Reconciling the Pigovian and Sandmo Principles of Emission Pricing

Ross McKitrick
Department of Economics and Finance
University of Guelph
Guelph ON Canada
N1G 2W1

ross.mckitrick@uoguelph.ca

March 2021

Key words: Emission taxes; Pigovian rule; tax interactions; damage thresholds; climate policy

JEL Codes: H21, H23, Q54, Q58

Acknowledgments: I am grateful to seminar participants at the Southern Economics Association, the University of British Columbia, the University of Victoria and the Canadian Resource and Environmental Economics Study Group for helpful comments.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of interest: none.

Highlights:

- The Sandmo second-best rule for emission pricing yields some policy paradoxes
- The Pigovian emission tax rule only holds in economies with no taxes
- Optimal emission taxes may vary by country even if marginal damages do not

- Regulation may reduce welfare even when marginal damages are positive
- The welfare ranking of marginal regulations and taxes must reverse at some point

Abstract: The classical Pigovian rule is attractively simple: polluters should pay the marginal damages (MD) of emissions, computed at the point where MD equals marginal abatement costs (MACs). In the context of climate policy this implies using the Social Cost of Carbon (SCC) as the marginal benefit criterion in regulatory cost-benefit analysis and in emission charge-setting. But beginning with Sandmo (1975), numerous economists have shown that in second-best economies the Pigovian rule is sub-optimal, and the optimal tax rate should be adjusted by a factor related to the marginal welfare cost of the tax system. While this point is theoretically well-established and has major implications, it is generally ignored in practice, which may be due to its apparent conflict with the intuitive Pigovian principle. Here I show that the two principles can be unified by noting that tax distortions drive a wedge between firms' private MACs and the social costs of abatement, and the Sandmo rule compensates for the difference. I provide a graphical exposition then present a formal derivation in general equilibrium. I also discuss some of the practical implications and surprising paradoxes created by the Sandmo rule. For instance the optimal emission tax will typically differ between jurisdictions even when MD is constant, which has important implications for carbon border tax adjustments. It also helps explain why cost-benefit analysis might conclude emissions should remain unregulated even when the optimal emission tax is positive, and why regulations can sometimes be marginally less costly than taxes.

1 INTRODUCTION

The standard Pigovian model states that an ideal emissions tax should equal the marginal social damages of emissions (denoted *MD* herein) evaluated at the optimal pollution level, which occurs where *MD* equals marginal abatement costs (*MAC*). This rule also implies that optimal *MD* is the appropriate metric to use in evaluating the benefits of environmental regulation. It is typically presented in textbooks based on partial equilibrium analysis, under the assumption that an emission tax neither affects nor is affected by the rest of the tax system. In the presence of pre-existing taxes however, revenue from a pollution charge can fund tax reductions elsewhere, while pollution taxes (or emission regulations) increase the marginal distortions associated with the overall tax system, which are known, respectively, as the revenue recycling and tax interaction effects. A result originally due to Sandmo (1975) is that the optimal emission tax taking these additional effects into account equals *MD* deflated by the marginal cost of public funds (*MCPF*, Sandmo 1975, Bovenberg and Goulder 1996, Goulder 1998, Parry et al. 1999, Schöb 2003), a result I refer to herein as the Sandmo rule. It implies that the optimal emissions charge should be lower than *MD* and the optimal emissions level typically exceeds the point where *MD* equals *MAC*, aka the Pigovian level.

Although the Sandmo pricing principle has long been known in the economics literature (see reviews in Goulder 1998, 2013) it is rarely referenced in either environmental economics textbooks or in discussions of emission pricing. For example the US InterAgency Working Group report (2013) on the Social Cost of Carbon made no reference to Sandmo-type adjustments. One of the key implications of the Sandmo rule is that the correct target for an emission tax is not the *MD* itself, but *MD* normalized by a jurisdiction's *MCPF*. This in turn implies that even when *MD* is constant across a large region (as the Social Cost of Carbon is assumed to be since carbon dioxide mixes globally), the optimal emission price should vary by tax jurisdiction and the variations can be substantial. Dahlby and Ferede (2018) estimate that even among Canadian provinces the *MCPF* of personal income taxes vary from a low of 1.4 to a high of 6.8. Additionally the optimal emission tax rate can depend even on the degree to which the *MCPF* differs within a jurisdiction across different factor taxes, and the extent to which emission tax revenue recycling widens or narrows the gap (see examples in Goulder 1998). Such considerations have important, but thus far overlooked, implications for border tax adjustments or other trade-related instruments to adjust for variations in climate policy stringency among nations. Finally the introduction of second-best elements in optimal emission models can imply different policy targets depending on whether taxes or regulations are used, such that cost-benefit analysis could conclude that emissions should be unregulated even if an optimal emissions tax would be positive.

The apparent inconsistency between the Pigovian and Sandmo rules can be reconciled by noting that the customary derivation of the *MAC* (e.g. McKitrick 1999) yields *private* marginal abatement costs, whereas *MD* is a social disutility. The proper optimal emission target is where social *MAC* crosses the *MD* line. The Sandmo rule compensates for the difference between private and social *MAC*s. I explain this herein in a simple graphical model and then elaborate on the results using a general equilibrium derivation. The graphical approach also clarifies why a corner solution can arise. As has been previously noted (Bovenberg and Goulder 1996, Goulder et al. 1997, Goulder 1998 and 2013, Bento and Jacobsen 2007) there can exist a positive damage threshold which *MD* must exceed for the first unit of abatement to yield a social welfare improvement, although this does not emerge in the Sandmo model. Bovenberg and Goulder (1996) and Goulder et al. (1997) respectively estimated that, in a model calibrated to the US economy, *MD* would have to be relatively high in the case of carbon dioxide (over $50 US per ton in 1996 dollars) and sulfur dioxide (over $100 per ton in

1997 dollars) for any non-revenue raising policy to be welfare-improving. Fullerton and Metcalf (2001) argued that the key issue is not revenue-raising *per se* but whether the policy creates scarcity rents that are left in private hands. If another fiscal instrument captures the rents and uses them to offset tax interactions the threshold effect may disappear.

Various other challenges to the Pigovian rule have been noted over the years. Turvey (1963) noted that if emissions are subject to prior regulations or Coasian bargaining the Pigovian rule will not lead to an optimum. Baumol and Oates (1988) noted that non-competitive market distortions also cause the Pigovian rule to break down. In sum, the environmental economics literature has shown that in many cases the simple Pigovian rule is not optimal and can even be welfare-reducing depending on the context in which it is applied. Nevertheless it holds such sway over emission pricing discussions and cost-benefit analysis that the rarity of its applicability is not widely-noted. In particular, distortions due to pre-existing taxes are ubiquitous and ought to be routinely taken into account. The Sandmo rule yields a relatively simple treatment for this case which deserves to be to be more widely understood and used.

The next section provides a graphical summary of the main argument. Section 3 sets up the theoretical model and derives the basic results. Section 4 offers additional points of discussion and Section 5 concludes.

2 GRAPHICAL MODEL

To begin with, note that *MCPF* is the value of the marginal welfare loss (measured in dollars) to the private sector arising from a one dollar increase in the revenue requirements of the public sector. For example an *MCPF* of 2.0 means the private sector loses two dollars' worth of economic welfare in order to increase public spending by one dollar. We typically define the welfare loss using the Hicksian Equivalent Variation. When an emissions tax is introduced, prices rise and the consumer's utility falls. If instead households were simply assessed a lump sum charge that resulted in the same reduction in utility the lump sum would be called the Equivalent Variation and would provide a monetary measure of the change in utility caused by the emissions tax. If that quantity were expressed per dollar of new revenue from the emission tax, the result would be the $MCPF$. Sandmo (1975) defined the $MCPF$ slightly differently, namely as the ratio of two Lagrange multipliers which defined, respectively, the marginal utility of private income and the marginal disutility of raising the public revenue requirements. The precise definition of $MCPF$ emerges from the specific set-up of a model, which is why it can vary by context. But it always seeks to measure the same thing: the loss in economic welfare from increasing the public budget requirement.

We proceed by drawing a distinction between private marginal abatement costs, denoted MAC_p, and social marginal abatement costs, denoted MAC_s. The former corresponds to the firm's marginal profits from emitting activity. The latter denotes the marginal social welfare costs associated with a requirement to reduce emissions. Expressed in this way MAC_s will not be invariant to the form of the policy. We will denote the emissions below using the letter C so we will call the emission tax τ_C and assume it is introduced alongside full recycling of revenue into tax reductions elsewhere. Define the tax as an affine transform of marginal damages: $\tau_C = aMD - b$ where $a > 0$ and b is a parameter to be determined. Profit-maximizing firms will respond to such a tax by cutting emissions to the point where $\tau_C = MAC_p$. The social optimum occurs where $MAC_s = MD$. Combining these yields

$$MAC_s = \frac{1}{a} MAC_p + \frac{b}{a}. \tag{1}$$

Figure 1 presents a case in which we assume *MD* is horizontal, for simplicity. The classical Pigovian rule assumes $a = 1$ and $b = 0$ which implies $\tau_C = MD$. The resulting optimum occurs where $MD = MAC_P$ at emissions E_P which is strictly less than the unregulated emissions level $\bar{E}$.

The analysis of Sandmo (1975) yielded $a = MCPF^{-1}$ and $b = 0$. Since $MCPF > 1$ in a second-best economy (namely an economy in which the tax system imposes marginal distortions) this implies $0 < a < 1$ which yields a clockwise rotation of MAC_p to MAC_{1S}. Now the social optimum occurs where $MAC_{1S} = MD$ which is at $E_1 > E_P$. But the firm's private optimum occurs where the emission price equals MAC_p. Consequently the emission price needs to be scaled down and the appropriate scaling factor is $aMD = MD/MCPF$.

The numerical analyses in Bovenberg and Goulder (1996) and the analytical models in Goulder et al. (1997) and Bento and Jacobsen (2007) yielded cases where $b > 0$. This implies social marginal abatement costs are not only a rotation but also a translation of MAC_p, yielding MAC_{2S}. As drawn, *MD* crosses MAC_{2s} at an emissions level above the unregulated level $\bar{E}$, which implies emissions should be subsidized. If we additionally impose a non-negativity constraint on τ_P then we obtain a threshold *Z*, which is the value of *MD* below which the optimal emissions tax would remain zero even when *MD* is positive. A positive value of *b* can arise for a variety of reasons. In Goulder (1998) it occurs if there are no revenue-recycling benefits and the tax interaction effect exceeds marginal damages at the unregulated emissions level. It also arose in Bovenberg and Goulder (1996) because revenue recycling via lump-sum transfers was permitted. In an optimizing context, if lump-sum instruments are available it will be optimal to use them for taxation and apply the proceeds to reducing distorting taxes, even to the extent of using a negative emissions tax to subsidize pollution. Applying a rule that emission taxes must be non-negative creates the corner solution with the positive threshold Z as shown. Since lump-sum revenue-recycling is functionally similar to use of regulatory measures, we will see in a subsequent section that use of command-and-control regulations implies a threshold effect.

Summarizing thus far, if the tax system imposes distortions such that $MCPF > 1$ (which is always the case) then full revenue-recycling implies the optimal emissions charge is $MD/MCPF$. If revenues are not recycled, or if emissions control is accomplished using non-revenue-raising policies, the optimal emission charge is shifted lower and may hit a zero bound even if marginal damages are positive. In the next section we re-cap these results in the context of a theoretical general equilibrium model, then explore additional variations. As is often the case in second-best economic analysis, the results can easily become surprising and counterintuitive.

3 THEORETICAL MODEL

3.1 BASIC SET-UP

Consider an economy with N identical households, a consumption good x with corresponding price p_x, an energy sector, a labour market and a government. Aggregate consumption is denoted $X = Nx$. Households also consume e_{hh} units of energy at price p_E and aggregate household energy demand is $E_{hh} = Ne_{hh}$. Households each have a time endowment t which can be allocated to labour l or leisure h, so the aggregate labour supply is $L = Nl$, aggregate leisure is $H = Nh$ and the aggregate time endowment is $T = Nt$. The before-tax nominal wage rate is w. We will set w as the numeraire so it is constant and equal to unity but for notational clarity I retain it in the derivations.

Energy is produced by a sector that uses only labour L_E so its production function is $E = F^E(L_E)$ and its profits are $\pi^E = p_E F^E(L_E) - wL_E$. Use of energy causes emissions $C = cE$ where c denotes the emissions intensity of energy and is assumed constant. Emissions are taxed at τ_C per unit so the tax-inclusive price of energy is

$$p'_E = p_E + c\tau_C. \tag{2}$$

Note that $'$ throughout denotes a tax-inclusive term.

The goods-producing firm has a single unit of fixed capital K_x equal to unity and a production function $F(L_x, E_x)K_x$ where the first argument denotes labour usage and the second denotes energy. As in Bento and Jacobsen (2007) assume that F is strictly concave and has decreasing returns to scale in L_x and E_x. The profit function for the consumption good firm is

$$\pi_x = p_x F(L_x, E_x) - wL_x - p'_E E_x$$

where $F_L > 0$ and $F_E > 0$. The first-order conditions imply $F_L = w/p_x$ and $F_E = p'_E/p_x$. Decreasing returns to scale imply that profits are positive and represent the return to capital for each firm. We assume shares in firms are distributed equally among all households.

We will assume that prices are initially normalized so that $p_x = p_E = w = 1$. The tax rate on household income is τ_y and net income is

$$y' = \left(\frac{\pi_x + \pi_E}{N} + wt\right)(1 - \tau_y). \tag{3}$$

The household budget constraint is $p_x x + p'_E e_{hh} + w'h = y'$ where $w' = w(1 - \tau_y\ )$. The corresponding national budget constraint (NBC) is

$$p_x X + p'_E E_{hh} + w'H = Y' \tag{4}$$

where $Y' = (\pi_x + \pi_E + wT)(1 - \tau_y)$.

The government does not use energy but provides a transfer G distributed equally to all households. It finances this through the tax τ_C on emissions C and the income tax τ_Y. Hence the Government Budget Constraint (GBC) is

$$G = \tau_y B + \tau_C C \tag{5}$$

where the income tax base B equals $\pi + wL$ and $\pi = \pi_x + \pi_E$.

Goods Market Equilibrium (GME) occurs where $X = F_x$. Energy Market Equilibrium (EME) occurs where $E_{hh} + E_x = E$. Labour Market Equilibrium (LME) occurs where $L_x + L_E = T - H$. It is straightforward to show that imposing GME, LME and EME on the NBC implies the GBC holds; likewise any four implies the fifth.

We assume that tax rates are adjusted to hold G constant. Differentiating Equation (5) and rearranging yields the revenue-neutral tax swap rule

$$\frac{d\tau_y}{d\tau_C} = -\frac{1}{B}\left(\tau_y \frac{dB}{d\tau_C} + \tau_C \frac{dC}{d\tau_C}\right). \tag{6}$$

We will assume that we are operating in a region of the economy for which the new tax revenue (represented by the second term) is sufficiently large as to make whole derivative negative, meaning that an increase in emission taxes finances a reduction in the income tax.

Household utility is $u(x, e_h, h) - \delta C$ where δ is the marginal disutility of each unit of emissions. We will use the indirect utility function $v(p_x, p_E, w', y')$ to define the national social welfare function

$$W = Nv(p_x, p_E', w', y') - \delta NC. \tag{7}$$

Note that by the envelope theorem, $\frac{d\pi_x}{d\tau_C} = -cE_x$, which can be thought of as defining the "demand" curve for emissions. Differentiating equation (7), then applying Roy's Theorem, equations (5) and (6) and collecting terms (See Appendix) yields the marginal welfare cost of the emissions tax

$$\frac{dW}{d\tau_C}\frac{1}{v_y} = -X\frac{dp_x}{d\tau_C} + \tau_y w\frac{dL}{d\tau_C} + \frac{dC}{d\tau_C}\left(\tau_C - \frac{\delta N}{v_Y}\right) \tag{8}$$

The term on the left side is the marginal utility from varying the emission tax, converted to a money measure by dividing by the marginal utility of income. The expression on the right side decomposes the welfare effect into three standard components (compare to Goulder 1998, Parry et al. 1999, and Bento and Jacobsen 2007, although note that these retain the labour tax whereas we have here imposed the GBC and substituted it out). The third term contains the difference between the emission tax and the marginal social costs of emissions ($\delta N/v_y$). The first two terms represent, respectively, the tax interaction effects and the revenue recycling benefit. If these disappear or exactly offset each other, we will want the third term to go to zero, which gives us the first best Pigovian outcome.

3.2 The Optimal Emission Tax

Our aim will be to define MD in terms of the model and then derive the coefficients of an affine transformation such that the optimal emission tax can be written $\tau_C = aMD - b$ where $a > 0$ and the sign of b is to be determined. The planner's problem is to maximize (7) with respect to τ_C subject to the non-negativity constraint $\tau_C \geq 0$, which means that we do not permit the regulator to subsidize emissions even if doing so would be optimal. The Lagrangian function is

$$\mathcal{L} = Nv(p_x, p_E', w', y') - \delta NC + \lambda\tau_c \tag{9}$$

where λ is the multiplier on the inequality constraint. The Kuhn-Tucker conditions are

$$v_y\left(-X\frac{dp_x}{d\tau_C} + \tau_Y w\frac{dL}{d\tau_C} + \tau_C\frac{dC}{d\tau_C} - \frac{\delta N}{v_Y}\frac{dC}{d\tau_C}\right) + \lambda = 0$$

$$\lambda \geq 0; \tau_C \geq 0;\ \lambda\tau_C = 0$$

Denote the optimum values using *. If $\lambda > 0$ and $\tau_C = 0$ then $-X\frac{dp_x}{d\tau_C} + \tau_Y w\frac{dL}{d\tau_C} - \frac{\delta N}{v_Y}\frac{dC}{d\tau_C} = -\frac{\lambda}{v_y} < 0$. Applying this to equation (8) implies $\frac{dW}{d\tau_C} < 0$. Since the welfare function is concave the implication is that at the corner solution where the non-negativity constraint binds, we have an underlying optimal value of $\tau_C^* < 0$ but we restrict the outcome to a non-negative tax rate.

A further rearrangement yields $X\frac{dp_x}{d\tau_C} - \tau_Y w\frac{dL}{d\tau_C} = \frac{\lambda}{v_y} - \frac{\delta N}{v_Y}\frac{dC}{d\tau_C}$. Since emissions are declining in τ_C the right hand side is strictly positive at the corner solution, which implies

$$\tau_C^* = 0 \Rightarrow X\frac{dp_x}{d\tau_C} - \frac{G}{B}w\frac{dL}{d\tau_C} > 0 \qquad (10).$$

This result will be useful below when characterizing the damage threshold.

If $\lambda = 0$ and $\tau_C > 0$ then the Kuhn-Tucker conditions imply (see Appendix)

$$\tau_C^* = aMD - b \qquad (11)$$

where

$$a = \frac{\frac{dC}{d\tau_C}}{\frac{dC}{d\tau_C} - \frac{C}{B}w\frac{dL}{d\tau_C}} \qquad (12)$$

and

$$b = -\frac{1}{\frac{dC}{d\tau_C} - \frac{C}{B}w\frac{dL}{d\tau_C}}\left(X\frac{dp_x}{d\tau_C} - \frac{G}{B}w\frac{dL}{d\tau_C}\right) \qquad (13)$$

We can show that the possibility of positive revenue recycling implies $0 < a < 1$ as follows. Note that the derivative of labour in the denominator of a is with respect to τ_C not τ_y. It could be broken up into its constituent partial derivatives, since L is a function of p_x, p_E', w' and τ_y. The emissions tax will cause the prices of the consumption good and energy to increase, which will cause the real labour supply to decrease. But it will also fund a reduction in the income tax rate, which will lead to an increase in the labour supply. We assume that the economy is operating where the latter revenue-recycling effect is sufficiently large to yield a net increase in labour, therefore $\frac{dL}{d\tau_C} > 0$. Since $\frac{dC}{d\tau_C} < 0$ it must be the case that $0 < a < 1$.

In Sandmo (1975), a corresponds to the ratio of the marginal utilities of public revenue and private income or $MCPF^{-1}$. A similar interpretation arises here which we can see by multiplying the top and bottom by τ_C and using the definitions of L and C, yielding

$$a = \frac{\tau_C c \dfrac{dE}{d\tau_C}}{\tau_C c \dfrac{dE}{d\tau_C} + \tau_C \dfrac{C}{B} w \dfrac{dH}{d\tau_C}}.$$

Recall from equation (2) that $c\tau_C$ is the wedge between the supply price of energy and the marginal willingness to pay, hence the numerator is the marginal welfare loss associated with a reduction in energy consumption due to an incremental increase in the emissions tax for the purpose of funding additional government spending. The same term appears in the denominator. The second term is the decline in leisure due to the emission tax, weighted by $\tau_C Cw/B$. To understand this term, note that solving the GBC for τ_Y would break it down to two components: G/B and $-\tau_C C/B$. The first is the portion required to cover government spending and the second is the offsetting reduction permitted by emission tax revenues. If government spending were zero but marginal damages necessitated $\tau_C > 0$ we could use the emission tax revenue to subsidize labour at the rate $-\tau_C C/B$. This term therefore represents the opportunity cost of needing to fund G. Hence the second term is the marginal decline of leisure weighted by the nominal wage rate times the portion of the income tax rate that represents the opportunity cost of needing to fund the government. Consequently, the denominator of a is the marginal (with respect to τ_C) opportunity cost of financing government spending through τ_C. The inverse of a is therefore this amount relative to the direct economic cost of the emission tax increase, giving a an interpretation similar to the inverse-MCPF weights found in previous models.

The coefficient b consists of the denominator of a multiplied by the term in equation (10). At the corner solution therefore $b > 0$. A damage threshold effect cannot arise in a Pigovian case or in the Sandmo model in which $\tau_c^* = aMD$. Nor did it arise in the Bovenberg and Goulder (1996) theoretical model, though it emerged as a property of their numerical model. In the present case, we can denote the threshold magnitude Z and derive it by setting $\tau_c^* = aMD - b = 0$ and solving for $Z = b/a$, or

$$Z = -\left(X\frac{dp_x}{d\tau_C} - \frac{G}{B}w\frac{dL}{d\tau_C}\right)/\frac{dC}{d\tau_C} \tag{14}$$

which represents the level of MD which observed damages must exceed for any positive emission tax to be welfare-improving. Equation [10] implies $Z > 0$.

Note that, *ceteris paribus*, the less price-elastic emissions are the higher will be the threshold which MD must exceed for emissions policy to be welfare-enhancing. It is also apparent from equation (14) that the threshold is increasing in G and decreasing in the size of the tax base B. Furthermore, note that in the Sandmo (1975) framework, if the government revenue requirement is low enough to be fully satisfied by an externality tax, the optimal policy would be a tax on the dirty good equal to marginal social damages and no other tax. But here if we set $\tau_Y = 0$ and set equation (8) to zero we obtain

$$\tau_C = MD + X\frac{dp_x}{d\tau_C}\left(\frac{dC}{d\tau_C}\right)^{-1}.$$

The second term is the marginal loss of consumer surplus in the market for X due to the introduction of the emission tax, per unit by which emissions are reduced by the tax. Since this is negative the optimal tax rate is strictly less than MD. In the Sandmo framework tractability requires all cross-price derivatives to be set equal to zero and all production is based on fixed input-output functions in which producer prices do not vary. If prices were similarly fixed herein the second term would disappear and the classical solution would emerge.

4 Some Second-Best Policy Paradoxes

4.1 Sufficient Conditions for the Pigovian Rule

We can use equations (11-13) to identify sufficient conditions for the classical Pigovian rule to hold. As a general rule, in any economy complex enough to impose Pigovian taxes, the Pigovian rule cannot hold. On the assumption that unregulated emissions are positive we require $\frac{dL}{d\tau_C} = 0$ and $\frac{dp_x}{d\tau_C} = 0$ in order to obtain $\tau_C^* = MD$ and $Z = 0$. The first condition could occur if the labour supply is assumed fixed. Alternatively it could occur if the labour supply is unresponsive to energy and consumption good price changes and the emission tax revenue is not used to fund a reduction in the income tax rate. But in the latter case the government budget balance will change. If we also stipulate that G is fixed we then have a contradiction. If we allow G to vary then it is a choice parameter and we need to re-solve for the optimum. The model would prescribe raising revenue using $G < 0$ as a lump-sum tax with the proceeds used to subsidize labour and energy consumption. If we impose the additional requirement that $G \geq \bar{g}$ where $\bar{g}$ is a fixed parameter we simply end up back where we started with fixed G. Therefore the first condition can only arise if the labour supply is assumed fixed.

The second condition could arise only if the supply of X is infinitely elastic or if some equivalent restriction is imposed that makes p_x unresponsive to the emission tax. Consequently we would only observe the Pigovian outcome in an economy in which there are no pre-existing tax distortions, and both consumption prices and labour supply are fixed. Among other things this would rule out the presence of governments or labour markets, and such an economy would be too primitive to implement emission taxes anyway. Once a government and a labour market are present the conditions for the Pigovian rule break down.

Interestingly, in this model, unlike those in Sandmo (1975) or Bovenberg and Goulder (1996), it is much harder to find a configuration that yields $0 < a < 1$ and $b = 0$. Such an outcome would require $X \frac{dp_x}{d\tau_C} = \frac{G}{B} w \frac{dL}{d\tau_C}$. If we are starting at the point where $\tau_C = 0$ we have $\tau_y = G/B$ so the condition equates to

$$X \frac{dp_x}{d\tau_C} = \tau_y w \frac{dL}{d\tau_C}. \qquad (15)$$

The left hand side is approximately the increased cost of consumption due to the marginal increase in the emission tax. The right hand side is the marginal revenue of the labour tax induced by a change in the emission tax. Only where these exactly coincide would $b = Z = 0$.

4.2 TAXING EMISSIONS THAT ARE ALREADY REGULATED

To the extent emission taxes are used they are typically introduced after emissions have already been subject to regulation. Suppose that prior to introducing an emissions tax, the regulator selects an emissions cap $\hat{C}$ to optimize welfare. The optimum occurs where (see Appendix)

$$MAC_p = MD + X\frac{dp_x}{dC} + E_{hh}\frac{dp_E}{dC} - \tau_y w\frac{dL}{dC}. \qquad (16)$$

Since compliance with a lower emission cap is costly the price derivatives are negative, while the derivative with respect to labour is positive. Therefore the optimum occurs where $MAC_p < MD$, which is above the Pigovian emissions level.

The threshold effect arises in this case as follows. As shown in the Appendix, the marginal welfare effect of tightening the emissions cap is

$$\frac{dW}{dC}\frac{1}{v_y} = -X\frac{dp_x}{dC} - E_{hh}\frac{dp_E}{dC} + \tau_y w\frac{dL}{dC} + MAC_p - MD.$$

At the unregulated emissions level ($MAC_p = 0$) to get an increase in W from a reduction in emissions (that is, $dW/dC < 0$) requires

$$\frac{dW}{dC}\frac{1}{v_y} = -X\frac{dp_x}{dC} - E_{hh}\frac{dp_E}{dC} + \tau_y w\frac{dL}{dC} - MD < 0 \Rightarrow MD > -X\frac{dp_x}{dC} - E_{hh}\frac{dp_E}{dC} + \tau_y w\frac{dL}{dC} > 0.$$

Thus MD must exceed a positive threshold level for the first unit of emission reduction to be welfare-improving. Optimal regulatory stringency may therefore be zero (i.e. emissions should be unregulated) even though $MD > 0$.

A paradoxical result can arise if a welfare-reducing regulation has been put in place capping emissions at $\hat{C}$. If an emission tax is then introduced at a rate below the shadow price associated with $\hat{C}$, in other words low enough that emissions do not change, such a tax will be strictly welfare-improving. Since prices, MAC_p and MD will remain constant the only change will be the labour supply effect resulting from a reduction in the income tax rate. By equation (8) we will therefore have a marginal welfare effect of

$$\frac{dW}{d\tau_C}\frac{1}{v_y} = \tau_y w\frac{dL}{d\tau_C} > 0.$$

Hence setting the emission cap may be welfare reducing if *MD* is less than the marginal welfare cost of the first unit of abatement. However, once the cap is in place, the emission tax is (initially) a rent-capture mechanism and raises welfare if used to fund a reduction in the income tax rate.

4.3 Saturated Labour Market Effect

We have assumed that the emission tax yields sufficient revenue to fund a reduction in the labour tax so that $\frac{dL}{d\tau_C} > 0$. But at a certain point the contractionary effect on the labour supply of the emission tax will fully offset the expansionary effect from revenue recycling and $dL/d\tau_C$ will go to zero. Equations (12) and (13) then imply that $a = 1$ and $b = Z = -X\frac{dp_x}{d\tau_C} / \frac{dC}{d\tau_C}$. The latter is strictly positive. This implies that, at this point, the policy has caused a parallel shift from MAC_p to MAC_S rather than a rotation as in Figure 1. It might seem counterintuitive that a rises in comparison to the Sandmo case. Equation (12) shows that a varies in a potentially complex way as the emission tax rises, based on changes in emissions C, the tax base B and the labour supply effect. It is only locally in the neighbourhood of $dL/d\tau_C = 0$ that a goes to unity. It is similar to the case described in Section 4.1 except that while labour is fixed prices are not.

4.4 Going to Zero: Taxes versus Regulations

Finally, as pointed out in Goulder (1998), if the aim is to drive emissions to zero, it does not matter whether taxes or regulatory instruments are used. Once emissions are zero there are no tax revenues to recycle, consequently the total costs will be the same between the two types of policy. But since the marginal welfare costs at low levels of abatement are higher for regulatory instruments, this implies the existence of a crossing point beyond which the marginal welfare costs of the tax policy must exceed regulations (or tradable quotas). In the current model that is the transition point identified in the previous section, where the labour supply no longer grows in response to the emission tax swap. Beyond that point, further tightening of an emission standard raises prices but does not raise the labour tax as much as would a further increase in the emissions tax. Consequently the marginal welfare advantage of emission taxes does not extend the whole way to zero emissions.

5 Conclusions and Policy Implications

The Pigovian rule prescribes an emissions tax equal to marginal damages. In the climate change context this would imply imposing a carbon tax equal to the Social Cost of Carbon. The rule is also routinely interpreted to mean that *MD* is the appropriate metric for computing benefits in a cost-benefit analysis. But as we have shown this is a special case that relies on exceptionally strong assumptions. The more general version is the Sandmo rule, in which the optimal price is *MD* deflated by the (local) marginal cost of public funds, which implies the optimal emissions are at level higher than the Pigovian outcome, and the optimal emission tax is lower than *MD*. The two results can be reconciled by noting that firms respond to the emission price according to their private *MAC*, but the social optimum is determined by the social *MAC*. The Sandmo rule compensates for the difference between these amounts.

Taking account of distortions in the tax system moves the discussion into the world of second-best analysis, which is notorious for yielding counter-intuitive and unexpected changes to first-best prescriptions. For instance it might seem odd that, the more burdensome is a country's tax system, the less stringently it should price pollution emissions. One might validly wonder what one has to do with the other, and moreover why, if the tax system is relatively more burdensome, it wouldn't argue for greater reliance on 'green' taxes rather than less. The key, as noted in Sandmo (1975), is that pollution control is like a public good. It stands to reason that an economy with a costlier tax system will have to settle for less provision of public goods compared to one with a more efficient and less costly tax system. If the public good in question is a reduction in pollution externalities, the same reasoning applies.

An implication of this analysis for global climate policy is that even though the Social Cost of Carbon is the same globally the optimal carbon price is not uniform across countries, nor even across tax jurisdictions within a country. Neither, therefore, should optimal climate policies be equally stringent across countries. This has major, but hitherto overlooked, implications for computing border tax adjustments, by which a country with a price on greenhouse gas emissions proposes to charge import tariffs on other countries which either have no such tax or impose it at a lower rate. Computing the appropriate level of such a tariff would need to take into account the *MCPF* in both jurisdictions. It is not automatically the case that if one country has a lower carbon price than another, or a relatively lax emissions policy, it therefore has an unfair advantage and should face punitive trade measures. It would be easy, in fact, to construct a case in which one jurisdiction with a low carbon tax would be justified in imposing a border tax on another jurisdiction with a higher carbon tax.

This is not the only paradox that arises once we have reconciled the Pigou and Sandmo principles. Because the position and slope of the social *MAC* is not invariant to the policy instrument, a case can arise in which an optimal emission tax would be positive, implying an optimal emission level below the unregulated level, yet the optimal regulation would be to leave emissions unregulated. The paradox deepens when we consider adding a tax on top of pre-existing regulations. If a welfare-reducing regulation has been imposed it may be possible to reduce the welfare loss by introducing an emissions tax, as long as the tax does not have any effect on emissions. Finally, if the target of the policy is to force emissions to zero it does not matter whether taxes or regulations are used, however along the path it must be the case that at a certain point the marginal welfare cost of emission taxes exceeds those of emission regulations, which is opposite to the case for the first unit of reductions.

6 REFERENCES

Baumol, William and Wallace Oates (1988) *The Theory of Environmental Policy*, 2nd ed. Cambridge: Cambridge University Press.

Bento, Antonio and Mark Jacobsen (2007) "Ricardian Rents, Environmental Policy and the 'Double-Dividend' Hypothesis." *Journal of Environmental Economics and Management* 53 pp. 17—31.

Bovenberg, A. Lans and Lawrence H. Goulder (1996). "Optimal Environmental Taxation in the Presence of Other Taxes: General-Equilibrium Analyses." *American Economic Review* 86(4) 985—1000.

Dahlby, Bev and Ergete Ferede (2018) "The Marginal Cost of Public Funds and the Laffer Curve: Evidence from the Canadian Provinces" *Finanz - Archiv : Zeitschrift für das Gesamte Finanzwesen* 74.2 (2018): 173-99. DOI:10.1628/fa-2018-0005.

Fullerton, Don and Gilbert E. Metcalf (2001) "Environmental Controls, Scarcity Rents and Pre-Existing Distortions." *Journal of Public Economics* 80 pp. 249—267.

Goulder, Lawrence H. (1998) "Environmental Policy Making in a Second-Best Setting." *Journal of Applied Economics* Vol I No 2 (Nov. 1998) 279—328.

Goulder, Lawrence H. (2013) "Climate Change Policy's Interactions with the Tax System." *Energy Economics* http://dx.doi.org/10.1016/j.eneco.2013.09.017

McKitrick, Ross R. (1999) "A Derivation of the Marginal Abatement Cost Function." Journal of Environmental Economics and Management, May 1999, pp. 306-314.

Parry, Ian, Roberton C. Williams III and Lawrence H. Goulder (1999). "When Can Carbon Abatement Policies Increase Welfare? The Fundamental Role of Distorted Factor Markets." *Journal of Environmental Economics and Management* 37: 52—84.

Sandmo, Agnar (1975) "Optimal Taxation in the Presence of Externalities." *Swedish Journal of Economics* 77 (1), 86–98.

Schöb, Ronnie (2003) "The Double Dividend Hypothesis of Environmental Taxes: A Survey." Fondazione Eni Enrico Mattei Working Paper No. 60.2003, 11 June 2003 http://papers.ssrn.com/abstract=413866

Turvey, Ralph (1963) "On Divergences Between Social Cost and Private Cost" *Economica* August 1963, reprinted in Dorfman R. and N. Dorfman *Economics of the Environment: Selected Readings* 2nd Ed., New York: Norton, 1977.

US Interagency Working Group on Social Cost of Carbon (IWG) (2013) "Technical Support Document: Technical update of the social cost of carbon for regulatory impact analysis Under Executive Order 12866." United States Government.

7 FIGURE

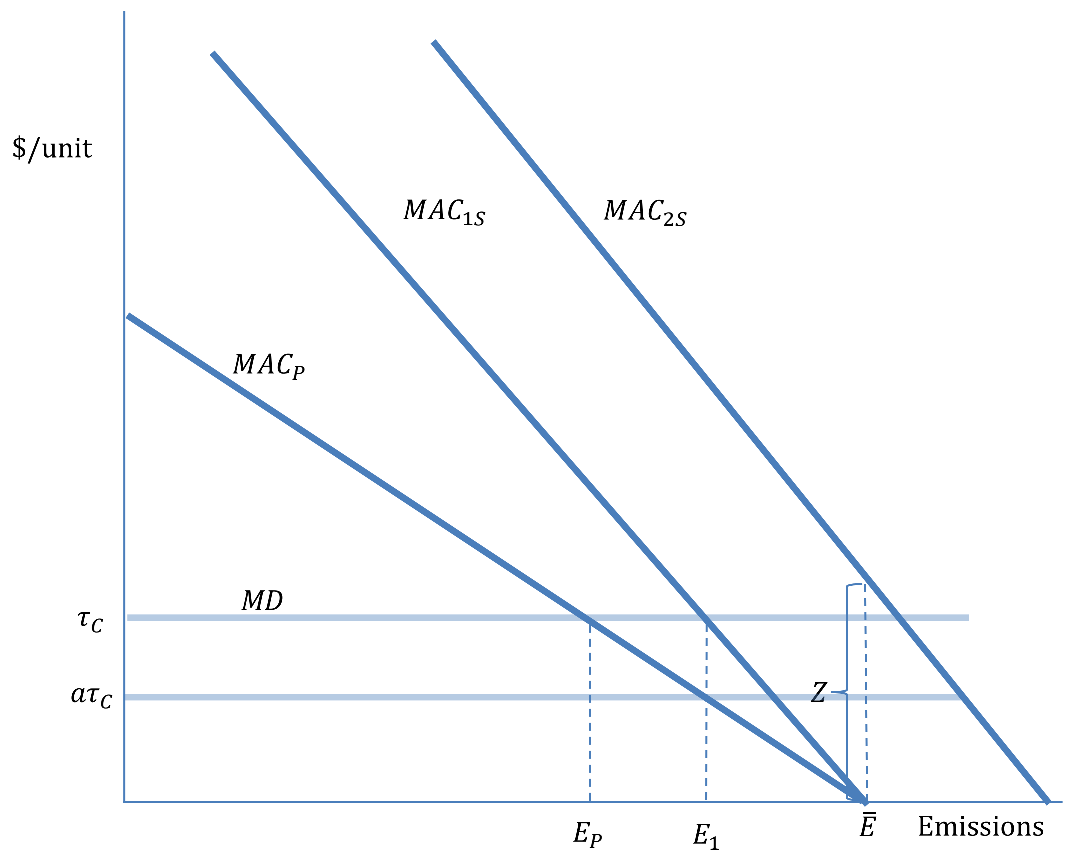

FIGURE 1: The classical Pigovian tax τ_C equals MD (here assumed constant), yielding equivalence between MD and private MAC (MAC_P) at emissions E_P. But when social MAC is rotated out to MAC_{1s} the optimal emissions tax is at $a\tau_C$ and optimal emissions is at E_1. If a positive threshold exists (Z) the marginal welfare costs of emission reductions follow MAC_{2S} and the optimal tax is below $a\tau_C$. As drawn the optimal emission tax is zero even though MD is positive.

8 APPENDIX

Derivation of Equation (8)

The first derivative of the social welfare function (equation 7) with respect to τ_C is

$$\frac{dW}{d\tau_C} = N\left(v_x \frac{dp_x}{d\tau_C} + v_E \frac{dp_E'}{d\tau_C} + \ v_w \frac{dw'}{d\tau_C} + v_y \frac{dy'}{d\tau_C}\right) - \delta N \frac{dC}{d\tau_C}$$

where derivatives of *v* are subscripted in order of the arguments. Divide this equation by v_y and apply Roy's theorem to obtain

$$\frac{dW}{d\tau_C}\frac{1}{v_y} = -X\frac{dp_x}{d\tau_C} - E_{hh}\frac{dp_E'}{d\tau_C} - H\frac{dw'}{d\tau_C} + \frac{dY'}{d\tau_C} - \frac{\delta N}{v_y}\frac{dC}{d\tau_C}.$$

Note $\frac{dp_E'}{d\tau_C} = c$. Use $\frac{dY'}{d\tau_C} = (1-\tau_Y)\frac{d\pi}{d\tau_C} - \pi\frac{d\tau_Y}{d\tau_C} + H\frac{dw'}{d\tau_C} + L\frac{dw'}{d\tau_C}$ and note that because w is the numeraire, $\frac{dw'}{d\tau_C} = -w\frac{d\tau_Y}{d\tau_C}$, to obtain

$$\frac{dW}{d\tau_C}\frac{1}{v_y} = -X\frac{dp_x}{d\tau_C} - cE_{hh} - H\frac{dw'}{d\tau_C} + H\frac{dw'}{d\tau_C} + L\frac{dw'}{d\tau_C} + (1-\tau_Y)\frac{d\pi}{d\tau_C} - \pi\frac{d\tau_Y}{d\tau_C} - \frac{\delta N}{v_y}\frac{dC}{d\tau_C}$$

$$= -X\frac{dp_x}{d\tau_C} - cE_{hh} - wL\frac{d\tau_Y}{d\tau_C} + (1-\tau_Y)\frac{d\pi}{d\tau_C} - \pi\frac{d\tau_Y}{d\tau_C} - \frac{\delta N}{v_y}\frac{dE}{d\tau_C}$$

$$= -X\frac{dp_x}{d\tau_C} - cE_{hh} - \frac{d\tau_Y}{d\tau_C}(\pi + wL) + (1-\tau_Y)\frac{d\pi}{d\tau_C} - \frac{\delta N}{v_y}\frac{dE}{d\tau_E}.$$

When firms optimize inputs the envelope theorem implies $\frac{d\pi}{d\tau_C} = -cE$. Note also that $\pi + wL = B$, so $\frac{dB}{d\tau_C} = -cE + w\frac{dL}{d\tau_C}$. Use these and equation (6) to obtain

$$\frac{dW}{d\tau_C}\frac{1}{v_y} = -X\frac{dp_x}{d\tau_C} - cE_{hh} + C + \tau_C\frac{dC}{d\tau_C} + \tau_y w\frac{dL}{d\tau_C} - \tau_y cE + \frac{d\pi}{d\tau_C} - \tau_y\frac{d\pi}{d\tau_C} - \frac{\delta N}{v_y}\frac{dC}{d\tau_C}$$

$$= -X\frac{dp_x}{d\tau_C} + \tau_Y w\frac{dL}{d\tau_C} + \tau_C\frac{dC}{d\tau_C} - \frac{\delta N}{v_y}\frac{dC}{d\tau_C}.$$

Derivation of Equations (11-13)

Set equation (8) to zero and rearrange to get

$$\tau_C\frac{dC}{d\tau_C} = \frac{\delta N}{v_Y}\frac{dC}{d\tau_C} + X\frac{dp_x}{d\tau_C} + \tau_Y w\frac{dL}{d\tau_C}$$

$$= \frac{\delta N}{v_Y}\frac{dC}{d\tau_C} + X\frac{dp_x}{d\tau_C} - \left(\frac{G}{B} - \tau_C\frac{C}{B}\right) w\frac{dL}{d\tau_C}$$

$$\Longrightarrow \tau_C\left(\frac{dC}{d\tau_C} - \frac{C}{B}w\frac{dL}{d\tau_C}\right) = \frac{\delta N}{v_Y}\frac{dC}{d\tau_C} + X\frac{dp_x}{d\tau_C} - \frac{G}{B}w\frac{dL}{d\tau_C}$$

Then $\tau_C = \frac{\delta N}{v_Y}\frac{dC}{d\tau_C}\left(\frac{dC}{d\tau_C} - \frac{C}{B}w\frac{dL}{d\tau_C}\right)^{-1} + \left(X\frac{dp_x}{d\tau_C} - \frac{G}{B}w\frac{dL}{d\tau_C}\right)\left(\frac{dC}{d\tau_C} - \frac{C}{B}w\frac{dL}{d\tau_C}\right)^{-1}$.

Derivation of Equation (16)

$$\frac{dW}{dC} = Nv_x\frac{dp_x}{dC} + Nv_E\frac{dp_E}{dC} + Nv_w\frac{dw'}{dC} + Nv_y\frac{dy'}{dC} - \delta N$$

$$\frac{dW}{dC}\frac{1}{v_y} = -X\frac{dp_x}{dC} - E_{hh}\frac{dp_E}{dC} - H\frac{dw'}{dC} + N\frac{dy'}{dC} - \frac{\delta N}{v_y}$$

With no emissions tax there is no revenue to recycle, but the emission cap may shrink the tax base requiring an adjustment to τ_y to maintain budget balance. The GBC thus implies $\tau_y\frac{d\pi}{dC} + \tau_y w\frac{dL}{dC} + B\frac{d\tau_y}{dC} = 0$. Also $\frac{dY}{dC} = \frac{d\pi}{dC}(1-\tau_y) - \pi\frac{d\tau_y}{dC} + T\frac{dw'}{dC}$. Then

$$\frac{dW}{dC}\frac{1}{v_y} = -X\,\frac{dp_x}{dC} - E_{hh}\frac{dp_E}{dC} + \frac{d\pi}{dC} - \tau_y\frac{d\pi}{dC} - \pi\frac{d\tau_y}{dC} + L\frac{dw'}{dC} - \frac{\delta N}{v_y}$$

$$= -X\,\frac{dp_x}{dC} - E_{hh}\frac{dp_E}{dC} + \frac{d\tau_y}{dC}(B - \pi - wL) + \tau_y w\frac{dL}{dC} + \frac{d\pi}{dC} - \frac{\delta N}{v_y}$$

$$= -X\,\frac{dp_x}{dC} - E_{hh}\frac{dp_E}{dC} + \tau_y w\frac{dL}{dC} + MAC_p - MD$$

where we use $MAC_p = d\pi/dC$. The rest follows by setting the above equation to zero.