operating a natural gas plant and x_2 be operating a coal plant to produce electricity. Assume that both activities are taxed and administrative cost increases with t. The private market maximizes $b(x_1, x_2) - t_1 x_1 - t_2 x_2$, leading to the first order conditions $\frac{\partial b}{\partial x_1} = t_1$ and $\frac{\partial b}{\partial x_2} = t_2$. The concavity of b ensures that $x_1(t_1, t_2)$ and $x_2(t_1, t_2)$ exist. The social planner's problem is:

$$\max_t b(x_1(t_1, t_2), x_2(t_1, t_2)) - e_1 x_1(t_1, t_2) - e_2 x_2(t_1, t_2) - c(t_1, t_2)$$

with first order conditions:

$$\tfrac{\partial b}{\partial x_1} \tfrac{\partial x_1}{\partial t_1} + \tfrac{\partial b}{\partial x_2} \tfrac{\partial x_2}{\partial t_1} - e_1 \tfrac{\partial x_1}{\partial t_1} - e_2 \tfrac{\partial x_2}{\partial t_1} - \tfrac{\partial c}{\partial t_1} = 0 \text{ and}$$
$$\tfrac{\partial b}{\partial x_1} \tfrac{\partial x_1}{\partial t_2} + \tfrac{\partial b}{\partial x_2} \tfrac{\partial x_2}{\partial t_2} - e_1 \tfrac{\partial x_1}{\partial t_2} - e_2 \tfrac{\partial x_2}{\partial t_2} - \tfrac{\partial c}{\partial t_2} = 0$$

Substituting the private market first order condition and manipulating the equation leads to:

$$t_1^* = e_1 + \frac{\partial c}{\partial t_1}\frac{\frac{\partial x_2}{\partial t_2}}{\frac{\partial x_1}{\partial t_1}\frac{\partial x_2}{\partial t_2} - \frac{\partial x_1}{\partial t_2}\frac{\partial x_2}{\partial t_1}} - \frac{\partial c}{\partial t_2}\frac{\frac{\partial x_2}{\partial t_1}}{\frac{\partial x_1}{\partial t_1}\frac{\partial x_2}{\partial t_2} - \frac{\partial x_1}{\partial t_2}\frac{\partial x_2}{\partial t_1}} \text{ and}$$
$$t_2^* = e_2 + \frac{\partial c}{\partial t_2}\frac{\frac{\partial x_1}{\partial t_1}}{\frac{\partial x_1}{\partial t_1}\frac{\partial x_2}{\partial t_2} - \frac{\partial x_1}{\partial t_2}\frac{\partial x_2}{\partial t_1}} - \frac{\partial c}{\partial t_1}\frac{\frac{\partial x_1}{\partial t_2}}{\frac{\partial x_1}{\partial t_1}\frac{\partial x_2}{\partial t_2} - \frac{\partial x_1}{\partial t_2}\frac{\partial x_2}{\partial t_1}}$$

Noting that $x'(t)^{-1} = t'(x)$, we have:

$$t_1^* = e_1 + \frac{\partial c}{\partial t_1}\frac{\partial t_1}{\partial x_1} + \frac{\partial c}{\partial t_2}\frac{\partial t_2}{\partial x_1} \text{ and}$$
$$t_2^* = e_2 + \frac{\partial c}{\partial t_2}\frac{\partial t_2}{\partial x_2} + \frac{\partial c}{\partial t_1}\frac{\partial t_1}{\partial x_2}$$

A subsidy on natural gas will be optimal if the cost of administering the subsidy is small ($\frac{\partial c}{\partial t_1}$ is negative and has a small magnitude), administering the tax on coal is costly ($\frac{\partial c}{\partial t_2}$ is positive and has a large magnitude), coal is more harmful than natural gas ($e_2 > e_1$), and coal x_2 and natural gas x_1 are very substitutable ($\frac{\partial x_2}{\partial t_1}$ is positive with large magnitude).

In arbitrary dimensions, the private market problem is the same as in the multiple activity, complete taxation, costless administration case. The social planner solves:

$$\max_t b(x(t)) - e^\top x(t) - c(x(t), t)$$

which leads to the first order condition $b'(x(t^*))x'(t^*) - e^\top x'(t^*) - c_1(x(t^*), t^*)x'(t^*) - c_2(x(t^*), t^*) = 0$, where c_1 denotes the partial derivative of c with respect to its first vector argument and c_2 is the partial derivative of c with respect to its second vector argument. Substituting in the private market first order condition and applying $x'(t^*)$'s invertibility yields:

$$t^{*\top} = e^\top + c_1 + c_2 x'(t^*)^{-1} \tag{4}$$

The table below describes some special cases. Note that for both cases with administrative cost the optimal tax expressions are implicit formulas.

Table 2: Multiple activity optimal taxes

Case	Cost function	Optimal Pigouvian tax matrix notation	Optimal Pigouvian tax element notation
No administrative cost	$c = 0$	$t^{*\top} = e^\top$	$t_i^* = e_i$
Measuring costs	$c = c(x)$	$t^{*\top} = e^\top + c'(x(t^*))$	$t_i^* = e_i + \frac{\partial c}{\partial x_i}$
Enforcement costs	$c = c(t)$	$t^{*\top} = e^\top + c'(t^*)x'(t^*)^{-1}$	$t_i^* = e_i + \sum_j \frac{\partial c}{\partial t_j}\frac{\partial t_j}{\partial x_i}$

Just as in the single activity case, if administrative cost is a function of only x, the administrative cost can be fully internalized.[25]

Proposition 3. *If administration is costly and the cost is a function of activity levels, the optimal tax vector is equal to the externality vector added to the marginal administrative cost.*

If administrative cost is a function of x, the social planner should raise the tax until *the marginal social benefit of the activity equals the marginal social cost of the activity*. The administrative cost can be interpreted as an additional externality. The characterization of the optimum provides an implicit solution—both sides of the equation are functions of t^*. Nonetheless there is no complementarity: $x'(t^*)$ does not enter into the equation. This is not the case when administration is a function of t.

Proposition 4. *If administration is costly and the cost is a function of the tax rates, the optimal tax on each activity is equal to the externality generated by that activity plus the sum of all the marginal administrative costs weighted by the responsiveness of the tax rate to changes in activity.*

If administrative cost varies with t, the optimal tax does not exhibit complementarity irrelevance. The social planner should raise taxes until *the marginal social benefit of the tax is equal to the tax's marginal social harm*. A subsidy may be optimal for some externally harmful activities.[26] c' will take negative values whenever there is a subsidy because the cost of administration will decrease the less negative the tax becomes.

[25] The administrative cost function could change if some taxes are not used. As with incomplete taxation, there is no analytical solution—the social planner must find the optimal tax for each set of taxes and chose the one that yields the highest welfare. There are up to 2^n different possible administrative cost functions, reflecting the monitoring costs for every possible set of activity taxes. Proposition 3 assumes that the cost function does not change with the tax base.

[26] The matrix $t'(x(t^*)) = b''(x(t^*))$ describes the effect of a change in the activity vector on tax rates at the optimal tax rate. Because b is concave, this matrix is negative definite, so the diagonals are all negative. An increase in t_i will, thus, reduce x_i although it may increase or have no effect on x_j. This implies that t^* may have negative entries.

3.4 Fixed administrative costs

If there are fixed administrative costs, it may not be optimal to tax some activities at all—leading to incomplete taxation. Leaving even one of the relevant activities untaxed alters the analysis (compared to complete taxation) because the tax vector no longer sets the marginal benefit of each activity equal to the marginal harm of each activity—for all untaxed activities, the marginal private benefit is equal to 0 regardless of the tax on the other variables. At the optimum each tax must not only account for the activity it is applied to but also every other untaxed activity.

If each activity tax has its own fixed cost, the social planner must optimize the social welfare function 2^n times—once for each possible combination of taxes—using incomplete taxes as described above. Let the fixed cost for each tax be f_i. Recall that the power set of $\{0, ..., n\}$ is Θ. The social planner's problem is:

$$\max_{\theta \in \Theta} \left\{ \max_{t_\theta} b(x(t_\theta)) - e^\top x(t_\theta) - \sum_{i \in \theta} f_i \right\}$$

No closed form solution exists, but the optimal tax expression for each subproblem is:

$$t_\theta^{*\top} = e_\theta^\top + \bar{e}_\theta^\top \bar{x}'_\theta(t_\theta^*) x'_\theta(t_\theta^*)^{-1} \tag{5}$$

Variable administrative cost $c(x_\theta(t_\theta), t_\theta)$ may be included in which case the optimal tax expression for each subproblem is:

$$t_\theta^{*\top} = e^\top x'(t_\theta^*) x'_\theta(t_\theta^*)^{-1} + c_1 + c_2 x'_\theta(t_\theta^*)^{-1} \tag{6}$$

The same intuition discussed above applies in this case.

3.5 Revenue requirement

Pigouvian taxes generate revenue in addition to aligning incentives. In this section we show how the optimal taxes described above are altered when there is a revenue constraint.[27] Assume that the

[27]This version of the Pigouvian tax problem has been analyzed by many papers. See Sandmo (1975), demonstrating that the optimal tax additively combines expressions relating to the Pigouvian and revenue raising taxes; Kopczuk (2003), generalizing Sandmo's additivity property; Bovenberg and De Mooij (1994), describing the relationship between a tax on labor and a Pigouvian tax on a polluting consumption good; and Kaplow (2012), demonstrating that changing a commodity tax to the first best Piouvian tax while making compensating changes in the income tax creates a Pareto improvement in a model where utility is separable in labor and other activities.

revenue requirement is $R \geq t_\theta^\top x_\theta(t_\theta)$. We begin with the Lagrangian:

$$\mathcal{L} = b(x(t_\theta)) + \lambda(R - t_\theta^\top x_\theta(t_\theta))$$

with first order condition:

$$b'(x(t_\theta^*))x'(t_\theta^*) - \lambda(t_\theta^{*\top} x_\theta'(t_\theta^*) + x_\theta(t_\theta^*)^\top) = 0$$

which sets the marginal social cost of each tax equal to the marginal revenue of that tax times the Lagrange multiplier and leads to the optimal tax expression:

$$t_\theta^{*\top} = \frac{\lambda}{1-\lambda} x_\theta(t_\theta^*)^\top x_\theta'(t_\theta^*)^{-1} \tag{7}$$

Note that the Karush Kuhn Tucker conditions require that $\lambda \leq 0$, which makes sense since increasing the required revenue should decrease private net benefit. Now we determine the optimal tax when there is a revenue constraint and Pigouvian taxation using the most general model with incomplete taxation and costly administration.

Proposition 5. *The optimal Pigouvian tax with a revenue constraint is equal to the expression for the optimal revenue constrained tax added to the expression for the optimal unconstrained Pigouvian tax times a scaling factor.*

Proof. Starting with the Lagrangian:

$$\mathcal{L} = b(x(t_\theta)) - e^\top x(t_\theta) - c(x_\theta(t_\theta), t_\theta) + \lambda(R - t_\theta^\top x_\theta(t_\theta))$$

with first order condition:

$$b'(x(t_\theta^*))x'(t_\theta^*) - e^\top x'(t_\theta^*) - c_1 x_\theta'(t_\theta^*) - c_2 - \lambda t_\theta^{*\top} x_\theta'(t_\theta^*) - \lambda x_\theta(t_\theta^*)^\top = 0$$

which sets the marginal social cost of each tax equal to the marginal revenue of that tax times the Lagrange multiplier. Rearranging:

$$(1-\lambda)t_\theta^{*\top} x_\theta'(t_\theta^*) = e^\top x'(t_\theta^*) + c_1 x_\theta'(t_\theta^*) + c_2 + \lambda x_\theta(t_\theta^*)^\top$$

The optimal tax expression is:

$$t_\theta^{*\top} = \frac{1}{1-\lambda} \underbrace{\left(e^\top x'(t_\theta^*) x_\theta'(t_\theta^*)^{-1} + c_1 + c_2 x_\theta'(t_\theta^*)^{-1}\right)}_{\text{optimal Pigouvian tax}} + \underbrace{\frac{\lambda}{1-\lambda} x_\theta(t_\theta^*)^\top x_\theta'(t_\theta^*)^{-1}}_{\text{optimal revenue constrained tax}} \tag{8}$$

□

The Karush Kuhn Tucker conditions require that $\lambda \leq 0$. If the revenue generated by the Pigouvian tax exceeds R, then $\lambda = 0$ and the optimal tax is equal to the Pigouvian tax.

The additivity property espoused by Sandmo (1975) only strictly applies when every activity is taxed and costs are only a function of activity levels—his definition of additivity hinged on there being no complementarity effects. However, in all cases the optimal tax is the optimal Pigouvian tax multiplied by $\frac{1}{1-\lambda}$ added to the optimal revenue constrained tax.

4 Model with output tax

Taxing every activity or even a large subset of activities would be expensive. Potential costs include drafting and disseminating the tax code and paying for the bureaucracy necessary to monitor each activity. An alternative to an activity tax scheme would be an output tax or potentially an output tax and a tax on a smaller subset of activities. Output may be much easier to measure because of the record keeping associated with market transactions. Determining whether it is better to tax output or activities requires knowing how much welfare is lost by taxing output as opposed to activities.[28]

An output tax cannot generally induce the private market to select the socially optimal combination of activities. Whereas activity taxes can induce firms to substitute one activity for another, an output tax generally cannot. Consider, for example, producing electricity from either coal or natural gas. Assume using coal is privately cheaper (by an arbitrarily small amount) but also has a larger externality. An output tax will not discourage coal use relative to natural gas use, but an activity tax on coal will. In this particular case, since coal is only barely cheaper than natural gas and since they are perfect substitutes, an infinite tax (or any tax that makes natural gas privately preferable to coal) on coal will be optimal. Output taxes can cause the private market to substitute

[28] As noted in the single activity case, when there is only one activity and production is a function of this activity, the social planner can achieve the same welfare with either a tax on the activity or a tax on output. With multiple externalities, an output tax can achieve the same welfare as the activity taxes in special cases, such as when the pollution generating production activities are perfect complements in production. In that case, because some set proportion of the activities is required to increase output, the optimal tax can be set considering only the marginal externality associated with a marginal change in output.

between activities if different combinations of activities are optimal at different scales of production. A tax on electricity could cause substitution from coal to natural gas if using coal exhibits better economies of scale.

4.1 Costless administration

The model used in the previous section is also used in this section albeit with some modification. In the previous section, the model made no explicit reference to output. In fact, the model could implicitly include many different outputs. In this section, we restrict the model to one output and introduce the increasing and concave function $q(x) : X \to \mathbb{R}$ which maps the vector of activities to the quantity of output produced.

Proposition 6. *At the optimal output tax, the marginal tax burden is equal to the marginal external harm of output.*

Proof. Let τ be the tax on output. Then the private market's problem is:

$$\max_{x} b(x) - \tau q(x)$$

with first order condition $b'(x(\tau)) - \tau q'(x(\tau)) = 0$, where $x(\tau) : \mathbb{R} \to \mathbb{R}^n$ is the private market's best response function to τ.[29] Note that the effective marginal tax on an activity is simply the tax rate times the marginal product of each activity. The planner's problem is then:

$$\max_{\tau} b(x(\tau)) - e^\top x(\tau)$$

with first order condition $b'(x(\tau^*))x'(\tau^*) - e^\top x'(\tau^*) = 0$. Substituting in the private market optimum:

$$\tau^* q'(x(\tau^*))x'(\tau^*) = e^\top x'(\tau^*)$$

□

The optimal tax is a ratio of marginal harm to marginal product:

$$\tau^* = \frac{e^\top x'(\tau^*)}{q'(x(\tau^*))x'(\tau^*)} \tag{9}$$

[29] $x(\tau)$ is well defined if $b(x) - \tau q(x)$ is concave and there exists y and z such that $0 < y < z$, $\frac{\partial b}{\partial x_i}(y) > 0$ for all i, and $\frac{\partial b}{\partial x_i}(z) < 0$ for all i.

The numerator is the marginal external harm of a change in the output tax. The output tax will be larger when the activities that decrease the most in response to the output tax have large externalities. The denominator is the marginal change in output in response to a change in the tax rate. The output tax will have smaller magnitude if the activities that are most responsive to the output tax have large marginal products.

Example 3. Let x_1 be operating a natural gas plant and x_2 be operating a coal plant to produce electricity. Assume there is a tax τ levied on electricity. The private market maximizes $b(x_1, x_2) - \tau q(x_1, x_2)$ with first order conditions $\frac{\partial b}{\partial x_1} = \tau \frac{\partial q}{\partial x_1}$ and $\frac{\partial b}{\partial x_2} = \tau \frac{\partial q}{\partial x_2}$. The social planner's problem is:

$$\max_{\tau} b(x_1(\tau), x_2(\tau)) - e_1 x_1(\tau) - e_2 x_2(\tau)$$

with first order condition $\frac{\partial b}{\partial x_1}\frac{\partial x_1}{\partial \tau} + \frac{\partial b}{\partial x_2}\frac{\partial x_2}{\partial \tau} - e_1\frac{\partial x_1}{\partial \tau} - e_2\frac{\partial x_2}{\partial \tau} = 0$. Substituting in the private market optimum $\tau^* \frac{\partial q}{\partial x_1}\frac{\partial x_1}{\partial \tau} + \tau^* \frac{\partial q}{\partial x_2}\frac{\partial x_2}{\partial \tau} - e_1\frac{\partial x_1}{\partial \tau} - e_2\frac{\partial x_2}{\partial \tau} = 0$ The optimal tax is:

$$\tau^* = \frac{e_1\frac{\partial x_1}{\partial \tau} + e_2\frac{\partial x_2}{\partial \tau}}{\frac{\partial q}{\partial x_1}\frac{\partial x_1}{\partial \tau} + \frac{\partial q}{\partial x_2}\frac{\partial x_2}{\partial \tau}}$$

By assumption both coal plants and natural gas plants pollute ($e_1 > 0$ and $e_1 > 0$) and more electricity is produced if plants increase operations ($\frac{\partial q}{\partial x_1} > 0$ and $\frac{\partial q}{\partial x_1} > 0$). The signs of $\frac{\partial x_1}{\partial \tau}$ and $\frac{\partial x_1}{\partial \tau}$ are ambiguous, but at least one of them must be negative. A subsidy will be optimal if the use of coal increases with the tax ($\frac{\partial x_2}{\partial \tau} > 0$), coal produces a large externality relative to natural gas ($e_2 > e_1$), and the marginal product of natural gas is large relative to the marginal product of coal ($\frac{\partial q}{\partial x_1} > \frac{\partial q}{\partial x_2}$).

4.2 Comparing activity and output taxes

In order to make a welfare comparison between the optimal output tax and the optimal activity tax, we find a map from the output tax to activity taxes.

Proposition 7. *For any output tax the private market will respond as if there is a tax on each activity equal to the output tax times the marginal product of that activity.*

Proof. We want to find $t(\tau)$ such that $x(t(\tau)) = x(\tau)$. Recall that $b'(x(t)) = t^\top$ and $b'(x(\tau)) - \tau q'(x(\tau)) = 0$. Thus

$$t(\tau)^\top = \tau q'(x(\tau)) \tag{10}$$

$\square$

Substituting the optimal output tax derived above:

$$t(\tau^*)^\top = \frac{e^\top x'(\tau^*)}{q'(x(\tau^*))x'(\tau^*)} q'(x(\tau^*)) \tag{11}$$

The higher the marginal product of an activity, the higher the effective tax on that activity. This relationship makes sense because with a tax on output, an increase in an activity causes an increase in the tax burden proportional to the activity's marginal product.

In the appendix we show that the welfare lost, assuming no administrative cost, when choosing the optimal output tax instead of the optimal activity tax is:

$$\Delta b - e^\top \Delta x = (t(\tau^*) - e)^\top \left[\int_0^1 r\, x'(\gamma(r))\ dr \right] (t(\tau^*) - e) \tag{12}$$

where $\gamma(r) = e + r(t(\tau^*) - e)$. Note that the integral is the sum of negative definite matrices and that the expression is a quadratic form. Thus the optimal activity tax is always weakly better than any output tax.[30] Note that the welfare lost increases with the square of the uncorrected externality and increases with the weighted average size of $x'(t)$ over the set $[e, t(\tau^*)]$. $x'(t)$ is weighed by r for the same reason dead weight loss approximations are a triangle; close to the optimum the marginal benefit and marginal social harm are equal but diverge the further $t(\tau^*)$ is from e.

In the appendix we also derive a Harberger (1964)-style approximation of the welfare change caused by moving from one activity tax vector to another, and we apply that method here to compute the lost welfare, assuming no administrative cost, caused by moving from the optimal activity tax to the optimal output tax.

$$\Delta b - e^\top \Delta x \approx \frac{1}{2}(t(\tau^*) - e)^\top (x(e) - x(\tau^*)) \tag{13}$$

$$= \frac{1}{2}(\tau^* q'(x(\tau^*)) - e)^\top (x(e) - x(\tau^*)) \tag{14}$$

At the optimal activity tax there are no uncorrected externalities. At the optimal output tax the uncorrected externality is $(t(\tau^*) - e)$. Because of the approximation, the average uncorrected externality for each activity is half of the uncorrected externality under the output tax, $\frac{1}{2}(t(\tau^*) - e)$. Multiplying the uncorrected externality by the 'excess' amount of the activity that occurs under the output tax, $(x(e) - x(\tau^*))$ yields the total welfare loss.

[30] If the optimal activity tax is equal to marginal product times the optimal output tax, then both taxes achieve the same welfare because $t(\tau^*) = e$.

4.3 Variable administrative cost

Output taxes with variable administrative cost are similar to a single activity tax with variable administrative cost. If administrative cost is a function of output level, the output tax will be higher compared to the costless case. If administrative cost is a function of the output tax rate the optimal output tax will be lower than the costless case.

Let $d : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ map quantity and output tax rate to administrative cost. Formally d is continuously differentiable, weakly convex, and has $\arg\min(d) = (0,0)$. The private market problem is the same as in the costless output tax case. The social planner solves:

$$\max_{\tau} b(x(\tau)) - e^\top x(\tau) - d(q(x(\tau)), \tau)$$

with first order condition:

$$b'(x(\tau^*))x'(\tau^*) - e^\top x'(\tau^*) - d_1 q'(x(\tau^*))x'(\tau^*) - d_2 = 0$$

where d_i represents the derivative of d with respect to its i^{th} argument. Substituting the private market first order condition $b'(x(\tau^*)) = \tau^* q'(x(\tau^*))$ yields:

$$\tau^* = \frac{e^\top x'(\tau^*) + d_1 q'(x(\tau^*))x'(\tau^*) + d_2}{q'(x(\tau^*))x'(\tau^*)} \tag{15}$$

$$= \frac{e^\top x'(\tau^*)}{q'(x(\tau^*))x'(\tau^*)} + d_1 + \frac{d_2}{q'(x(\tau^*))x'(\tau^*)} \tag{16}$$

If administrative cost varies with output, the optimal output tax will be larger. If administrative cost varies with the output tax rate, the optimal output tax will be smaller. To see this note that $d_1 > 0$, $d_2 > 0$, and $q'x' < 0$.

If the problem includes both activity and output taxes the private market solves:

$$\max_{x} b(x) - t^\top x - \tau q(x)$$

With the most general form of variable administrative cost, the social planner solves:

$$\max_{t,\tau} b(x(t,\tau)) - e^\top x(t,\tau) - c(x(t,\tau), t) - d(q(x(t,\tau)), \tau)$$

However, the private market's problem does not yield a one-to-one best response function, $x(t,\tau)$—there are infinitely many t's and τ's that lead to the same x. Because of this the implicit expressions that come from the first order conditions of the social planner's problem are only informative about t and τ jointly. The end result is that little analysis can be performed.

5 Conclusion

Pigou's insight, that taxes can fully correct externalities, relies on the assumed insignificance of administrative cost and the assumption that all the externalities are subject to taxation. There are important settings in which these assumptions are unrealistic.

Relaxing these assumptions results in several policy implications. First, policymakers should determine the drivers of administrative cost. The existence of measuring costs (administrative costs that increase with activity levels) will tend to increase tax rates while enforcement costs (administrative costs that increase with tax rates) tend to decrease tax rates. Second, the policymaker should use the cheapest tax instruments that effect the greatest reduction in external harm. If fully correcting all externalities is too costly, policymakers should use output taxes or the complementarity of activities to cost-effectively reduce external harm.

Subsidies for harmful activities are impossible in classical Pigouvian analysis. Including administrative cost in the analysis helps reconcile this result with some real world policies. For example, needle exchanges are a form of subsidy to the use of heroin with clean needles. This activity is socially harmful but is certainly less socially harmful than the near-perfect substitute of heroin use with shared needles.

References

Atkinson, A. B., Stern, N. H., 1974. Pigou, taxation and public goods. The Review of Economic Studies, 119–128.

Atkinson, A. B., Stiglitz, J. E., 1976. The design of tax structure: direct versus indirect taxation. Journal of Public Economics 6 (1), 55–75.

Auerbach, A. J., 1985. The theory of excess burden and optimal taxation. Handbook of Public Economics 1, 61–127.

Auerbach, A. J., Hines, J. R., 2002. Taxation and economic efficiency. Handbook of Public Economics 3, 1347–1421.

Ballard, C. L., Fullerton, D., 1992. Distortionary taxes and the provision of public goods. Journal of Economic Perspectives 6 (3), 117–131.

Bovenberg, A. L., De Mooij, R. A., 1994. Environmental levies and distortionary taxation. The American Economic Review, 1085–1089.

Bovenberg, A. L., Goulder, L. H., 1996. Optimal environmental taxation in the presence of other taxes: general-equilibrium analyses. The American Economic Review, 985–1000.

Bovenberg, A. L., van der Ploeg, F., 1994. Environmental policy, public finance and the labour market in a second-best world. Journal of Public Economics 55 (3), 349–390.

Corlett, W. J., Hague, D. C., 1953. Complementarity and the excess burden of taxation. The Review of Economic Studies, 21–30.

Cremer, H., Gahvari, F., 1993. Tax evasion and optimal commodity taxation. Journal of Public Economics 50 (2), 261–275.

Cremer, H., Gahvari, F., 2001. Second-best taxation of emissions and polluting goods. Journal of Public Economics 80 (2), 169–197.

Cremer, H., Gahvari, F., Ladoux, N., 1998. Externalities and optimal taxation. Journal of Public Economics 70 (3), 343–364.

Dixit, A., et al., 1985. Tax policy in open economies. Handbook of Public Economics 1, 313–374.

Fullerton, D., 1997. Environmental levies and distortionary taxation: comment. The American Economic Review, 245–251.

Fullerton, D., Wolverton, A., 2005. The two-part instrument in a second-best world. Journal of Public Economics 89 (9), 1961–1975.

Goulder, L. H., Parry, I. W., Burtraw, D., 1997. Revenue-raising versus other approaches to environmental protection: The critical significance of preexisting tax distortions. The Rand Journal of Economics 28 (4), 708.

Green, J., Sheshinski, E., 1976. Direct versus indirect remedies for externalities. Journal of Political Economy, 797–808.

Harberger, A. C., 1964. The measurement of waste. American Economic Review, 58–76.

IRS, 2016. Budget in Brief FY2017.

Kaplow, B. L., 2012. Optimal control of externalities in the presence of income taxation. International Economic Review 53 (2), 487–509.

Kaplow, L., 1990a. A note on the optimal use of nonmonetary sanctions. Journal of Public Economics 42 (2), 245–247.

Kaplow, L., 1990b. Optimal taxation with costly enforcement and evasion. Journal of Public Economics 43 (2), 221–236.

Kaplow, L., 1996. The optimal supply of public goods and the distortionary cost of taxation. National Tax Journal, 513–533.

Kopczuk, W., 2003. A note on optimal taxation in the presence of externalities. Economics Letters, 81–86.

Lipsey, R. G., Lancaster, K., 1956. The general theory of the second best. The Review of Economic Studies 24 (1), 11–32.

Mayshar, J., 1991. Taxation with costly administration. The Scandinavian Journal of Economics, 75–88.

Ng, Y. K., 1980. Optimal corrective taxes or subsidies when revenue raising imposes an excess burden. The American Economic Review, 744–751.

Pigou, A. C., 1920. The economics of welfare, 1920. McMillan&Co., London.

Pirttilä, J., Tuomala, M., 1997. Income tax, commodity tax and environmental policy. International Tax and Public Finance 4 (3), 379–393.

Polinsky, A. M., Shavell, S., 1992. Enforcement costs and the optimal magnitude and probability of fines. Journal of Law and Economics, 133–148.

Polinsky, M., Shavell, S., 1982. Pigouvian taxation with administrative costs. Journal of Public Economics 19 (3), 385–394.

Ramsey, F. P., 1927. A contribution to the theory of taxation. The Economic Journal, 47–61.

Sandmo, A., 1975. Optimal taxation in the presence of externalities. Swedish Journal of Economics, 86–98.

Sandmo, A., 1978. Direct versus indirect pigovian taxation. European Economic Review 7 (4), 337–349.

Slemrod, J., 1990. Optimal taxation and optimal tax systems. The Journal of Economic Perspectives 4 (1), 157–178.

Slemrod, J., Yitzhaki, S., 2002. Tax avoidance, evasion, and administration. Handbook of Public Economics 3, 1423–1470.

Stiglitz, J. E., Dasgupta, P., 1971. Differential taxation, public goods, and economic efficiency. The Review of Economic Studies, 151–174.

Yitzhaki, S., 1979. A note on optimal taxation and administrative costs. The American Economic Review, 475–480.

A Comparing two activity taxes

Consider two arbitrary activity tax vectors, v and w. Assuming no administrative cost, the change in welfare of moving from v to w is:

$$\begin{aligned}\Delta b - e^\top \Delta x &= b(x(w)) - b(x(v)) - e^\top [x(w) - x(v)] \\ &= b(x(\gamma(1))) - b(x(\gamma(0))) - e^\top [x(\gamma(1)) - x(\gamma(0))]\end{aligned}$$

where $\gamma : \mathbb{R} \to \mathbb{R}^n,\ \gamma(r) = v + r(w - v)$. Then by the fundamental theorem of calculus:

$$= \int_0^1 [b'(x(\gamma(r))) - e^\top]\, x'(\gamma(r))\, \gamma'(r)\, dr$$

Recalling that $b'(x(t)) = t^\top$:

$$= \int_0^1 (\gamma(r) - e)^\top\, x'(\gamma(r))\, \gamma'(r)\, dr$$

Integrating by parts:

$$\begin{aligned}&= (\gamma(r) - e)^\top\, x(\gamma(r))\Big|_{r=0}^{1} \ - \ \int_0^1 \gamma'(r)^\top\, x(\gamma(r))\, dr \\ &= (w - e)^\top x(w) - \ (v - e)^\top x(v) \ - \ (w - v)^\top \ \int_0^1 x(\gamma(r))\, dr\end{aligned}$$

Assuming $\int_0^1 x(\gamma(r))\, dr \approx \frac{1}{2}(x(w) + x(v))$, which is analogous to Harberger (1964)'s approximation:

$$\begin{aligned}&= (w - e)^\top x(w) - \ (v - e)^\top x(v) \ - \ \frac{1}{2}(w - v)^\top (x(w) + x(v)) \\ &= \left[(w - e) - \ \frac{1}{2}(w - v)\right]^\top x(w) - \ \left[(v - e) \ + \ \frac{1}{2}(w - v)\right]^\top x(v) \\ &= \frac{1}{2}(w + v - 2e)^\top (x(w) - x(v))\end{aligned}$$

If $v = e$ (i.e. the social planner shifts away from the optimal activity tax), then:

$$\Delta b - e^\top \Delta x = \int_0^1 r(w - e)^\top\, x'(\gamma(r))\, (w - e)\, dr \tag{17}$$

Note that the integral is the sum of negative definite matrices and that the expression is a quadratic form. Thus the optimal activity tax is strictly superior to any other activity tax. Note that the welfare lost increases with the square of the uncorrected externality and increases with the weighted average size of $x'(t)$ over the set $[e, w]$. $x'(t)$ is weighed by r because close to the optimum the marginal benefit and marginal social harm are equal but diverge more the further w is from e—for the same reason dead weight loss approximations are a triangle.

As an aside, this appendix generalizes the *Harberger Triangle* to a setting with more than one taxed good and an initial set of taxes.[31] Assuming there are no externalities, the approximate change in welfare from moving to one tax vector to another is:

$$\Delta b \approx \frac{1}{2}(w+v)^{\top}(x(w)-x(v)) \tag{18}$$

This equation corresponds to the *Harberger Trapezoid* associated with a change from an existing set of taxes to another set of taxes. Setting $v = 0$ simplifies this equation to the Harberger Triangle associated with a change from no taxes to a new set of taxes, in a setting with more than one good and more than one tax:

$$\Delta b \approx \frac{1}{2}w^{\top}(x(w)-x(0)) \tag{19}$$

[31] See generally Auerbach (1985) and Auerbach and Hines (2002).