Redistribution and pollution taxes with non-linear Engel curves[☆]

Bas Jacobs [a, *], Frederick van der Ploeg [b, **]

[a] Erasmus University Rotterdam, Tinbergen Institute and CESifo, the Netherlands
[b] University of Oxford, Tinbergen Institute, CEPR and CESifo, the Netherlands

ARTICLE INFO

ABSTRACT

This paper analyses optimal corrective taxation and optimal income redistribution. The Pigouvian pollution tax is higher if pollution damages disproportionally hurt the poor due to equity weighting of pollution damages. Moreover, under general utility functions, optimal pollution taxes should be set below the Pigouvian tax if the poor spend a disproportionate fraction of their income on polluting goods. However, if Engel curves are linear, optimal pollution taxes should follow the first-best rule for the Pigouvian corrective tax even if the government wants to redistribute income and the poor spend a disproportional part of their income on polluting goods. The often-used quasi-linear, CES and Stone-Geary utility functions all have linear Engel curves. If Engel curves are linear, and if pollution taxes are not optimised, Pareto-improving green tax reforms exist that move the pollution tax closer to the Pigouvian tax. Simulations demonstrate that optimal corrective taxes should be Pigouvian if the demand for polluting goods is derived from a LES demand system, but deviate from the Pigouvian taxes if demand for polluting goods demand is derived from a PIGLOG demand system.

© 2019 Elsevier Inc. All rights reserved.

1. Introduction

Pollution taxes typically hurt the poor more than the rich, since the poor consume relatively more polluting goods than the rich.[1] Moreover, the poor may be more exposed to pollution damages than the rich.[2] Therefore, our first question is whether and, if so how, the optimal setting of pollution taxes should take into account distributional concerns. Our second question is whether Pareto-improving green tax reforms exist that raise pollution taxes, while no one is being made worse off. To answer these two questions, we extend the standard analyses of pollution and income taxes in Sandmo (1975) and Bovenberg and De Mooij (1994) with heterogeneous individuals that differ in their gross earnings per hour worked and by how much they suffer from pollution damages. Besides the externalities from pollution, there are no other pre-existing tax and non-tax distortions in the economy. Maximisation of social welfare implies that the government cares about pollution damages, income redistribution, and distortions in the labour market. Pollution taxes thus serve potentially three goals: internalising externalities, redistributing income, and alleviating labour market distortions.

Our first main result is that pollution taxes should generally take into account distributional concerns if Engel curves for polluting goods are non-linear. Empirical studies document that linear Engel curves are rejected by the data (e.g. Banks and Lewbel, 1997; Tovar Reaños and Wölfing, 2018).[3,4] If pollution taxes are regressive, they cause distributional losses and optimal pollution taxes should be set below the Pigouvian tax — ceteris paribus.[5] Consequently, the government sacrifices on its environmental priorities to redistribute more income. Moreover, pollution taxes can be driven below Pigouvian rates for efficiency reasons, since they generally raise the marginal tax wedge on labour, and thus exacerbate labour market distortions caused by the labour income tax, as is known from the literature on the double dividend (e.g. Sandmo, 1975, 2000; Bovenberg and De Mooij, 1994; Bovenberg and van der Ploeg, 1994; Goulder, 1995; Bovenberg, 1998, 1999).

However, we show that if Engel curves for polluting goods are linear, then optimal corrective taxes follow the first-best Pigouvian rule even in second-best situations with distortionary taxation and heterogeneous pollution damages. However, the first-best Pigouvian rule for the optimal pollution tax is evaluated at the second-best rather than the first-best allocation.[6] Consequently, if Engel curves for polluting goods are linear, then distributional concerns should *not* play a role in setting pollution taxes. Intuitively, with linear Engel curves in demand for polluting goods, the linear pollution tax has the same distributional benefits and the same labour market distortions as the linear income tax, but in addition also distorts commodity demands. Hence, the pollution tax should be aimed only at internalising pollution externalities. We show that all utility functions belonging to the Gorman polar class yield linear Engel curves.[7] We calibrate our model to empirical estimates for fuel demand and confirm our result that pollution taxes can deviate substantially from Pigouvian levels if Engel curves are non-linear.

We like to stress that, even if optimal second-best pollution taxes could be Pigouvian, this does not imply that distributional motives do not matter at all for setting optimal corrective taxes, only that the incidence of the corrective tax should not guide the optimal corrective tax. Indeed, if society cares more about inequality, and the poor are hurt relatively more by pollution damages, the equity-weighted Pigouvian pollution tax is higher (cf. Mirrlees, 1978; Anthoff et al., 2009; Anthoff and Tol, 2010).[8]

---

[1] The incidence of taxes on fuels and electricity falls more on the poor than on the rich, but taxes on transport fuels vary non-linearly with disposable income in the OECD (Flues and Thomas, 2015). UK energy outlays are the biggest expenditure item after food for the poorest 10 percent; energy is the smallest expense for the richest 10 percent (Advani et al., 2013). US climate policy is regressive (Burtraw et al., 2009; Rausch et al., 2011; Hassett et al., 2009; Fullerton et al., 2012). Pizer and Sexton (2017) survey the literature on incidence of pollution taxes and household surveys for Mexico, the US and the UK. The distributional impact depends on patterns in energy demand, the physical, social and climatic characteristics of the tax jurisdictions, and how the revenue is used. West and Williams III (2004) estimate an AIDS demand system for the US for gasoline, leisure, and other goods. They show that the gasoline tax is regressive and cuts in the labour tax can make climate policy less regressive, whilst lump-sum transfers can make it more progressive. Crainger and Kohlstad (2010) find that carbon permits or taxes in the United States are regressive, both from a current and from a lifetime income perspective. Both theory and empirics suggest that efficiency standards are less regressive than gasoline taxes (Levison, 2016). India's coal tax might be mildly progressive if the poor do not have complete access to the power grid (Parry et al., 2017).

[2] At the local level, poor people and their children suffer more from fine-particles pollution (e.g. Miranda et al., 2011). Similarly, traffic congestion has ambiguous effects on the welfare distribution. The poor and their children typically live and go to school closest to busy roads. However, the rich have the highest opportunity cost of time (De Palma and Lindsey, 2004).

[3] A review of econometric estimates of energy demand studies can be found in Bohi and Zimmerman (1984). Many studies only consider income effects and not price effects. An excellent example of a recent empirical study of energy demand is Papageorgiou et al. (2017), which obtains an estimate of the substitution elasticity between clean and dirty energy inputs around 2—3. Apart from Banks and Lewbel (1997), there are almost no empirical studies that estimate fully specified demand systems of the non-Gorman polar form for energy and other consumption demands that allow energy to be a necessity.

[4] In macroeconomics, it is also recognised that non-linear Engel curves often match the empirics better and thus help to explain structural transformation and growth (e.g. Boppart, 2014; Comin et al., 2015).

[5] Baumgärtner et al. (2017) show that adjustments for the willingness to pay for environmental goods is substantially affected by income inequality. Adler et al. (2016) examine how priority for the poor affects the social cost of carbon.

[6] Aigner (2014) finds that, even if second-best rules for optimal corrective taxes are the same as the first-rules, the second-best *level* of Pigouvian taxation depends on the second-best allocation, which is different from the first-best allocation due to distortions and available tax instruments.

[7] Williams III (2006) estimates an AIDS demand system for the U.S., which is not of the Gorman polar form. He finds that the optimal gasoline tax exceeds marginal social damages, though distributional concerns cause this tax rate to be substantially less than in a representative-agent model. Cremer et al. (2003) study the incidence of pollution taxes in France with four different income groups and find that pollution taxes should be set much below marginal social damages for redistributional reasons. Mayeres and Proost (2001) discuss the distributional consequences of welfare-improving, revenue-neutral reforms of congestion pricing during peak periods of car transport. Abrell et al. (2017) analyse production externalities and corrective pollution taxes in a model with heterogeneous agents to analyse whether deviations from uniform emission pricing can be optimal.

[8] Moreover, the value of environmental quality is lower if inequality is larger, since public resources are more valuable in more unequal societies (Aigner, 2014).

Our second question is whether Pareto-improving green reforms are feasible starting from a situation of no or insufficient pollution taxation. We show that Pareto-improving tax reforms are feasible only if polluting goods demand features linear Engel curves. Intuitively, if Engel curves are linear, the government can always engineer a change in the linear income tax so that every individual gets a perfect compensation for the increase in linear pollution taxes. Such a tax reform keeps non-environmental welfare of all individuals constant, while reducing aggregate pollution and lowering government revenue. Hence, if pollution taxes are sub-optimally low in the initial situation, such a green tax reform unambiguously boosts social welfare. However, with non-linear Engel curves it is not possible to engineer Pareto-improving green tax reforms, since it is no longer possible to engineer a green tax reform that keeps everyone's non-environmental welfare fixed.

Our paper is related to four strands in the literature. A first strand in the literature analyses optimal corrective taxation and optimal income redistribution with non-linear income taxes, see Cremer and Gahvari (2001), Cremer et al. (1998, 2003), Micheletto (2008), Kaplow (2012), and Jacobs and de Mooij (2015). These papers show that optimal corrective taxes are set at Pigouvian levels if preferences are uniform and weakly separable between leisure and consumption. We extend this literature by analysing linear taxes, since real-world tax systems are not continuously non-linear like in the Mirrlees (1971) model, but (piece-wise) linear. We demonstrate that preferences not only need to be weakly separable, but also need to belong to the Gorman polar class of utility functions to have optimal corrective taxes at Pigouvian levels in second-best settings with distortionary income taxes.

A second strand in the literature analyses optimal corrective taxation and redistribution with linear taxes, see Diamond (1973), Sandmo (2000), Pirttilä and Schöb (1999) and Jacobs and de Mooij (2015). Diamond (1973) analyses optimal linear corrective taxes on 'non-atmospheric' externalities that can differ across individuals and where individualised lump-sum taxes can achieve income redistribution without distortions. Sandmo (2000), Pirttilä and Schöb (1999) and Jacobs and de Mooij (2015) analyse optimal linear corrective taxes on 'atmospheric' externalities and optimal linear income taxation in second-best settings with homogeneous externalities. These contributions show that optimal corrective taxes are generally geared towards income redistribution. Pirttilä and Schöb (1999) recognise that weak separability and linear Engel curves are needed to have optimal corrective taxes at Pigouvian levels, but they do not derive the conditions under which this is the case. Jacobs and de Mooij (2015) show that Pigouvian taxes result in the special case where preferences are weakly separable and homothetic. We extend these contributions on optimal linear corrective and income taxation by allowing for atmospheric externalities that may cause heterogeneous environmental damages across individuals. We derive that only Gorman polar utility functions yield linear Engel curves for polluting goods. Furthermore, we analyse green tax reforms with Gorman polar utility. Finally, we provide simulations with non-linear Engel curves for polluting goods to illustrate the importance of non-linear Engel curves for optimal corrective taxes.

A third and recent strand in the literature analyses environmental tax reforms, the recycling of revenue and the effects on income distribution (e.g. Klenert and Mattauch, 2016; Klenert et al., 2018a; Aubert and Chiroleu-Assouline, 2017).[9] All these contributions employ Stone-Geary utility functions belonging to the Gorman polar class and argue that distributional consequences of pollution taxes are critically important to evaluate the consequences of green tax reforms. However, they do not recognise that optimal pollution taxes should *not* be targeted at income redistribution with Stone-Geary utility functions. Therefore, if pollution taxes are not optimised, one can *always* implement a green tax reform that is distribution neutral by a suitable adjustment of the income tax (i.e. a combination of a lower tax rate and a larger lump-sum transfer). By analysing different tax reforms, these papers may thus give the incorrect impression that distributional consequences of pollution taxes are key in the evaluation of green tax reforms if preferences are of the Stone-Geary form. Moreover, these authors do not recognise that tax reforms exist that unambiguously generate a better environment and while keeping equality constant under Stone-Geary utility. Our findings do not imply that the distributional impact of pollution taxes cannot be important. However, one must then move away from assuming Stone-Geary preferences.

A fourth and older strand in the literature studies the potential double dividend of pollution taxes, see for example Bovenberg and De Mooij (1994), Bovenberg and van der Ploeg (1994), Goulder (1995), Bovenberg (1998, 1999) and Sandmo (2000). By allowing for distributional concerns to motivate tax distortions, we show that both the weak and strong double dividends of environmental taxes no longer exist in the tax optimum. The weak double dividend does not apply, since recycling the proceeds of the pollution tax to cut distortionary labour taxes is no longer preferred over cutting lump-sum taxes, since in the tax optimum the marginal cost of public funds is equal for both labour and lump-sum taxes. Moreover, in the tax optimum the strong double dividend does not exist either. Intuitively, income taxes reduce income inequality, which has social marginal benefits that are in the optimum equal to the deadweight losses of the labour tax. Therefore, the marginal cost of public funds is equal to one for distortionary income taxes, and the government is not interested in cutting the labour tax with the proceeds of the pollution tax.

The outline of this paper is as follows. Section 2 sets up our model of income distribution and pollution damages. Section 3 derives the optimal taxes on income and pollution. Section 4 derives our result that the optimal pollution tax should not be directed at income distributional concerns if preferences are of the Gorman polar form. Section 5 shows that Pareto-improving green tax reforms exist if pollution taxes are not at their Pigouvian levels and preferences are of the Gorman polar form. Section

---

[9] Fullerton and Monty (2013) analyse pollution taxes and distribution between unskilled and skilled workers in a general-equilibrium model with endogenous wages. They show that a pollution tax makes both groups worse off. In their calibration, the income loss of the unskilled workers is so large that recycling all revenue from the pollution tax only to them is not sufficient to compensate them.

6 offers some stylised simulations to illustrate our analytical results. Section 7 discusses the instrument set of the government, heterogeneous preferences, externalities in production, and general equilibrium effects. Section 8 concludes.

2. A model of income distribution and pollution damages

A heterogeneous population of individuals differ in their (exogenous) earning ability $n \in [0, \infty)$. The population size is normalised to one. Earning ability of each individual refers to labour productivity per hour worked, which is distributed across the population according to probability density function $f(n)$. Individuals derive sub-utility $u(c_n, b_n, 1 - l_n)$ from consumption of clean goods c_n, consumption of polluting ('bad') goods b_n, leisure $1 - l_n$, where l_n is labour supply and the time endowment of each individual is normalised to one.[10,11] The sub-utility function $u(\cdot)$ is identical for all individuals, increasing, strictly concave, and twice continuously differentiable. It features positive but diminishing returns in all its arguments. In our theoretical analysis, we assume that all commodities are normal goods.[12]

Individuals derive disutility $D_n \equiv D(b, n)$ from pollution b, which is a function of the aggregate demand for polluting goods: $b \equiv \int_0^{\infty} b_n f(n) \mathrm{d}n$. Environmental disutility $D(\cdot)$ is increasing, convex and twice continuously differentiable in aggregate pollution b. Marginal pollution damages $D_b(b, n)$ are allowed to differ between individuals of different productivity n. Typically, $D_{bn}(b, n) < 0$ so that the poor are hurt relatively more by pollution than the rich as $D_b(b,n)$ decreases with n. Our formulation of type-dependent externalities may therefore also reflect occupation-specific environmental damages resulting from the structure of the economy, for example, if low-skilled jobs are more affected by environmental damages than high-skilled jobs. Our formulation thus considers atmospheric externalities, where the environmental damages differ across individuals, which contrasts with Diamond (1973), who analyses non-atmospheric externalities that differ across individuals.

Total utility U_n of each individual is the sub-utility from commodities minus disutility from pollution damages:

$$U_n = u(c_n, b_n, 1 - l_n) - D(b, n). \tag{1}$$

By assuming that environmental disutility enters utility additively, we focus on the distributional aspects of the pollution tax and environmental quality and abstract from second-best interactions of environmental quality with distortions in labour and goods markets.[13]

The government sets a linear income tax with rate t and non-individualised lump-sum transfer s. The ad valorem pollution tax is also linear and has rate q. The informational requirements to employ linear tax instruments are that the government must be able to verify aggregate labour income and aggregate consumption of polluting commodities.[14] The household budget constraint is:

$$c_n + (1 + q)b_n = (1 - t)nl_n + s. \tag{2}$$

Wages per efficiency unit of labour, the price of clean goods, and the before-tax price of polluting goods are all normalised to one. The first-order conditions for labour supply and commodity demands are given by $u_{1-l}/u_c = (1 - t)n$ and $u_b/u_c = 1 + q$. Indirect utility v_n is given by $v_n \equiv v(t, q, s, b, n)$. Its derivatives follow from Roy's identity: $\partial v_n/\partial s = \eta_n$, $\partial v_n/\partial t = -\eta_n n l_n$, $\partial v_n/\partial q = -\eta_n b_n$, and $\partial v_n/\partial b = D_b(b, n)$, where η_n denotes the marginal private utility of income of individual n. We define the *compensated* elasticities of labour supply and demand for dirty goods as follows: $\varepsilon_{lt} \equiv (\partial l_n^c/\partial t)(1 - t)/l_n < 0$, $\varepsilon_{bt} \equiv (\partial b_n^c/\partial t)$ $(1 - t)/b_n$, $\varepsilon_{lq} \equiv (\partial l_n^c/\partial q)(1 + q)/l_n$ and $\varepsilon_{bq} \equiv (\partial b_n^c/\partial q)(1 + q)/b_n < 0$, where a superscript c denotes a compensated (Hicksian) change. If leisure and polluting goods are Hicksian substitutes (complements), then ε_{bt} is positive (negative) and ε_{lq} is negative (positive).

The government budget constraint states that revenue from income and pollution taxes minus the cost of lump-sum transfers must cover an exogenous public revenue requirement R:

$$R = \int_0^{\infty} (tnl_n + qb_n - s)f(n)\mathrm{d}n. \tag{3}$$

The government maximises social welfare W, which is a weighted sum of the individual utilities:

---

[10] In the remainder, Roman letters denote functions and subscripts denote derivatives (except ability n).

[11] One may interpret labour supply in broad terms, including all responses on the intensive margin, such as intensity of work, entrepreneurial effort, occupational choice, human capital investment, labour supplied to the formal sector rather than to the informal sector (e.g. household production), and tax avoidance and evasion (e.g. via the black market). Hence, the labour supply elasticity in the model corresponds to the elasticity of taxable income (ETI).

[12] In the simulations, however, we allow polluting goods demand to be inferior under the PIGLOG demand system.

[13] See Jacobs and de Mooij (2015) for a complete characterisation of the optimal tax system using general utility functions and a traditional social welfare function.

[14] We postpone the discussion of the consequences of allowing for non-linear instruments that depend on individual labour earnings or polluting goods demands to Section 7.

$$W \equiv \int_0^\infty \omega_n v_n \mathrm{f}(n)\mathrm{d}n, \quad \int_0^\infty \omega_n \mathrm{f}(n)\mathrm{d}n = 1, \quad \omega_n > 0, \quad \forall n, \tag{4}$$

where ω_n is the Pareto weight of individual n. Without loss of generality, we normalise the sum of the Pareto weights to one, i.e. $\int_0^\infty \omega_n \mathrm{f}(n)\mathrm{d}n = 1$. By denoting λ as the multiplier on the government budget constraint in eq. (3), we define the marginal social welfare weight of individual n as $g_n \equiv \omega_n \eta_n / \lambda$. Here, g_n captures the increase in social welfare − measured in money units − if individual n receives an additional euro in income. This equals the welfare weight ω_n applied to the increase in private utility η_n, which is divided by the marginal social value of public revenue to convert from utility to money units. Our formulation in terms of Pareto weights allow us to remain agnostic about the specific distributional objectives of the government, since we do not *a priori* impose any structure on the pattern of Pareto weights.

Naturally, we can derive the social welfare weights from a social welfare function, as is usual in most optimal tax analyses. For example, the concave social welfare function $\Psi(v_n), \Psi' > 0, \Psi'' \leq 0$, gives $\omega_n = \Psi'(v_n)$, and social welfare weights $g_n \equiv \Psi'(v_n)\eta_n / \lambda$ that smoothly decline with income. The weights g_n can also be interpreted as generalised social welfare weights that are a function of individual characteristics, taxes or non-welfarist objectives such as concerns for fairness, hard work, or sacrifice (e.g. Saez and Stantcheva, 2016).

Finally, our characterisation of the optimal tax system describes a Pareto-optimal tax structure. This follows from minimising public resources R in eq. (3) subject to a set of constraints on the distribution of utilities $v_n \geq \bar{v}_n, \forall n$, where $\bar{v}_n$ is an exogenously given level of utility for individual n (e.g. the utility in absence of government intervention). If we define ω_n / λ as the multiplier on each utility constraint, then the social welfare weights are equal to $g_n \equiv \omega_n \eta_n / \lambda$ if the utility constraint is binding and $g_n = 0$ if the utility constraint does not bind.

3. Optimal income and pollution taxes

The government aims to maximise social welfare in eq. (4) subject to the government budget constraint in eq. (3) with λ as its Lagrange multiplier. In order to characterise the optimal tax formulae below, we need to introduce three more economic concepts. First, we need to define the Pigouvian corrective tax in our model with heterogeneous agents and heterogeneous pollution damages. Second, to connect our findings to the second-best literature on optimal corrective taxation, we require a definition for the marginal cost of public funds. Third, we will aggregate the distributional impacts income and pollution taxes on social welfare with the so-called Feldstein (1972) distributional characteristics.

3.1. Pigouvian tax

The primary goal of corrective taxes is to internalise pollution externalities. In our model with heterogeneous agents and heterogeneous pollution damages, the equity-weighted Pigouvian tax, or Pigouvian tax for short, can be defined as:

$$q^P \equiv \int_0^\infty \omega_n \frac{\mathrm{D}_b(b,n)}{\lambda} f(n)\mathrm{d}n = \int_0^\infty g_n \frac{\mathrm{D}_b(b,n)}{\eta_n} f(n)\mathrm{d}n, \tag{5}$$

The Pigouvian tax q^P exactly internalises all damages associated with polluting consumption by all individuals in society. The Pigouvian tax depends on distributional concerns for two reasons. First, the shadow value of public funds λ is higher in countries with a more unequal income distribution. Second, the impacts of environmental damages are distributed unequally across the population. We discuss both in turn.

First, to illustrate the impact of the shadow value of public funds λ, consider the case where marginal environmental damages are constant across individuals (i.e. $\mathrm{D}_{bn}(b,n) = 0$), and write these as $\mathrm{D}_b(b,n) = \mathrm{D}'(b)$. Then, the Pigouvian tax (5) becomes $q^P = \mathrm{D}'(b)/\lambda$. Moreover, if there are no distortionary taxes ($t = 0$), and optimal corrective taxes are Pigouvian ($q = q^P$), the shadow value of public funds boils down to $\lambda = \int_0^\infty \omega_n \eta_n \mathrm{f}(n)\mathrm{d}n$, which follows from eq. (10) below with $g_n^* = g_n = \omega_n \eta_n / \lambda$. Consequently, the shadow value of public funds is larger in more unequal societies (i.e. more dispersion in η_n) or if governments are more inequality averse (i.e. Pareto weights ω_n decline faster). Hence, Pigouvian taxes are then lower − ceteris paribus (cf. Aigner, 2014).

Second, environmental damages are not necessarily constant across individuals. The Pigouvian tax in eq. (7) weights the environmental damages of individual n with the social welfare weight g_n of that individual. Consequently, our formulation captures the literature that applies 'equity weights' to discount environmental damages (cf. Mirrlees, 1978; Anthoff et al.,

2009; Anthoff and Tol, 2010).[15] If poor individuals have higher social welfare weights g_n, and suffer larger marginal damages from aggregate pollution (i.e. $D_{bn} < 0$), then the Pigouvian tax is higher. Similarly, the Pigouvian tax is lower if the rich have lower social welfare weights g_n, and they bear the consequences of environmental damage relatively more (i.e. $D_{bn} > 0$).

To illustrate the equity weighting of environmental damages further, consider a linear pollution damage function of the form $D(b,n) \equiv \omega_n^b b$, where ω_n^b measures the marginal utility loss to individual n if aggregate pollution b increases by one unit. ω_n^b is thus the 'pollution weight' of individual n. Defining the aggregate pollution weight as $\omega^b \equiv \int_0^\infty \omega_n^b f(n) dn$, the equity-weighted Pigouvian corrective tax in eq. (5) becomes[16]:

$$q^P = \left(\omega^b + \text{cov}\left[\omega_n, \omega_n^b \right] \right) \big/ \lambda. \tag{5a}$$

Hence, if the poor suffer disproportionately from pollution (i.e. ω_n^b is higher for lower n), and the government cares more about the poor than the rich (i.e. ω_n is higher for lower n), the covariance between the pollution weights ω_n^b and the welfare weights ω_n is positive, and eq. (5a) indicates that the Pigouvian tax is bigger than the unweighted sum of the marginal pollution damages, ω^b/λ, to reflect the higher (lower) welfare weight given to the poor. Conversely, if the rich suffer more from pollution damages, the Pigouvian tax in eq. (5a) is less than the sum of unweighted pollution damages ω^b/λ.

3.2. Diamond's social marginal welfare weights and the marginal cost of public funds

The marginal cost of public funds (MCPF) measures the scarcity of public resources in terms of private resources. The MCPF is defined as the social marginal value of public resources (in money units) divided by the social marginal value of private resources (in money units) (Jacobs, 2018):

$$MCPF \equiv 1 \bigg/ \int_0^\infty g_n^* f(n) dn, \quad g_n^* \equiv g_n + tn \frac{\partial l_n}{\partial s} + (q - q^P) \frac{\partial b_n}{\partial s}, \tag{6}$$

where g_n^* are Diamond (1975) social marginal welfare weights. These weights measure the social marginal value of private income for each individual n in money units. The social marginal value of private income includes the changes in social welfare due to the change in government revenue if individual n receives additional income. The term $tn\partial l_n/\partial s$ is due to the income effect on labour supply if an individual receives an additional euro. The latter depresses labour tax revenue, and thereby reduces social welfare (provided the labour income tax rate is strictly positive). Similarly, the term $(q - q^P)\partial b_n/\partial s$ is due to the income effect in demand for polluting goods. Raising the income of an individual lowers (raises) social welfare if the pollution tax is below (above) the Pigouvian tax (i.e. $q < q^P$).

Jacobs (2018) argues that the MCPF should be defined by taking the Diamond-based social welfare weights to measure the social marginal value of private income and not the standard social welfare weights $g_n \equiv \omega_n \eta_n/\lambda$, which ignore changes in social welfare due to changes in government revenue. Taking the average of g_n rather than g_n^* to calculate the MCPF results in three important problems, which render the usefulness of the standard MCPF-measure problematic. First, the MCPF for non-distortionary lump-sum taxation is not equal to one in the tax optimum. Second, the MCPF for distortionary taxes is not related to its marginal excess burden in the absence of distributional concerns. Third, the MCPF is highly sensitive to the normalisation of the tax code: it swaps from a number below one if income is taxed to above one if consumption is taxed for exactly the same optimal second-best allocation. All three problems disappear using the Diamond-based social welfare weights defined in (6) to calculate the MCPF (Jacobs, 2018).

3.3. Feldstein distributional characteristics

To aggregate the distributional impacts of taxes on income and polluting goods on social welfare, we employ Feldstein (1972)'s distributional characteristics of labour income nl_n and polluting consumption b_n:

---

[15] The climate-economics literature has adopted equity weighs in calculating the social cost of carbon (i.e. the Pigouvian tax on carbon emissions) so as to capture trade-offs between present and future generations (e.g. Bruce et al., 1995; Anthoff and Tol, 2010). Equity weights that decline with income or consumption typically push up the social cost of carbon. However, if equity weights are applied both across space and time, equity weighting can decrease the cost of carbon (e.g. Anthoff et al., 2009; Nordhaus, 2014).

[16] Use the definition for the covariance between two variables x and y ($\text{cov}[x,y] = E[xy] - E[x]E[y]$, where E is the expectations operator) and substitute $\omega^b \equiv \int_0^\infty \omega_n^b f(n) dn$, to express the Pigouvian tax as: $q^P \equiv \int_0^\infty \omega_n \omega_n^b f(n) dn/\lambda = \left(\int_0^\infty \omega_n f(n) dn \int_0^\infty \omega_n^b f(n) dn + \text{cov}[\omega_n, \omega_b] \right)/\lambda = (\omega^b + \text{cov}[\omega_n, \omega_n^b])/\lambda$.

$$\xi_l \equiv \frac{\int_0^\infty (1 - g_n^*) n l_n f(n) \mathrm{d}n}{\int_0^\infty g_n^* f(n) \mathrm{d}n \int_0^\infty n l_n f(n) \mathrm{d}n} = \frac{-\mathrm{cov}\left[g_n^*, n l_n\right]}{\mathrm{E}\left[g_n^*\right] \mathrm{E}[n l_n]}, \tag{7a}$$

$$\xi_b \equiv \frac{\int_0^\infty (1 - g_n^*) b_n f(n) \mathrm{d}n}{\int_0^\infty g_n^* f(n) \mathrm{d}n \int_0^\infty b_n f(n) \mathrm{d}n} = \frac{-\mathrm{cov}\left[g_n^*, b_n\right]}{\mathrm{E}\left[g_n^*\right] \mathrm{E}[b_n]}, \tag{7b}$$

where $\mathrm{cov}[g_n^*, x_n]$ is the covariance between Diamond's social welfare weight g_n^* and x_n for $x_n = l_n, b_n$, and $\mathrm{E}[x_n] \equiv \int_0^\infty x_n f(n) \mathrm{d}n$ denotes the expectations operator for variable x_n. The characteristics ξ_l and ξ_b, capture how strongly the Diamond social welfare weights correlate with the tax bases. They measure the marginal increase in social welfare — expressed in money units as a fraction of the tax base — of the additional income redistribution that is achieved by marginally raising the tax rate on that base. Alternatively, the Feldstein characteristics can be interpreted as the weighted net benefit, i.e. the weighted average of $(1 - g_n^*)$, of raising one euro with particular tax instrument, where the respective weights are labour income or the demand for polluting goods. The net benefit consists of the gross benefit of raising one euro (1) in revenue minus the utility loss (g_n^*) inflicted on all tax payers that need to pay the additional euro. The economic interpretation of the Feldstein characteristic is thus akin to the deadweight loss of the tax rate. The deadweight loss measures the marginal reduction in social welfare — in money units as a fraction of the taxed base — if the tax rate is increased on that base over and above raising the same amount of revenue with lump-sum taxes.

Feldstein's distributional characteristic lies between minus and plus one, since it is the normalised covariance between the social welfare weights g_n^* and the contribution of each individual to the tax base. If we assume that Diamond's marginal social welfare weights g_n^* decline with income, the normalised covariance is negative for the income tax base, and thus $\xi_l > 0$. Intuitively, taxing income provides social benefits, since the highest income earners have the lowest welfare weights. If social welfare weights decline with polluting consumption demand, then $\xi_b > 0$. Taxing polluting goods then yields distributional benefits as well. An example where this could be the case is airline travel. However, Diamond's social welfare weights g_n^* (see eq. (6)) and polluting consumption demand b_n may be *positively* correlated if (a) the government cares more about the poor than the rich (i.e. ω_n declines with n) and (b) the poor consume more polluting goods than the rich do, that is, if polluting goods are inferior. In that case, $\mathrm{cov}[g_n^*, b_n] > 0$ in eq. (7b). Feldstein's distributional characteristic for polluting goods is then negative $(\xi_b < 0)$ and taxing polluting goods yields distributional losses. Note that we have ruled this case out by assuming that all commodities are normal in the theoretical analysis. Hence, both distributional characteristics are always positive $(\xi_l > 0, \xi_b > 0)$ provided that the social welfare weights g_n^* are non-increasing.[17]

The Feldstein characteristics in eqs. (7a) and (7b) are zero only if Diamond's marginal social welfare weights g_n^* are constant across all individuals. This is the case if the social welfare function is utilitarian ($\omega_n = 1$ for all n) *and* the private marginal utility of income η_n is constant, hence $g_n = \omega_n \eta_n / \lambda = \eta_n / \lambda$ is independent of n. In that case, we have no distorting taxes, i.e. $t = 0$ and $q = q^P$, so that $g_n^* = g_n = 1$, see Proposition 1 below. With smoothly declining Pareto weights ω_n, the Feldstein characteristics are positive and smaller than 1. For a Rawlsian social welfare function ($\omega_0 > 0$, and $\omega_n = 0$ for all other n), ξ_l and ξ_b converge to 1 if the earnings ability of the lowest individual converges to zero.

3.4. Optimal taxation

After these preliminaries, we are in the position to derive our first proposition on optimal income redistribution and pollution taxes.

Proposition 1. *The optimal pollution tax, the optimal marginal income tax rate, and the marginal cost of public funds are determined by:*

$$\xi_b = \frac{t}{1-t} \left(\frac{-\overline{\varepsilon_{lq}}}{\overline{\beta}} \right) + \frac{q - q^P}{1 + q} \left(\frac{-\overline{\beta}\varepsilon_{bq}}{\overline{\beta}} \right), \tag{8}$$

$$\xi_l = \frac{t}{1-t} \left(-\overline{\varepsilon_{lt}} \right) + \frac{q - q^P}{1 + q} \left(-\overline{\beta}\varepsilon_{bt} \right), \tag{9}$$

---

[17] However, in our simulations with PIGLOG preferences, polluting goods demand can be inferior at higher income levels.

$$MCPF = 1. \tag{10}$$

where a bar denotes an income-weighted variable x, i.e. $\bar{x} \equiv \int_0^\infty n l_n x_n \mathrm{f}(n) \mathrm{d}n / \int_0^\infty n l_n \mathrm{f}(n) \mathrm{d}n$, *and* $\beta_n \equiv (1+q)b_n/((1-t)nl_n))$ *is the marginal propensity to consume dirty goods out of labour income.*

Proof. *See* Appendix A.

The condition for the optimal corrective tax is given in eq. (8). In the absence of labour market distortions ($t = 0$) and redistributional objectives ($\xi_b = 0$), the optimal pollution tax equals the Pigouvian tax: $q = q^P$. However, the optimal pollution tax in eq. (8) generally differs from the Pigouvian tax stated in eq. (5). The optimal pollution tax in eq. (8) equates the distributional benefits of setting a higher corrective tax (the left-hand side) to the deadweight losses in the labour market (first term on the right-hand side) and the deadweight losses in the goods market (second term on the right-hand side). In second-best settings with distortionary taxation and income redistribution, the corrective pollution tax thus serves two other roles besides internalising pollution externalities. First, if the corrective tax entails distributional gains, $\xi_b > 0$, pollution taxes contribute to redistributing income. Consequently, optimal pollution taxes should be set at rates that are higher than dictated by the Pigouvian level − ceteris paribus. If, however, the incidence of corrective taxes is regressive, ξ_b is lower and the distributional gains are lower. Hence, corrective taxes should be lowered accordingly − ceteris paribus. See also Sandmo (2000), Pirttilä and Schöb (1999), and Jacobs and de Mooij (2015).

Second, a corrective pollution tax generally exacerbates labour market distortions, since indirect taxes also increase the total tax wedge on labour supply. This curbs labour supply further below socially optimal levels (provided that $\varepsilon_{lq} < 0$).[18] Hence, if pre-existing labour market distortions are more severe (higher t or higher $-\varepsilon_{lq}$) optimal corrective taxes should be set more below the Pigouvian level − ceteris paribus. This argument for setting non-Pigouvian corrective taxes goes back to Corlett and Hague (1953), who demonstrate that − in settings without pollution externalities − non-uniform commodity taxes are optimal if some commodities are stronger complements with leisure than others.[19] In our setting, optimal pollution taxes should be higher (lower) than the Pigouvian rate if polluting goods are stronger (weaker) complements to leisure than clean goods (cf. Pirttilä and Schöb, 1999; Jacobs and de Mooij, 2015).

Redistributing income or alleviating labour market distortions comes at a price of distorting commodity demands. If the optimal pollution tax is set above the Pigouvian level (i.e. $q > q^P$), raising the corrective tax generates distortions in the demand for polluting goods, and more so if this demand is more elastic (ε_{bq} larger). But, if the optimal pollution tax is set below the Pigouvian rate ($q < q^P$), raising the corrective tax alleviates goods market distortions by moving the pollution tax closer to the Pigouvian rate.

The optimal income tax in eq. (9) equates the distributional benefits of setting a higher income tax (the left-hand side) to the deadweight losses in the labour market (first term on the right-hand side) and the deadweight losses in the goods market (second term on the right-hand side). Intuitively, the larger are the distributional benefits of income taxes, the larger should the optimal income tax rate be. The more elastically labour supply responds to taxation ($-\varepsilon_{lt}$ larger), the lower should the optimal income tax rate be − ceteris paribus. If the corrective tax is set above the Pigouvian level ($q > q^P$), and income taxes reduce demand for polluting commodities ($\varepsilon_{bt} < 0$), the income tax exacerbates goods market distortions, and more so if the demand for polluting commodities is more elastic ($-\varepsilon_{bt}$ larger). However, if the corrective tax is set below the Pigouvian rate ($q < q^P$), the income tax reduces goods market distortions. A higher income tax then reduces the demand for polluting commodities. This raises social welfare if the optimal pollution tax is set below the Pigouvian rate − ceteris paribus.[20]

Finally, eq. (10) shows that the marginal cost of public funds (*MCPF*) equals 1 in the optimal tax system. The reason is that the government changes the transfer s such that the marginal value of one unit of resources is the same in the public sector (1) and in the private sector ($\int_0^\infty g_n^* \mathrm{f}(n) \mathrm{d}n$). Alternatively, distortions in labour or goods markets are present only because they generate either distributional or environmental benefits, as also eqs. (8) and (9) demonstrate. Consequently, economic distortions are the price of more income redistribution or a better environment, and not the price of raising public revenue. After all, the government always has access to a non-distortionary marginal source of finance to raise revenue (i.e. the non-individualised lump-sum transfer). Raising public revenue is therefore not a task of the income tax or of the corrective tax (Jacobs and de Mooij, 2015).[21]

The finding that the *MCPF* equals one in the tax optimum has important implications for both the weak and the strong double dividend hypotheses of environmental tax reforms (Goulder, 1995; Bovenberg, 1998, 1999; Sandmo, 2000). The *weak*

---

[18] We note that the sign of ε_{lq} is generally ambiguous, and the arguments reverse if $\varepsilon_{lq} > 0$. The reason is that clean and polluting goods can have different degrees of Hicksian complementarity with leisure. For most often-used utility functions, labour supply falls with the pollution tax, see Jacobs and de Mooij (2011). However, it can be the case that labour supply increases with the pollution tax if polluting goods and leisure are strong complements. This is what we have in our simulations with PIGLOG preferences.

[19] These results have been generalised to settings with linear taxes and income redistribution (Diamond and Mirrlees, 1971a,b; Diamond, 1975) and non-linear taxes and income redistribution (Atkinson and Stiglitz, 1976; Jacobs and Boadway, 2014).

[20] We note that in general the sign of ε_{bq} is not unambiguous, see also footnote 18.

[21] If the analysis is extended to allow for public goods, the optimal provision follows from the modified Samuelson rule, and no corrections need to be made for the *MCPF*. See also Jacobs (2018).

double dividend hypothesis posits that using the proceeds of an increase in environmental taxes to cut distortionary taxes yields larger social welfare compared to using the same proceeds to cut non-distortionary lump-sum taxes. Intuitively, this is because the *MCPF* for distortionary taxes is larger than the *MCPF* for non-distortionary taxes. We find that the weak double dividend hypothesis is no longer satisfied in the tax optimum because we allow for redistributional concerns to motivate tax distortions. Indeed, while deadweight losses tend to push the *MCPF* above 1, marginal distributional benefits lower the *MCPF* below 1. In the tax optimum, both are equal so that *MCPF* = 1 for both distortionary and non-distortionary taxes, unlike in models with representative agents. As a result, cutting the income tax is no longer preferred over cutting the lump-sum taxes. Hence, the weak double dividend hypothesis no longer applies. See also Jacobs and de Mooij (2015).[22]

According to the *strong* double dividend hypothesis, recycling the revenue from environmental taxes to cut distortionary labour taxes yields a simultaneous increase in environmental and non-environmental welfare.[23] In the absence of distributional concerns, and employing distortionary taxes as the marginal source of public finance, recycling the revenue from higher pollution taxes in the form of lower labour taxes generally harms labour market performance, because the revenue-recycling effect is typically smaller than the tax-interaction effect. Consequently, there is generally no strong double dividend of environmental tax reforms (Bovenberg and De Mooij, 1994; Bovenberg and van der Ploeg, 1994; Goulder, 1995; Bovenberg, 1998, 1999; Sandmo, 2000).[24] The latter result can also be understood with the aid of the *MCPF*. If the *MCPF* > 1, a tax-shift from taxes on labour to taxes on pollution exacerbates the pre-existing distortions of the tax system by raising the effective marginal tax rate on labour, which includes the pollution tax.

However, with distributional concerns (and allowing for non-individualised lump-sum transfers), non-environmental welfare is not only determined by the distortions in the labour market, but also by the distributional impact of income and environmental taxes. Suppose we start from an optimal income tax system, so that eqs. (9) and (10) hold and the *MCPF* = 1, and that the pollution tax is not optimised, so that eq. (8) does not hold. Raising the pollution tax then raises environmental welfare if the initial pollution tax is lower than the optimal pollution tax. Moreover, recycling the revenue from a marginally larger pollution tax via lower income taxes, not only gives a negative welfare effect of larger labour market distortions, but also results in an exactly offsetting welfare increase of lower income inequality. A strong double dividend of better environmental quality and larger non-environmental welfare does not occur if the labour tax is optimised. Intuitively, non-environmental welfare does not change if the income tax is optimal, and the *MCPF* = 1 for the distortionary labour tax. Therefore, the government is not interested in lowering the distortionary tax on labour income with the proceeds of the pollution tax. It is sufficient to judge the desirability of raising the pollution tax by looking at its impact on environmental welfare only.

The strong double dividend might re-appear if taxes on labour income are not optimised from a redistributional perspective, for example because there is a constraint on the use of lump-sum taxation. In this case, the initial tax code is suboptimal to start with and the environmental tax can either exacerbate or alleviate the sub-optimal redistribution by the income tax. Let us suppose that eq. (10) no longer holds, so that the *MCPF* for the income tax rate can be both smaller or larger than 1 in the constrained optimum. This depends on whether distributional gains of income taxes are larger or smaller than labour market distortions and distortions in the market for polluting goods (Jacobs and de Mooij, 2015; Jacobs, 2018). If *MCPF* < 1, there is insufficient income redistribution and a strong double dividend exists, since environmental and non-environmental welfare both increase with a green tax reform. Intuitively, the larger corrective tax results in a larger effective tax burden on labour, which compensates for the sub-optimally low level of the income tax.[25] However, if *MCPF* > 1, then a strong double dividend cannot occur, like in Bovenberg and De Mooij (1994), Bovenberg and van der Ploeg (1994), Goulder (1995), Bovenberg (1998, 1999) and Sandmo (2000). The income tax is then set sub-optimally high, and a higher pollution tax only exacerbates this distortion, so that non-environmental welfare unambiguously declines.

4. Gorman polar preferences

Many applications assume that utility is weakly separable between commodities and labour, i.e. $U_n = \mathrm{u}(\mathrm{h}(c_n, b_n), 1 - l_n) - \mathrm{D}(b, n)$, with a sub-utility function $\mathrm{h}(\cdot)$ that belongs to the Gorman (1961) polar form.[26] This section demonstrates that in this case the (regressive) incidence of pollution taxes should not determine optimal corrective taxes.

---

[22] Pirttilä and Schöb (1999) do not recognise that the *MCPF* equals one in the tax optimum and thus inadvertently connect their results to the double dividend literature.

[23] Non-environmental welfare corresponds to the weighted sum of utilities excluding the environmental damages. Environmental welfare corresponds to total environmental damages.

[24] This presumes that initial tax systems are not sub-optimal from a revenue-raising perspective and that there are no other non-tax distortions.

[25] This finding generalises to settings where the initial tax system is not optimal from a non-environmental perspective or where there are market failures. Indeed, a strong double dividend occurs if the environmental tax helps to correct initial tax or non-tax distortions, for example, due to involuntary unemployment, non-taxed fixed factors, rents, subsidies on energy use, see also Bovenberg and van der Ploeg (1994), Goulder (1995) and Bovenberg (1998, 1999).

[26] We also assume weak separability, since we wish to focus on the case where all commodity taxes are uniform in the absence of environmental externalities if preferences are Gorman polar (cf. Deaton, 1979). Empirical studies generally find weak evidence that labour supply is weakly separable from commodity demand (e.g. Meghir and Browning, 1991; Crawford et al., 2010; Pirttilä and Suoniemi, 2014). Nevertheless, no clear-cut patterns in commodity demands and labour supply emerge from this literature, only in some obvious cases, such as child-care services. Moreover, many studies in environmental economics and macroeconomics assume weakly separable preferences between commodity demands and labour supply.

Definition 1.   – *The sub-utility function $h_n = \mathrm{h}(c_n, b_n)$ for individual n is of the* Gorman (1961) *polar form if its corresponding indirect utility function can be written as* $h_n = \upsilon(q, m_n) = (m_n - \varphi(q))/\mathrm{p}(q)$, *where* $m_n \equiv (1-t)nl_n + s$ *is net disposable income,* $\partial h_n/\partial m_n = 1/\mathrm{p}(q)$ *denotes the marginal sub-utility of income, and* $\varphi(q)$ *is the expenditure to obtain zero sub-utility. Both* $\mathrm{p}(q)$ *and* $\varphi(q)$ *are the same for all individuals. The expenditure function is* $m_n = \mathrm{p}(q)h_n + \varphi(q)$.

Marshallian demand for polluting goods follows from application of Roy's identity:

$$b_n = -\frac{\partial \upsilon/\partial q}{\partial \upsilon/\partial m_n} = \varphi'(q) + \frac{\mathrm{p}'(q)}{\mathrm{p}(q)}(m_n - \varphi(q)). \tag{11}$$

Consequently, the Engel curve is linear for all preferences that are of the Gorman polar form, which allows for aggregation over all individuals. For later reference, we rewrite the demand for polluting goods as:

$$(1+q)b_n = \alpha(q) + \phi(q)m_n, \tag{12}$$

where the coefficients $\alpha(q) \equiv (1+q)(\varphi'(q) - \varphi(q)\mathrm{p}'(q)/\mathrm{p}(q))$ and $\phi(q) \equiv (1+q)\mathrm{p}'(q)/\mathrm{p}(q)$ are the same for everyone.[27]

We now provide some examples of the Gorman polar form (see Appendix B for derivations):

- *Quasi-linear preferences*: $\mathrm{h}(c_n, b_n) = c_n + \mathrm{k}(b_n)$, $\mathrm{k}' > 0$, $\mathrm{k}'' < 0$. Demand for polluting goods is $b_n = \kappa(q)$ with $\kappa'(q) = 1/\mathrm{k}''(b_n)$ from $\mathrm{k}'(b_n) = 1 + q$. Hence, we have $\alpha(q) \equiv (1+q)\kappa(1+q)$, $\phi(q) = 0$, $\mathrm{p}(q) = 1$, and $\varphi(q) = (1+q)\kappa(1+q) - \mathrm{k}(\kappa(1+q))$.
- *Standard Stone-Geary preferences*: $\mathrm{h}(c_n, b_n - \overline{b}) = (c_n/(1-a))^{1-a}((b_n - \overline{b})/a)^a$, $0 < a < 1$, where $\overline{b}$ is the subsistence level for the consumption of polluting goods. Demand for polluting goods is $b_n = (1-a)\overline{b} + am_n/(1+q)$. Hence, we have $\alpha(q) \equiv (1+q)(1-a)\overline{b}$, $\phi(q) = a$, $\mathrm{p}(q) = (1+q)^a$ and $\varphi(q) = (1+q)\overline{b}$. This specification leads to the Linear Expenditure System. See also (18) below.
- *Generalised Stone-Geary preferences*: $\mathrm{h}(c_n, b_n - \overline{b}) = [(1-a)c_n^{(\sigma-1)/\sigma} + a(b_n - \overline{b})^{(\sigma-1)/\sigma}]^{\sigma/(\sigma-1)}$, $0 < a < 1$ and $\sigma > 0$, where $\overline{b}$ is the subsistence level for the consumption of polluting goods. Demand for polluting goods is: $b_n = (1 + (1+q)^{\sigma-1}((1-a)/a)^\sigma)^{-1}m_n/(1+q) + (1 + (1+q)^{1-\sigma}(a/(1-a))^\sigma)^{-1}\overline{b}$. Hence, we have $\alpha(q) \equiv (1 + a^\sigma(1-a)^{-\sigma}(1+q)^{1-\sigma})^{-1}(1+q)\overline{b}$, $\phi(q) = (1 + (1-a)^\sigma a^{-\sigma}(1+q)^{\sigma-1})^{-1}$, $\mathrm{p}(q) = (1-a)^{\sigma/(1-\sigma)}(1 + a^\sigma(1-a)^{-\sigma}(1+q)^{1-\sigma})^{1/(1-\sigma)}$, and $\varphi(q) = (1+q)\overline{b}$.

These three examples correspond to non-homothetic preferences. The class of homothetic preferences is nested within the class of Gorman polar form preferences by setting $\varphi(q) = 0$. For example, Cobb-Douglas and CES preferences are nested in the standard and generalised Stone-Geary preferences, respectively, if the subsistence level of consumption of polluting goods is set to zero (i.e. $\overline{b} = 0$).

Proposition 2 shows that with Gorman polar form preferences the optimal rule for the optimal pollution tax is the same as the first-best rule, i.e. the pollution tax is set at the Pigouvian level.

Proposition 2.   *If the utility function is* $U_n = \mathrm{u}(\mathrm{h}(c_n, b_n), 1 - l_n) - \mathrm{D}(b, n)$, *where* $\mathrm{h}(c_n, b_n)$ *is of the* Gorman (1961) *polar form and identical for all individuals, the optimal pollution tax, the optimal marginal income tax rate, and the marginal cost of public funds are determined by:*

$$q = q^P \equiv \int_0^\infty g_n \frac{\mathrm{D}_b(b, n)}{\eta_n}\, \mathrm{f}(n)\mathrm{d}n, \tag{13}$$

$$\frac{t}{1-t} = \frac{\xi_l}{-\varepsilon_{lt}}, \tag{14}$$

$$MCPF = 1. \tag{15}$$

Proof.   *See* Appendix C.

If preferences are weakly separable and sub-utility is of the Gorman polar form, the optimal corrective pollution tax in eq. (13) is solely aimed at internalising pollution externalities, not at redistributing income or reducing labour market distortions. Consequently, the optimal corrective tax is set at the Pigouvian level, even if the poor spend a larger fraction of their income on consumption of polluting goods. The reason is that with Gorman polar preferences the Engel curves of polluting goods are linear in labour earnings. At the margin, individuals have the same marginal propensities to consume polluting goods out of their net income. Consequently, linear corrective pollution taxes have equivalent marginal distributional benefits as linear taxes on labour earnings. Similarly, if the demand for polluting consumption goods features a linear Engel curve, linear pollution taxes cause identical distortions in labour supply as linear income taxes with the same revenue. Although corrective pollution taxes have the same distributional benefits and the same labour market distortions as income taxes, they also

---

[27] $\phi(q) = ((1+q)b_n - \alpha(q))/m_n$ is the *marginal* propensity to consume polluting goods out of net labour income. If $\alpha(q) > 0$, then $\phi(q)$ is smaller than the expenditure share $(1+q)b_n/m_n$, which is the *average* propensity to consume polluting goods out of net labour income, and vice versa if $\alpha(q) < 0$.

distort commodity demands. Consequently, there are no redistributional or efficiency reasons to set corrective taxes at non-Pigouvian rates if preferences are Gorman polar. Therefore, the optimal second-best pollution tax is aimed solely at internalising pollution externalities and follows the first-best, Pigouvian rule.[28]

Proposition 2 derives the general properties of the utility function that yield optimal Pigouvian corrective taxes in second-best settings with distortionary taxation. In doing so, it generalises Corollary 1 in Jacobs and de Mooij (2015) for weakly separable homothetic sub-utility functions without income effects to any Gorman polar form sub-utility function with income effects. Moreover, it proves the claim in Pirttilä and Schöb (1999) that linear Engel curves are needed for optimal corrective taxes in second-best settings to be Pigouvian.

Although we have shown that the optimal decision *rule* for second-best pollution taxes is the same as in first-best settings, optimal pollution taxes are generally different in second-best settings as optimal allocations differ in first- and second-best outcomes. The total amount of pollution b, private marginal utilities of income η_n, and social welfare weights g_n are typically not the same in the first-best and second-best optima.

Corollary 1 derives the optimal tax system for a constant private marginal utility of income, no income effects in labour supply, and uniform marginal environmental damages over the population.

Corollary 1. *Assume that utility is quasi-linear:* $u(h(c_n, b_n), 1 - l_n) = h(c_n, b_n) - l_n^{1+1/\vartheta}/(1 + 1/\vartheta)$, *where* $\vartheta > 0$ *and* $h(c_n, b_n)$ *is of the* Gorman (1961) *polar form. Assume that marginal environmental damages are independent of ability:* $D_b(b, n) \equiv D'(b)$. *Then, labour supply is* $l_n = ((1 - t)n/p(q))^{\vartheta}$, *the optimal pollution tax is* $q = q^P = p(q)D'(b)$, *and the optimal labour income tax rate follows from* $t/(1 - t) = \xi_l/\vartheta$, *where the Diamond welfare weights are* $g_n^* = g_n, \forall n$. *If Pareto weights are the same for all individuals (* $g_n = 1, \forall n$ *), then the optimal income tax rate is zero (* $t = 0$ *).*

Proof. *See* Appendix D.

If sub-utility is quasi-linear, the optimal pollution tax is Pigouvian, and equals marginal pollution damages $q = p(q)D'(b)$. The optimal income tax rate depends positively on demand for income redistribution ξ_l and varies inversely with the compensated elasticity of labour supply ϑ.[29]

5. Green tax reforms

Policy makers may find it hard to implement an *optimal* income and pollution tax system. It is therefore relevant to investigate the question whether welfare-improving tax reforms exist that raise pollution taxes and reduce income taxes if the starting position is one where the initial income and pollution taxes are not at their optimal levels from Proposition 1. We do so by considering a net-income neutral tax reform, which is similar in spirit to Kaplow (2012). In particular, we analyse a tax reform that keeps net incomes, and, thereby, non-environmental utility $u(\cdot)$, of all individuals fixed, while environmental damages are allowed to change. This reform is not necessarily budgetary neutral.[30] We ask the question whether such a net-income neutral increase in the pollution tax raises social welfare. That is, are reductions in aggregate pollution damages, which raise social welfare, sufficiently large that they compensate for reductions in public revenue, which lower social welfare? Proposition 3 provides the answer.

Proposition 3. *Assume that the utility function is* $U_n = u(h(c_n, b_n), 1 - l_n) - D(b, n)$, *where* $h(c_n, b_n)$ *is of the Gorman polar form. Consider an environmental tax reform that raises pollution taxes (* $dq > 0$ *) and keeps net incomes and non-environmental welfare of all individuals constant (* $du_n = 0$ *) by lowering the income tax and raising the transfer such that*

$$dt = -\frac{\phi(q)(1 - t)}{1 + q}dq < 0 \ and \ ds = \left(\frac{\alpha(q) + \phi(q)s}{1 + q}\right)dq > 0. \tag{16}$$

Social welfare W rises (falls) under this tax reform if the pollution tax is below (above) the Pigouvian tax:

$$\frac{1}{\lambda}\frac{dW}{dq}\bigg|_{reform} = \int_0^\infty \left[\frac{(q^P - q)}{(1 + q)}\left(\phi(q)\varepsilon_{bt} - \varepsilon_{bq}\right)\right]b_n f(n)dn. \tag{17}$$

Proof.   See Appendix E for the proof of (16) and (17) and Appendix C for the proof that $\phi(q)\varepsilon_{bt} > \varepsilon_{bq}$.

---

[28] This result is the environmental-tax counterpart to the findings by Sandmo (1975) and Deaton (1979) on the desirability of uniform linear commodity taxes. Our findings apply to the pollution tax net of the Pigouvian tax, i.e. $q - q^P$.

[29] See also Boadway and Tremblay (2008), who analyse optimal corrective taxation and optimal redistribution with preferences that are quasi-linear in labour.

[30] This tax reform is therefore different than the typical tax reform that is analysed in the double dividend literature, which maintains budget balance, see e.g. Bovenberg and De Mooij (1994).

If preferences are identical for all individuals and are of the Gorman polar form, both the slope and the intercept of the Engel curve for polluting goods are the same for all individuals. The increase in the pollution tax shifts the intercept of the Engel curve and changes its slope. Hence, in order to keep everyone's non-environmental utility fixed, the government needs a policy instrument to off-set the change in the intercept of the Engel curve − the transfer s − and an instrument to off-set the change in the slope of the Engel curve − the income tax rate t. The government can thus implement a tax reform that simultaneously raises pollution taxes q and the transfer s, while lowering the tax rate t, such that everyone's non-environmental utility u(·) remains constant, see eq. (16). Since the changes in policy instruments are identical for all individuals, the distribution-neutral green tax reform in Proposition 3 can be implemented with policy instruments that are identical for everyone.

The only reason why the government would implement such a tax reform is that it brings the tax system closer to the second-best optimum, since it does not use the pollution tax for reducing labour market distortions or redistributing incomes. Eq. (17) of Proposition 3 demonstrates that if corrective taxes q are lower (higher) than the Pigouvian rate q^P, the distribution-neutral environmental tax reform in eq. (16) raises (lowers) social welfare. Hence, if the government prices pollution insufficiently, it is possible to recycle revenues from a higher pollution tax in such a way that none of the individuals' non-environmental utilities is affected, whilst aggregate damages from pollution fall. This striking result follows (again) from the fact that Engel curves for polluting goods are linear.

Proposition 3 does not hold if preferences are not of the Gorman polar form, i.e. if Engel curves are non-linear.[31] The reason is that it is not possible to design a tax reform that simultaneously changes the pollution tax rate, the income tax rate and the lump-sum transfer such that non-environmental utility remains fixed for all individuals. Intuitively, the income tax rate needs to off-set the impact of the pollution tax on the slope of the Engel curve of polluting goods, while the lump-sum transfer adjusts to off-set the impact of the pollution tax on the intercept of the Engel curve for polluting goods. If either the slope or the intercept of the Engel curve of the polluting good varies with income, then the government requires a different income tax rate and a different transfer for each individual to keep non-environmental utility fixed in the tax reform. Given that tax rates are flat and transfers are uniform, it is therefore not possible to design net-income neutral tax reforms if Engel curves are not linear. Then, the distributional incidence of pollution taxes and the tax-interaction effects with labour supply determine their desirability, as we have demonstrated in Proposition 1.

Proposition 3 does not require that the income tax system is optimised and holds for any arbitrary income tax system (t, s) with any arbitrary MCPF. Hence, Proposition 3 is more general than Proposition 2, which assumes that (income) taxes are optimised. If the income tax is optimised, but the pollution tax is not, then the MCPF is unity (cf. eq. (15) of Proposition 2). Therefore, Proposition 3 casts even more doubt on both the weak and the strong double-dividend hypotheses, which rely on the MCPF for income taxes to be smaller than the MCPF for lump-sum taxes (weak double dividend hypothesis) and the MCPF for income taxes to be larger than one (strong double dividend hypothesis). Proposition 3 demonstrates that, as long as distributional-neutral green tax reforms are feasible, the only criterion whether to increase pollution taxes is whether doing so would bring the pollution tax closer to the optimal pollution tax, irrespective of how large the MCPF is for either lump-sum or income taxes. Intuitively, with distribution-neutral tax reforms, both the deadweight losses and redistributional benefits of the entire tax-transfer system remain constant. As a result, there is no change in labour market distortions (the labour market dividend in the double dividend literature) and no change in redistribution. Indeed, the only criterion to judge the desirability of pollution taxes is whether the increase in environmental welfare (due to better environmental quality) is larger than the decline non-environmental welfare (lower public revenue due to larger pollution taxes), which boils down to setting corrective taxes at Pigouvian levels if utility belongs to the Gorman polar class.

If preferences are of the Gorman polar form, Proposition 3 demonstrates that the government can design net-income neutral adjustments of the income tax even if the poor − on average − spend a larger fraction of their income on polluting goods. Proposition 3 applies irrespective of whether pollution damages are equally or unequally distributed in the population. However, this does not imply that distributional concerns do not affect the desirability of the tax reform. The distributional impact of pollution determines the optimal Pigouvian pollution tax ($q = q^P$), and thereby the desirability of the environmental tax reform.

Proposition 3 once more reveals that one should be careful to conclude that the regressive incidence of pollution taxes determines the desirability of green tax reforms. Klenert and Mattauch (2016), Klenert et al. (2018a), and Aubert and Chiroleu-Assouline (2017) analyse the environmental and distributional consequences of green tax reforms under different assumptions regarding the recycling of tax revenue via the income tax. They show that tax reforms *may* exist that yield a better environment and more income redistribution under some recycling schedules for the income tax. These papers might give the impression that the distributional incidence of pollution taxes is important to judge the desirability of green tax reforms. However, this impression is misleading, since all papers employ Stone-Geary utility functions that belong to the Gorman polar class. Proposition 2 shows that in this case optimal pollution taxes should *not* be aimed at income redistribution. Moreover, Proposition 3 demonstrates that the government can always achieve a better environment while keeping the distribution of non-environmental welfare constant under the specific tax reform being analysed (a decrease in the tax rate jointly with an increase in the demogrant). Klenert and Mattauch (2016), Klenert et al. (2018a), and Aubert and Chiroleu-Assouline (2017) do

---

[31] It does not matter whether Engel-curves are (locally) concave or convex for Proposition 3 to break down. Engel curves are often logarithmic. In our calibrated example with PIGLOG preferences, the Engel curve for polluting goods is concave over the whole range and thus the Engel curve for clean goods is convex over the whole range. More generally, Engel curves can have non-linear forms, for example in the EASI demand system.

not analyse this reform, which explains why they do not unambiguously find that the environment can improve without sacrificing the goal of income redistribution.

6. Simulations of optimal climate and redistribution policies and tax reform

To what extent do non-Gorman polar forms for preferences affect the optimal pollution tax and to what extent does it differ from the Pigouvian tax? To illustrate our main theoretical findings, this section presents simulations of optimal pollution taxes and redistributive income taxes for two different types of demand systems. In Section 6.1 we formally analyse the behaviour of the demand for polluting goods in the Linear Expenditure System (LES) and the PIGLOG (i.e. price-independent generalised logarithmic) demand system. The LES is based on Stone-Geary preferences that are of the Gorman polar form and features linear Engel curves. In contrast, the PIGLOG system is not of the Gorman polar form and thus features empirically more plausible non-linear Engel curves (Banks and Lewbel, 1997). Section 6.2 discusses the calibration of our demand systems to empirical estimates for fuel demand of Banks and Lewbel (1997). Sections 6.3 and 6.4 give the benchmark and various sensitivity runs for the optimal tax system under the LES and PIGLOG demand system, respectively.

6.1. LES and PIGLOG demand systems

In applied econometric studies, demand functions for energy and other commodities are usually estimated with flexible demand systems based on price-independent generalised linear (PIGL) preferences (Muellbauer, 1976), which contain PIGLOG, translog and AIDS (i.e. almost ideal demand system) preferences as special cases (e.g. Deaton and Muellbauer, 1980).[32] Although the Gorman polar form is a limiting case of PIGL preferences, PIGLOG, translog or AIDS preferences do not satisfy the Gorman polar form.[33] These latter demand systems satisfy the axioms of order and permit aggregation over consumers without demanding parallel linear Engel curves.

In our application, we assume that utility is quasi-linear, utility from clean and dirty goods is separable from leisure, and that the Frisch elasticity of labour supply is constant and denoted (as in Corollary 1) by ϑ. The sub-utility function of individual n is thus $u_n = \mathrm{h}(c_n, b_n) - l_n^{1+1/\vartheta}/(1 + 1/\vartheta), \vartheta > 0$, which implies that there are no income effects in labour supply if $\mathrm{h}(\cdot)$ is linearly homogeneous. We employ two specifications of sub-utility $\mathrm{h}(c_n, b_n)$. The first one is the Stone-Geary utility function, $\mathrm{h}(c_n, b_n - \overline{b}) = (c_n/(1-a))^{1-a}((b_n - \overline{b})/a)^a, 0 < a < 1$, where $\overline{b}$ is the subsistence level for the consumption of polluting goods. This yields the Linear Expenditure System (LES) discussed above with budget shares, and (un)compensated labour supplies that are given by:

$$\frac{(1+q)b_n}{m_n} = a + (1-a)(1+q)\overline{b}/m_n, \quad \frac{c_n}{m_n} = (1-a)(1 - (1+q)\overline{b}/m_n), \quad 0 < a < 1, \quad \overline{b} > 0, \tag{18}$$

$$l_n = l_n^c = ((1-t)n/(1+q)^a)^\vartheta. \tag{19}$$

Our second specification employs PIGLOG preferences that are characterised by the following expenditure function:

$$\mathrm{m}(h_n, q) = \Theta(q)\Pi(q)^{\ln(h_n)}, \tag{20}$$

where $m_n = \mathrm{m}(h_n, \mathrm{q}) = c_n + (1+\mathrm{q})b_n$ is the minimal expenditure required to reach sub-utility level h_n for individual n if the pollution tax is q. PIGLOG preferences have no neatly defined direct utility function. Denote by $\Theta(q) \equiv (1+q)^{\chi_0}$ and $\ln(\Pi(q)) \equiv (1+q)^{-\chi_1}$ with $\chi_0 > 0$ and $\chi_1 > 0$, then Shephard's lemma gives the compensated demand function: $b_n = \partial\mathrm{m}(h_n, q)/\partial q = (\chi_0 - \chi_1(1+q)^{-\chi_1} \ln(h_n))m_n/(1+q)$. Hence, with PIGLOG preferences the budget shares of dirty and clean goods are given by:

$$\frac{(1+q)b_n}{m_n} = \chi_0 - \chi_1 \ln\left(\frac{m_n}{(1+q)^{\chi_0}}\right), \quad \frac{c_n}{m_n} = 1 - \chi_0 + \chi_1 \ln\left(\frac{m_n}{(1+q)^{\chi_0}}\right). \tag{18a}$$

By inverting $\ln(h_n) = (1+q)^{\chi_1}(\ln(m_n) - \chi_0 \ln(1+q))$, we obtain the indirect utility function for the PIGLOG system: $\upsilon(q, m_n) = (m_n/(1+q)^{\chi_0})^{(1+q)^{\chi_1}}$. The corresponding expenditure function is thus given by: $m(h_n, q) = (1+q)^{\chi_0} h_n^{-(1+q)^{\chi_1}}$. Maximising utility $\upsilon(q, m_n) = \left(\frac{(1-t)nl_n+s}{(1+q)^{\chi_0}}\right)^{(1+q)^{\chi_1}} - \frac{l_n^{1+1/\vartheta}}{1+1/\vartheta}$ gives the uncompensated labour supply function with PIGLOG preferences:

---

[32] Bohi and Zimmerman (1984) give a survey of econometric estimates of energy demand functions.

[33] Empirical studies often use flexible functional forms based on, for example, translog functional forms (e.g. Berndt and Wood, 1975; Pindyck, 1979) or Diewert's generalised Cobb-Douglas functional forms (e.g. Magnus, 1979) to estimate energy demand. These are not of the Gorman polar form either.
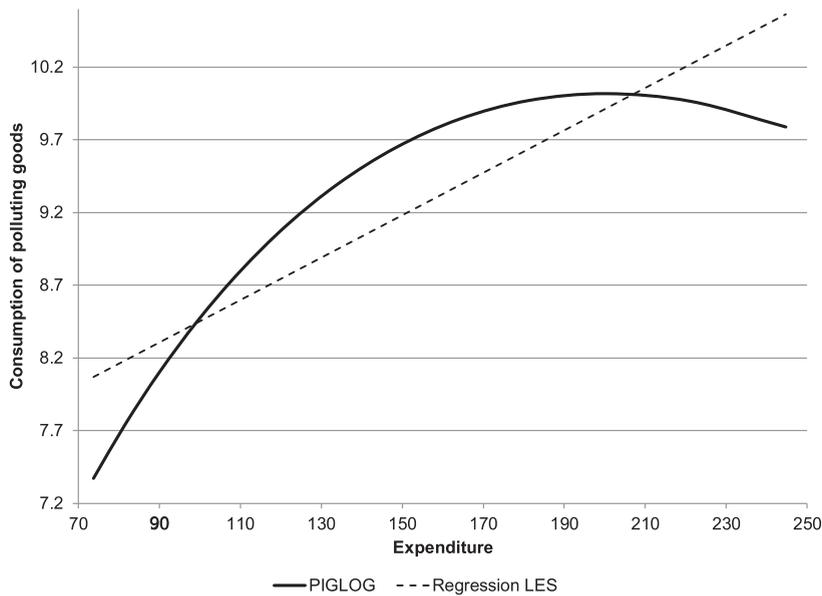
Fig. 1. Calibration of LES system to data generated by PIGLOG system.

$$l_n = \left(\frac{(1+q)^{\chi_1}(1-t)nh_n}{(1-t)nl_n + s} \right)^{\vartheta}.$$ (19a)

Labour supply is thus an increasing function of the pollution tax under PIGLOG demand (i.e. $\varepsilon_{lq} > 0$), but a negative function of the pollution tax under LES demand. This implies that polluting goods and leisure are stronger complements than clean goods and leisure.

In both demand systems, polluting goods are necessities and clean goods are luxuries. The Engel curves of the LES (18) are linear. In contrast, in the PIGLOG system (18a), the Engel curves are non-linear. PIGLOG preferences are thus not of the Gorman polar form and Proposition 2 does not apply.[34] The budget share of polluting goods (18) and (18a) rises in the pollution tax for both demand systems. Dirty goods are normal goods under LES demand at any level of income but become inferior goods at high levels of income under PIGLOG demand.

6.2. Calibration

We calibrate both demand systems to the same data. Fig. 1B of Banks and Lewbel (1997) plots the budget share of domestic fossil fuel in the United Kingdom as a negative linear function of the logarithm of fuel expenditure. This budget share varies linearly from 0.1 at ln (expenditure) = 4.3 down to 0.04 at ln (expenditure) = 5.5. Consequently, the budget share of polluting goods declines, but the budget share of clean goods rises in the logarithm of total expenditure. For our illustration, we let domestic fossil fuel correspond to 'polluting' goods and we lump food, clothing, alcohol and other consumer items together as 'clean' goods. This gives estimates $\chi_0 = 0.315$ and $\chi_1 = 0.05$ for the PIGLOG expenditure function.

As shown in Fig. 1, we approximate the LES system in eq. (18) by fitting a least squares regression $b_n = (1-a)\overline{b} + am_n$ (dashed line) to the data generated by the PIGLOG system (solid line). This gives $a = 0.0146$ and $\overline{b} = 7.098$ ($R^2 = 0.77$). Fig. 1 shows that estimated demand for polluting goods with the LES demand system is too high for low and high levels of expenditures and too low for medium levels of expenditures. Since we assume quasi-linear utility, the private marginal utility of income is constant. Consequently, only the government's aversion to inequality generates a motive for income redistribution. Further details on the calibration of labour supply, inequality, environmental damages and the social welfare weights are presented in Appendix F.

6.3. Optimal taxation with LES demand

Table 1 gives the results if commodity demands follow from the LES for the baseline and various sensitivity checks. The main finding is that the pollution tax is always set at its Pigouvian level, i.e. $q = q^P = p(q)D'(b)$. As prescribed by Corollary 1,

---

[34] One non-parametric analysis of a long time series of consumer expenditure patterns in the U.K. suggests that Engel curves are quadratic in the log of expenditure (Banks and Lewbel, 1997). This corresponds to QUAIDS demand systems, which generalises PIGLOG and AIDS demand systems, and are of the non-Gorman polar form.

Table 1
Optimal pollution and income tax system with LES demand.

	Baseline	Less inequality aversion ($\psi = 2$)	More income inequality ($\varpi = 0.8$)	Larger pollution damages ($\Xi_0 = 0.2$)	Regressive pollution incidence ($\Xi_1 = 1$)	Higher labour elasticity ($\vartheta = 0.35$)	Lower public revenue ($R = 30$)
Marginal tax rate t (%)	65.9	55.7	77.9	65.9	65.9	59.4	65.9
Income transfer s	120	103	154	121	121	185	140
Av. income tax (%)	19.1	17.9	18.8	18.8	18.8	12.4	11.3
Pollution tax q (%)	10.0	10.0	10.0	20.1	18.3	10.0	10.0
Distr. char. l ξ_l	0.48	0.31	0.88	0.48	0.48	0.51	0.48
Distr. char. b ξ_b	0.06	0.05	0.07	0.05	0.06	0.09	0.06
Av. labour supply	2.40	2.56	2.07	2.39	2.40	3.62	2.40
Pollution b	9.75	9.98	9.80	9.53	9.56	11.58	10.01

we know that for LES demand the optimal pollution tax is directed exclusively at internalising pollution externalities, and not at income redistribution. The linear income tax redistributes income with a positive marginal tax rate and positive lump-sum transfers. The income tax is progressive, since the marginal tax rate always exceeds the average tax rate. Our simulations confirm that higher inequality aversion, more ex-ante inequality, a higher labour supply elasticity, or a larger public revenue requirement do not change the optimal pollution tax. The income tax system does adjust, however.

Halving inequality aversion (reducing ψ from 4 to 2) reduces the demand for redistribution and thus leads to a lower income tax rate and a lower lump-sum transfer. Consequently, there are fewer income tax distortions and more economic activity and pollution. The pollution tax does not adjust to fight the increase in pollution. More pre-tax income inequality (increasing ϖ from 0.4 to 0.8) has the opposite effects. Aggregate pollution increases despite the fall in output because there are now more poor people that spend relatively more on polluting goods. Again, the pollution tax does not react to the rise in total pollution.

If pollution damages double (increasing Ξ_0 from 0.1 to 0.2), the optimal pollution tax roughly doubles, but the income tax rate and the lump-sum transfer hardly change. Aggregate economic activity is not much affected, but aggregate pollution falls due to the higher pollution tax.

A higher wage elasticity of labour supply (raising ϑ from 0.25 to 0.35) raises distortions of income taxation and induces a lower income tax rate, since the government faces a more severe trade-off between equity and efficiency. It thus sacrifices on income redistribution. As a result, aggregate economic activity expands and pollution increases. A higher elasticity also raises the lump-sum transfer. The main reason is that average labour supply increases with more than 50 percent. This is mainly a mechanical increase in labour supply due to changing the utility function, and only partly a behavioural response.[35] With Gorman polar preferences, the optimal pollution tax does not respond to this additional pollution.

A smaller public revenue requirement (cutting R from 50 to 30) leaves the income tax rate (and the pollution tax) unaffected and, due to the assumption of quasi-linearity, leads to a one-to-one increase in lump-sum transfers. Economic activity is unaffected, but higher net incomes lead to more spending, and thus to more aggregate pollution.

Hence, with LES demand the government does not use the pollution tax for redistribution and sets the optimal pollution tax at the Pigouvian level even if the poor spend relatively more on polluting goods.

6.4. Optimal taxation with PIGLOG demand

Table 2 presents the results for PIGLOG demand. Comparing the benchmark outcome with LES in Table 1, the optimal pollution tax now differs from the Pigouvian tax, since the Engel curves are now non-linear. The benchmark optimal pollution tax is 0.8 percentage point lower than the Pigouvian tax. With PIGLOG preferences, the government thus also employs the pollution tax to redistribute incomes and alleviate labour-market distortions caused by the income tax. The lump-sum transfer is not that different compared to the optimal transfer with the LES, since income and pollution taxes are not that different and there are no income effects in labour supply.

To gain more intuition for our results, we rewrite eq. (8) in Proposition 1 as:

$$\frac{q - q^P}{1 + q} \underbrace{\left(\frac{-\overline{\beta \varepsilon_{bq}}}{\overline{\beta}} \right)}_{+} = \underbrace{\xi_b}_{(-)} + \underbrace{\frac{t}{(1-t)} \frac{\overline{\varepsilon_{lq}}}{-\overline{\varepsilon_{lt}\beta}}}_{+}, \tag{21}$$

where we have signed the various terms of the optimal pollution tax for the PIGLOG case, based on our simulations. The left-hand side of eq. (21) gives the marginal deadweight loss of distorting polluting goods demand away from the Pigouvian level.

---

[35] If taxes decline from 66 to 59%, and the labour supply elasticity is 0.35, average labour supply approximately increases with $dl/l = -\varepsilon_{lt} \times dt/(1-t) = 7$ percent. This is a fraction of the increase in labour supply in the simulations.

Table 2
Optimal pollution and income tax system with PIGLOG demand.

	Baseline	Less inequality aversion ($\psi = 2$)	More income inequality ($\varpi = 0.8$)	Larger pollution damages ($\Xi_0 = 0.2$)	Regressive pollution incidence ($\Xi_1 = 1$)	Higher labour elasticity ($\vartheta = 0.35$)	Lower public revenue ($R = 30$)
Marginal tax rate t (%)	65.9	55.6	77.8	65.8	65.8	59.3	65.9
Income transfer s	120	103	154	120	120	185	140
Av. income tax (%)	19.2	18.0	18.8	18.9	19.0	12.5	11.4
Pollution tax q (%)	9.2	8.3	12.0	16.6	15.3	9.9	8.8
Pigouvian tax q^P (%)	10.0	10.0	10.0	20.0	18.2	10.0	10.0
Distr. char. l ξ_l	0.48	0.31	0.88	0.48	0.48	0.51	0.48
Distr. char. b ξ_b	−0.01	−0.03	−0.03	−0.01	−0.01	−0.30	−0.03
Av. labour supply	2.40	2.56	2.07	2.40	2.40	3.73	2.40
Pollution b	8.931	8.767	8.465	8.365	8.461	6.601	8.906

The government optimally sets a larger (smaller) gap between the pollution tax and the Pigouvian tax, the less (more) elastic is the demand for polluting goods is, i.e. if $-\overline{\beta \varepsilon_{bq}}/\overline{\beta}$ is smaller (larger) − ceteris paribus benefits of doing this the right-hand side of eq. (21). Intuitively, deviations from the Pigouvian tax generate smaller (larger) distortions in demand for polluting goods if the demand for polluting goods is less (more) elastic.

The right-hand side of eq. (21) gives the benefits in terms of redistribution or lower labour market distortions of setting pollution taxes at non-Pigouvian rates. First, with PIGLOG preferences, polluting goods can be inferior. Expenditure on dirty goods first rises and then declines with income, as also shown in Fig. 1. In our simulations, the distributional characteristic for polluting goods ξ_b is negative. Hence, the inferiority of polluting goods at high levels of income dominates the normality of polluting goods at lower incomes. Consequently, the government likes to lower the tax on polluting goods below the Pigouvian level for redistributional reasons − ceteris paribus. Second, polluting goods and leisure are found to be (strong) complements with PIGLOG preferences, since labour supply increases if we raise the pollution tax, i.e. $\varepsilon_{lq} > 0$, which follows from eq. (21). The government thus likes to set pollution taxes above the Pigouvian rate to alleviate distortions of the income tax on labour supply − ceteris paribus. This is the Corlett and Hague (1953) motive for differential commodity taxation. The gap between the optimal pollution tax and the Pigouvian tax thus depends on these two opposing effects and can be positive or negative, as we can see in Table 2.

Cutting income inequality aversion in half (reducing ψ from 4 to 2) implies that the government has a smaller desire to redistribute income. The income tax system becomes less progressive due to a lower income tax rate and lump-sum transfer. The optimal pollution tax slightly decreases for two reasons. First, less redistribution implies that net incomes of high-income individuals increase − for whom polluting goods are inferior − while net incomes of low-income individuals decrease − for whom polluting goods are normal. On average polluting goods are thus more inferior, so that the distributional characteristic of polluting goods decreases, and optimal pollution taxes should be lowered. Second, labour market distortions are less severe if labour taxes are lower. Consequently, the optimal tax on polluting goods falls, since there is less reason to tax polluting goods to alleviate labour market distortions.

Doubling pre-tax inequality (raising ϖ from 0.4 to 0.8) makes the optimal income tax system more progressive. Intuitively, the distributional characteristic of labour income increases substantially and a higher income tax rate is optimal. The optimal pollution tax is raised above the Pigouvian pollution tax for two reasons. First, larger pre-tax inequality lowers the distributional characteristic of the polluting good, because it becomes more inferior − see also previous paragraph where we discussed the effect of larger after-tax inequality. Hence, larger pre-tax inequality tends to lower the pollution tax. Second, labour-market distortions increase if optimal labour taxes are set at higher rate. For that reason, the optimal tax on polluting goods should increase, since taxing polluting goods has larger benefits in terms of reducing labour market distortions. The last effect dominates, so that optimal pollution taxes are driven above the Pigouvian rates. Doubling marginal pollution damages (from $\Xi_0 = 0.1$ to 0.2) leads to a doubling of the Pigouvian tax, but the optimal pollution tax falls relatively more below the Pigouvian tax compared to the baseline. If global warming damages hit the poor relatively more ($\Xi_1 = 1$ instead of 0), the equity-weighted Pigouvian tax increases from 10 percent to 18.2 percent, since the government cares relatively more about the poor more than about the rich. As with doubling marginal damages, the gap between the optimal pollution tax and the Pigouvian increases. Aggregate economic activity is hardly affected in both cases compared to the baseline, but aggregate pollution falls substantially.

A higher wage elasticity of labour supply ($\vartheta = 0.35$ instead of 0.25) increases the efficiency loss from income taxation and thus lowers the optimal income tax rate. A higher elasticity also raises the lump-sum transfer, because average incomes increase due to a near doubling of labour supply, which is mainly the mechanical result of changing the utility function. The pollution tax increases compared to the baseline, but still slightly below the Pigouvian rate. Intuitively, a higher elasticity of labour supply indicates that pollution taxes should be increased to reduce larger labour-market distortions. On the other hand, a higher distributional characteristic of polluting goods − because polluting goods have become more inferior on average − calls for lower pollution taxes to redistribute more income. The net effect is that pollution taxes increase compared to the baseline.

Whether optimal corrective taxes should substantially deviate from Pigouvian taxes is thus an empirical question. The gap between the optimal corrective tax and the Pigouvian tax may be negative or positive, depending on whether polluting goods are normal or inferior and whether they are stronger complements to leisure than clean goods are. Setting the corrective tax at the Pigouvian rate is a useful first approximation and doing so is certainly better than setting no pollution tax at all. However, our simulations also demonstrate that one must be careful in drawing policy conclusions if preferences are not of the Gorman polar form. If the data for the demand for polluting goods in fact suggest non-Gorman polar preferences, then setting pollution taxes at the Pigouvian rate can be misleading.

7. Robustness

This section discusses how our results are affected if several important assumptions in our model are relaxed. First, we discuss the implications of constraining the government instruments to exclude non-individualised lump-sum transfers. Second, we analyse the consequences of expanding the government instrument set by allowing for non-linear income taxes. Third, we analyse preference heterogeneity. Fourth, we verify whether our results can be generalised to production externalities. Fifth, we discuss the implications of general-equilibrium effects on wages and prices. Sixth, we discuss the role of pre-existing tax and non-tax distortions for the desirability of pollution taxes and green tax reforms.

7.1. Constraints on lump-sum transfers

The double dividend literature analyses corrective taxes in models with homogenous agents, where non-individualised lump-sum transfers are not available or are fixed at levels before pollution was a concern (e.g. Sandmo, 1975; Bovenberg and De Mooij, 1994; Bovenberg and van der Ploeg, 1994; Goulder, 1995; Bovenberg, 1998, 1999; Sandmo, 2000). If non-individualised lump-sum transfers are no longer available, the government thus loses its non-distortionary marginal source of public finance. Consequently, the marginal cost of public funds is no longer equal to one and governments face a trade-off between correcting externalities and redistributing income (Jacobs and de Mooij, 2015; Proposition 3). Hence, both the distortionary income and pollution taxes are needed to raise government revenue in the most efficient manner.

Whereas ruling out individualised lump-sum transfers is necessary to obtain a second-best problem in representative-agent models, such constraints on the government instrument set are no longer needed if heterogeneous agents are allowed for, as this and other papers have shown (Diamond, 1973; Pirttilä and Schöb, 1999; Sandmo, 2000; Jacobs and de Mooij, 2015). Moreover, although one might have objections to allowing for lump-sum *taxes*, it is not clear from a political or policy perspective why the government cannot provide its citizens with lump-sum *transfers*. We have shown that the optimal redistributive tax system generally features a positive lump-sum transfer and positive marginal tax rates to make tax systems progressive. Similarly, it is not clear why these transfers cannot be reduced at the margin. Most real-world tax-benefit systems feature a general tax exemption/credit or a guaranteed minimum income, all of which can be marginally reduced. Therefore, we think that allowing for non-individualised transfers as part of the government instrument set is a realistic and important extension of existing studies that assume such transfers away.

It is of course not clear whether governments optimise income taxes in the real world. We have shown that in that case the *MCPF* can be both above one (excessive redistribution) or below one (insufficient redistribution). Ultimately, it requires a political judgement to assess whether income taxes are too high or too low. More research can shed light on the deeper political reasons why certain policy instruments can and why others cannot be used.

7.2. Non-linear income taxation

Our paper complements other contributions on optimal corrective taxation and optimal income redistribution under non-linear, Mirrleesian income taxes. Cremer et al. (1998, 2003), Micheletto (2008) and Jacobs and de Mooij (2015) show that Proposition 1 of our paper is sensitive to assuming linear income taxes: if the government can levy non-linear income taxes, then optimal pollution taxes should be set at Pigouvian rates and should not take into account distributional concerns (except for the equity weighting of pollution damages). Moreover, the structure of the optimal non-linear income tax should be the same as in the absence of pollution. Intuitively, a fully non-linear tax can always compensate individuals for any increase in pollution taxes as long as preferences of individuals are the same. However, in the real world, a fully non-linear Mirrleesian income tax system is generally not implemented.[36] Governments generally levy (piece-wise) linear income taxes. Therefore, our analysis of linear instruments becomes relevant, since the piece-wise linear tax schedule resembles the linear tax more than the non-linear tax.[37]

To see why, the government has access to a flat rate and a bracket-specific intercept of the tax schedule with a piece-wise linear tax system. Therefore, to be able to neutralise the distributional impact of a pollution tax with a piece-wise linear tax system, the same conditions apply as in the linear tax system: Engel curves for polluting goods should be linear, but now *within each tax bracket*. The slopes and the intercepts of the Engel curves can thus vary across brackets. However, if Engel

---

[36] An exception is Germany with a quadratic tax schedule up to an upper boundary.

[37] Our analysis extends in modified form to piece-wise linear tax schedules. A piece-wise linear schedule contains a number of flat rates and tax brackets. The bracket length implicitly defines a separate non-individualised transfer for each tax bracket.

curves are not linear within brackets, then even piece-wise linear tax schedules cannot undo the distributional consequences of pollution taxes. Naturally, as the number of tax brackets becomes large, a piece-wise linear tax system approximates a non-linear schedule better and can better off-set the distributional impacts of a pollution tax. But given that real-world tax systems typically have a limited number of tax brackets, the case with linear or piece-wise linear taxes is in our view the natural benchmark to analyse redistribution and corrective taxation in real-world applications.

Furthermore, our paper relates to Kaplow (2012). Kaplow analyses non-linear income taxes and corrective taxation. He shows that distribution-neutral green tax reforms are *always* feasible under weakly separable preferences if non-linear taxes are not optimal and if preferences are identical, but not of the Gorman polar form. Intuitively, the government can adjust the non-linear tax schedule at each point in the income distribution to fully compensate for the change in the corrective tax if preferences are identical across individuals. Moreover, if preferences are weakly separable, such a distribution-neutral tax change does not impact labour supply incentives. A distributional-neutral change in the corrective tax thus neither affects inequality nor tax distortions. Therefore, if the corrective tax is below (above) the Pigouvian level, an increase (decrease) in corrective taxation is desirable. Kaplow's analysis can in general not be applied to linear tax schedules since linear schedules cannot be perfectly tailored to neutralise all distributional effects of corrective taxes. Moreover, Kaplow assumes sub-optimal non-linear income taxes, while at the same time the non-linear income tax schedule can be adjusted at each income level. If there are valid reasons why the non-linear tax schedule cannot be optimised, then it may not be possible to implement perfect benefit-absorbing tax changes for the same reasons. We complement Kaplow (2012) by identifying the conditions under which benefit-absorbing tax changes do in fact exist under much more restricted linear tax schedules. We show that this requires preferences to be of the Gorman polar form. Moreover, we assume that the constraints on setting the income tax are the same before and after the distribution-neutral green tax reforms.

7.3. Preference heterogeneity

Our model assumed separable and identical preferences for clean goods, polluting goods, leisure and environmental quality. Cremer and Gahvari (2001) and Cremer et al. (2003) demonstrate that corrective taxes need to take into account the distributional concerns if preferences for polluting goods are heterogeneous and taxes are non-linear.[38] Intuitively, the government can then no longer perfectly compensate every individual for a change in the pollution tax. Hence, pollution taxes complement the non-linear income tax to redistribute income, just as in our analysis with linear income taxes with non-Gorman polar preferences. In contrast to the literature on optimal non-linear taxation, our Proposition 1 remains the same if we have heterogeneous preferences (while maintaining separability between private goods and the environment).[39] Therefore, optimal linear income and pollution taxes are robust to including preference heterogeneity, in contrast to the papers that analyse non-linear income taxes. However, preferences are no longer of the Gorman polar form if the utility function is heterogeneous across types. Hence, our Propositions 2 and 3 do not generalise if we allow for preference heterogeneity.

7.4. Production externalities

One may wonder how robust our findings are with respect to the assumption that all externalities are non-pecuniary and originate from consumption of polluting goods. This section shows that our results carry over in only slightly modified form to the case where pollution originates in the production process. To see why, suppose that there is a representative firm that maximises profits using a constant-returns-to-scale production technology with aggregate labour $L \equiv \int_0^\infty nl_n f(n) \mathrm{d}n$ as its only input. The production function is designated by $Y = A(b)L$, where labour productivity A depends negatively on aggregate pollution b, i.e. $A'(b) < 0$. Firms take aggregate pollution b as given. Profit maximisation implies that the wage rate per efficiency unit of labour is given by $w = A(b)$, so that labour demand is perfectly elastic at the wage rate w. The individual now maximises $u(c_n, b_n, 1 - l_n)$ subject to $c_n + (1 + q)b_n = (1 - t)A(b)nl_n + s$. As a result, environmental damages enter indirect utility $v(t, q, s, b, n)$ via the wage rate per efficiency unit of labour $A(b)$. Appendix G derives that the optimal tax formulae of Proposition 1 apply if we redefine the social welfare weights as $g_n^* \equiv g_n + tA(b)n\frac{\partial l_n}{\partial s} + (q - q^P)\frac{\partial b_n}{\partial s}$ and expenditure share as $\beta_n \equiv \frac{(1+q)b_n}{(1-t)A(b)nl_n}$. Moreover, the Pigouvian tax under production externalities is modified to:

$$q^P \equiv \frac{\mu}{\lambda} = -A'(b) \int_0^\infty nl_n f(n) \mathrm{d}n = -\frac{A'(b)}{A(b)} Y. \tag{22}$$

---

[38] This argument for differential commodity taxation under heterogeneous preferences goes back to Mirrlees (1976) and Saez (2002) who show that governments should use commodity taxes for income redistribution, besides non-linear income taxes, if individual preferences for commodities are heterogeneous.
[39] This follows from redoing the derivations in Appendix A with utility function $U_n = u(c_n, b_n, 1 - l_n, n) - D(b, n)$. This also gives indirect function $v(t, q, s, b, n)$. The entire derivation thus remains the same.

Eq. (22) also implies that production is efficient, since the social marginal costs of cleaner production − the Pigouvian tax − (left-hand side) are equal to the social marginal benefits of cleaner production (right-hand side). The main difference with consumption externalities is that the Pigouvian tax with production externalities requires no equity weighting of pollution damages, since production externalities affect all wages of workers symmetrically. There is otherwise no difference compared to the model with consumption externalities. Therefore, our finding that optimal corrective taxes does (not) need to take into account distributional concerns with non-Gorman polar (Gorman polar) preferences extends to the case with production externalities.

If one were to allow for multiple production sectors, each causing different externalities, production efficiency would be obtained, also in second-best settings with distortionary taxation, provided the government can tax or subsidise use of polluting inputs each sector (Cremer and Gahvari, 2001). Furthermore, this also implies that abatement investments, reducing the aggregate amount of pollution, should be efficient. Intuitively, investments in abatement are intermediate goods in the production of environmental quality. If production of environmental quality should be efficient, then investment in abatement should also satisfy first-best principles in second-best settings (Diamond and Mirrlees, 1971a,b). If the conditions for the production-efficiency theorem would fail, see next point, then distributional concerns will come into play in setting optimal corrective taxes. We leave this as an extension for future research.

7.5. General equilibrium

We have abstracted from general-equilibrium effects on factor prices, which could affect the distributional consequences of taxes on income and pollution (see, for example, Fullerton and Heutel, 2007; Fullerton and Monty, 2013). However, our partial-equilibrium assumption is without loss of generality if Diamond and Mirrlees (1971a, b) production efficiency theorem applies. In that case, the optimal tax formulae are the same in general as in partial equilibrium.

The Diamond-Mirrlees production efficiency theorem rests on the following two premises. First, production takes place under constant returns to scale or, if production takes place under decreasing returns to scale, profits should be fully taxed away.[40] If the government cannot tax all pure profits, the initial tax system is not optimal from a revenue-raising perspective, and interactions between correcting this initial tax distortion and optimal taxes on income and pollution emerge. Optimal taxes then also aim to shift the tax burden towards untaxed profits (Goulder, 1995; Bovenberg, 1998, 1999). See also next subsection.

Second, all transactions between firms and households should be verifiable so that consumer prices can be set independently from producer prices. This condition is violated if labour types are imperfect substitutes in production and the government cannot tax/subsidise each occupation separately. In that case, distributional effects originating from general-equilibrium effects on factor or commodity prices might become relevant for evaluating green tax reforms. For example, Fullerton and Monty (2013) analyse pollution taxes in a general-equilibrium model with high- and low-skilled occupations. Their simulations show that an increase in pollution taxes, while rebating all revenue to low-skilled labour does not prevent a reduction in their overall net wage, since pollution taxes reduce their before-tax wage rate. Similarly, Gahvari (2014) analyses optimal non-linear taxation with endogenous wages. Since the government cannot condition the tax schedule on the occupation of the individual, the second condition for production efficiency does not hold. Consequently, optimal pollution taxes take into account the general-equilibrium effects on wage rates.[41] However, if the government has sufficient instruments to tax/subsidise each occupational type separately, for example via education subsidies or occupation-specific tax schedules, the government no longer relies on production distortions to redistribute income, and production efficiency is restored (Scheuer and Werning, 2016). Intuitively, production distortions emerge only if the government lacks the instruments needed to directly redistribute between different occupations.

Our results thus generalise to settings with general-equilibrium effects on wages and prices if the government can tax or subsidise each factor in the production function with a separate schedule and if there are no untaxed pure profits in production. However, if these conditions are not applicable, then general-equilibrium effects will come into play in setting optimal corrective taxes. In particular, optimal income and corrective taxes exploit general-equilibrium effects for income redistribution (Allen, 1982; Stiglitz, 1982; Tsyvinski et al., 2016). We leave such an extension for future research.[42]

7.6. Pre-existing distortions

We abstracted from pre-existing tax and non-tax distortions in the economy to analyse optimal pollution taxes and green tax reforms. However, from the double dividend literature follows that untaxed fixed factors, rents, complementarity of

---

[40] If there is no profit tax, there will be scarcity rents so that the production efficiency theorem no longer holds and environmental command-and-control regulations have benefits (e.g. Fullerton and Metcalf, 2001).

[41] Chiroleu-Assouline and Fodha (2014) analyse green tax reforms in a very different setup than the standard Mirrleesian optimal tax framework. They study an OLG model with exogenous labour supply (hence labour taxes are lump sum), homogeneous production externalities (rather than heterogeneous consumption externalities), heterogeneous discount rates and wage rates. Labour types are perfect substitutes in production. They find that Pareto-improving green tax reforms could be feasible with the proper adjustments of the income tax.

[42] Ruling out factor-specific taxes and allowing for non-taxed pure profits and general-equilibrium effects has also implications for the double dividend debate, since green tax reforms that would shift the tax burden towards untaxed profits or provoke general-equilibrium effects that compress the wage distribution become more desirable.

polluting goods with leisure, complementarity of work with the environment and unemployment, can also be reasons why pollution taxes are desirable and why a strong double dividend of green tax reforms may exist (Bovenberg and van der Ploeg, 1994; Goulder, 1995; Bovenberg, 1998, 1999; Fullerton and Metcalf, 2001). If pollution taxes alleviate pre-existing tax and non-tax distortions in the economy, both environmental and non-environmental welfare may increase in a green tax reform. Our paper analyses distributional concerns. In such a setting, we argued that the weak and strong double dividend of green tax reforms are not relevant in the tax optimum, since tax distortions are present for distributional concerns. However, if there are other tax and non-tax distortions, from which we abstracted, it is still not clear whether this affects the desirability of pollution taxes or green tax reforms. This ultimately depends on whether income inequality rises or falls if pre-existing tax and non-tax distortions in the economy are alleviated. It will be an interesting avenue for further research to explore various tax and non-tax distortions and income inequality in models of optimal pollution taxes and green tax reforms.

8. Conclusions

Our main conclusion is that corrective pollution taxes should take account of distributional concerns if Engel curves for polluting goods are non-linear, which is the empirically relevant case (e.g. Banks and Lewbel, 1997; Tovar Reaños and Wölfing, 2018). Moreover, pollution taxes should be at their Pigouvian level, and not be directed at redistribution, if Engel curves are linear, despite a possible regressive incidence (e.g. as with LES demand). Engel curves are linear with Gorman polar preferences, which include quasi-linear, CES and Stone-Geary preferences as special cases. We show that if Engel curves are linear, a Pareto-improving green tax reform is possible provided the pollution tax is not set at the Pigouvian pollution tax. With non-linear Engel curves and linear tax systems it is generally not possible to have Pareto-improving tax reforms.

Three directions of further research are important. First, estimation of commodity demand systems should be undertaken using flexible forms for demand functions that allow for non-linear Engel curves (e.g. Lewbel and Pendakur, 2009; Tovar Reaños and Wölfing, 2018). Only with more accurate information regarding the shape of the Engel curve for polluting goods demands, one can determine whether deviations from the equity-weighted Pigouvian tax are really socially desirable. Second, our normative analysis of optimal pollution taxation and income distribution should be complemented with the analysis of political economy aspects of environmental tax reforms given the widespread political resistance to tax fossil fuel emissions, for example the protests of the yellow vests in France. Political constraints may prevent setting the optimal second-best pollution tax even if Engel curves are linear and individuals can be compensated for the distributional consequences of green tax reforms. Third, understanding behavioural biases may be important to increase social support for green tax reforms. For example, individuals may respond differently to general income tax cuts than a clearly visible, transparent and uniform "carbon dividend" that is handed out to all citizens (e.g. Klenert et al., 2018b).

Appendix A. Proof of Proposition 1

The Lagrangian Λ for the maximisation of social welfare is defined by − where we use aggregate pollution b as a control variable as in Jacobs and de Mooij (2015):

$$\max_{\{t,q,s,b\}} \Lambda \equiv \int_0^\infty \omega_n \mathrm{v}(t,q,s,b,n) + \lambda[tnl_n + qb_n - s - R]\mathrm{f}(n)\mathrm{d}n + \mu\left[b - \int_0^\infty b_n\mathrm{f}(n)\mathrm{d}n\right]. \tag{A1}$$

The following necessary first-order conditions have to hold:

$$\frac{\partial \Lambda}{\partial s} = \int_0^\infty \left[\omega_n \eta_n - \lambda + \lambda tn\frac{\partial l_n}{\partial s} + (\lambda q - \mu)\frac{\partial b_n}{\partial s}\right]\mathrm{f}(n)\mathrm{d}n = 0, \tag{A2}$$

$$\frac{\partial \Lambda}{\partial t} = \int_0^\infty \left[-\omega_n \eta_n nl_n + \lambda nl_n + \lambda tn\frac{\partial l_n}{\partial t} + (\lambda q - \mu)\frac{\partial b_n}{\partial t}\right]\mathrm{f}(n)\mathrm{d}n = 0, \tag{A3}$$

$$\frac{\partial \Lambda}{\partial q} = \int_0^\infty \left[-\omega_n \eta_n b_n + \lambda b_n + \lambda tn\frac{\partial l_n}{\partial q} + (\lambda q - \mu)\frac{\partial b_n}{\partial q}\right]\mathrm{f}(n)\mathrm{d}n = 0, \tag{A4}$$

$$\frac{\partial \Lambda}{\partial b} = \int_0^\infty \left[-\omega_n \mathrm{D}_b(b,n) + \mu + \lambda tn\frac{\partial l_n}{\partial b} + (\lambda q - \mu)\frac{\partial b_n}{\partial b}\right]\mathrm{f}(n)\mathrm{d}n = 0, \tag{A5}$$

where we have used Roy's identity in each line to substitute for the derivatives of indirect utility.

From the first-order condition for b in (A5), it follows that the Pigouvian pollution tax equals

$$q^P \equiv \frac{\mu}{\lambda} = \frac{\int_0^\infty \omega_n D_b(b,n) f(n) dn}{\lambda} = \int_0^\infty \frac{\omega_n \eta_n}{\lambda} \frac{D_b(b,n)}{\eta_n} f(n) dn = \int_0^\infty g_n \frac{D_b(b,n)}{\eta_n} f(n) dn. \tag{A6}$$

From the first-order condition for s in (A2) follows that the marginal cost of public funds equals one:

$$MCPF = 1 \left/ \left(\int_0^\infty g_n + tn \frac{\partial l_n}{\partial s} + (q - q^P) \frac{\partial b_n}{\partial s} f(n) dn \right) \right. = 1. \tag{A7}$$

We simplify the first-order conditions for t and q, i.e. (A3) and (A4), in a number of steps. First, we substitute the Slutsky eqs. in eq. (A3) and eq. (A4) and use $q^P = \mu/\lambda$:

$$\frac{\partial \Lambda}{\partial t} = \int_0^\infty \left[-\omega_n \eta_n n l_n + n l_n + \lambda t n \left(\frac{\partial l_n^c}{\partial t} - n l_n \frac{\partial l_n}{\partial s} \right) + \lambda (q - q^P) \left(\frac{\partial b_n^c}{\partial t} - n l_n \frac{\partial b_n}{\partial s} \right) \right] f(n) dn = 0. \tag{A8}$$

$$\frac{\partial \Lambda}{\partial q} = \int_0^\infty \left[-\omega_n \eta_n b_n + \lambda b_n + \lambda t n \left(\frac{\partial l_n^c}{\partial q} - b_n \frac{\partial l_n}{\partial s} \right) + \lambda (q - q^P) \left(\frac{\partial b_n^c}{\partial q} - b_n \frac{\partial b_n}{\partial s} \right) \right] f(n) dn = 0. \tag{A9}$$

Then, substitute definitions for the Diamond social welfare weights in eq. (5) and the compensated elasticities:

$$\int_0^\infty \left[(1 - g_n^*) n l_n + \frac{t}{(1-t)} \varepsilon_{lt} n l_n + \frac{(q - q^P)}{(1+q)} \frac{(1+q) b_n}{(1-t) n l_n} \varepsilon_{bt} n l_n \right] f(n) dn = 0. \tag{A10}$$

$$\int_0^\infty \left[(1 - g_n^*) b_n + \frac{t}{(1-t)} \frac{(1-t)}{(1+q)} \varepsilon_{lq} n l_n + \frac{(q - q^P)}{(1+q)} \frac{b_n}{n l_n} \varepsilon_{bq} n l_n \right] f(n) dn = 0. \tag{A11}$$

Finally, use the definitions for the distributional characteristics ξ_l and ξ_b in eqs. (6a) and (6b), and use the expenditure quote $\beta_n \equiv \frac{(1+q) b_n}{(1-t) n l_n}$ to find the results stated in Proposition 1.

Appendix B. Examples of Gorman Polar Form Preferences

Quasi-linear preferences: The sub-utility function for this case is $h(c_n, b_n) = c_n + k(b_n), k' > 0, k'' < 0$. Maximising utility subject to the private budget constraint $c_n + (1+q) b_n = m_n$ yields the first-order condition $k'(b_n) = 1 + q$, so demand for polluting goods is $b_n = \kappa(q)$ with $\kappa'(q) = 1/k''(b_n)$. Demand for clean goods follows residually from the private budget constraint: $c_n = m_n - (1+q)\kappa(q)$. Hence, the indirect utility function belongs to the class of Gorman polar preferences as $\upsilon(q, m_n) = (m_n - \varphi(q))/p(q) = m_n - (1+q)\kappa(q) + k(\kappa(q))$ with $p(q) = 1$ and $\varphi(q) = (1+q)\kappa(1+q) - k(\kappa(1+q))$. This is a special case of the demand for polluting dirty goods in eq. (12) with $\alpha(q) \equiv (1+q)\kappa(1+q)$ and $\phi(q) = 0$.

Standard Stone-Geary preferences: The sub-utility function for this case is $h(c_n, b_n - \overline{b}) = (c_n/(1-a))^{1-a}((b_n - \overline{b})/a)^a$ with $0 < a < 1$ and $\overline{b} > 0$. Maximising utility subject to the private budget constraint gives the first-order condition $((b_n - \overline{b})/a)^{a-1} = (1+q)(c_n/(1-a))^{-a}$. The demand functions for polluting and for clean goods are: $b_n = (1-a)\overline{b} + am_n/(1+q)$ and $c_n = (1-a)(m_n - (1+q)\overline{b})$. Upon substitution into the utility function, the Gorman polar form indirect utility function is obtained: $\upsilon(q, m_n) = (m_n - \varphi(q))/p(q) = (1+q)^{-a}(m_n - (1+q)\overline{b})$, so that $p(q) = (1+q)^a$ and $\varphi(q) = (1+q)\overline{b}$. This yields a special case of the demand function in eq. (12) with $\phi(q) = a$ and $\alpha(q) \equiv (1+q)(1-a)\overline{b}$. Using $m_n = (1+q)^a h_n + (1+q)\overline{b}$, the compensated demand for polluting goods follows as $b_n = \overline{b} + a(1+q)^{a-1} h_n$.

Generalised Stone-Geary preferences: The sub-utility function for this case is $h(c_n, b_n - \overline{b}) = [(1-a)c_n^{(\sigma-1)/\sigma} + a(b_n - \overline{b})^{(\sigma-1)/\sigma}]^{\sigma/(\sigma-1)}$ with $0 < a < 1$ and $\sigma > 0$. Maximising utility subject to the budget constraint $c_n + (1+q) b_n = m_n$ gives the optimality condition $b_n - \overline{b} = ((1+q)(1-a)/a)^{-\sigma} c_n$. Solving this together with the budget constraint gives the Marshallian demand functions:

$$b_n = \frac{m_n/(1+q)}{1+(1+q)^{\sigma-1}((1-a)/a)^\sigma} + \frac{\overline{b}}{1+(1+q)^{1-\sigma}(a/(1-a))^\sigma},\tag{B1}$$

$$c_n = \frac{m_n}{1+(1+q)^{1-\sigma}(a/(1-a))^\sigma} - \frac{(1+q)\overline{b}}{1+(1+q)^{1-\sigma}(a/(1-a))^\sigma}.\tag{B2}$$

Substituting eq. (B1) and eq. (B2) into the utility function, one finds the Gorman polar form indirect utility function $\upsilon(q, m_n) = (m_n - \varphi(q))/\mathrm{p}(q) = (1-a)^{\sigma/(\sigma-1)}(1 + (1+q)^{1-\sigma}(a/(1-a))^\sigma)^{1/(\sigma-1)}(m_n - (1+q)\overline{b})$ with

$$\mathrm{p}(q) = (1-a)^{\sigma/(\sigma-1)}\left(1 + (1+q)^{1-\sigma}(a/(1-a))^\sigma\right)^{1/(1-\sigma)} \text{ and } \varphi(q) = (1+q)\overline{b}.\tag{B3}$$

Demand for polluting goods is thus given by eq. (12) with

$$\phi(q) = \left(1 + (1+q)^{\sigma-1}((1-a)/a)^\sigma\right)^{-1} \text{and } \alpha(q) = \left(1 + (1+q)^{1-\sigma}(a/(1-a))^\sigma\right)^{-1}(1+q)\overline{b}.\tag{B4}$$

Appendix C. Proof of Proposition 2

The Proof requires three steps. First, if preferences are weakly separable and sub-utility belongs to the class of Gorman polar preferences the demand for dirty commodities features linear Engel curves. Second, we derive the elasticities of labour supply with respect to income and the optimal corrective pollution tax for this class of utility functions. Third, with this information, we can solve the first-order conditions to prove the proposition.

C.1. Gorman polar preferences and linear Engel curves

Utility is $\mathrm{u}(\mathrm{h}(c_n, b_n), 1 - l_n) - \mathrm{D}(b, n)$, where sub-utility $\mathrm{h}(c_n, b_n)$ belongs to the class of Gorman polar utility functions. The household budget constraint is given by $(1-t)nl_n + s = c_n + (1+q)b_n \equiv m_n$. We solve this problem in two stages using so-called conditional demand functions (Mirrlees, 1976; Jacobs and Boadway, 2014). In the first stage, individuals optimally choose clean and dirty consumption to maximise utility $\mathrm{u}(\cdot)$ for a given level of labour supply l_n subject to $c_n + (1+q)b_n \equiv m_n$. Labour supply will be determined in the second stage. This gives the following first-order condition:

$$\mathrm{h}_b/\mathrm{h}_c = 1 + q.\tag{C1}$$

The first-order condition, joint with the budget constraint, yield the conditional demand functions for clean and dirty goods, which are functions of the pollution tax and net income: $c_n \equiv c^c(q, m_n)$, $b_n \equiv b^c(q, m_n)$. Due to the weak separability of $\mathrm{u}(\cdot)$, the conditional demand functions are independent from labour supply (Jacobs and Boadway, 2014). These conditional demand functions then give a conditional indirect sub-utility function, which is a function of the pollution tax q and net income m_n only: $v(q, m_n) \equiv \mathrm{h}(c^c(q, m_n), b^c(q, m_n))$. Moreover, from Definition 1 follows that the conditional indirect utility function is given by:

$$\mathrm{h}_n = v(q, m_n) = \frac{m_n - \phi(q)}{\mathrm{p}(q)} = \frac{(1-t)nl_n + s - \phi(q)}{\mathrm{p}(q)}.\tag{C2}$$

In the second stage, individuals thus maximise utility $\mathrm{u}(((1-t)nl_n + s - \phi(q))/\mathrm{p}(q), l_n)$ by optimally choosing labour supply l_n subject to $(1-t)nl_n + s = m_n$. This gives the first-order condition:

$$\frac{-\mathrm{u}_l}{\mathrm{u}_h} = \frac{(1-t)n}{\mathrm{p}(q)}.\tag{C3}$$

C.2. Compensated elasticities with Gorman polar preferences

To find the compensated elasticities with respect to the policy instruments, we log-linearise the first-order conditions in eq. (C1) and eq. (C3) and the utility function $\mathrm{u}(\cdot)$ around an initial equilibrium, while keeping total utility constant. A tilde designates a relative change, e.g. $\tilde{l}_n \equiv dl_n/l_n$, except for the relative changes in the tax rates, which are denoted $\tilde{t} \equiv dt/(1-t)$ and $\tilde{q} \equiv dq/(1+q)$.

Log-linearising the sub-utility function $\mathrm{h}(\cdot)$ gives:

$$\tilde{h}_n = \gamma_c \tilde{c}_n + \gamma_b \tilde{b}_n, \quad \gamma_c \equiv \frac{h_c c_n}{h_n}, \quad \gamma_b \equiv \frac{h_b b_n}{h_n}. \tag{C4}$$

Next, log-linearising the first-order condition for consumption in eq. (C1) gives:

$$\delta^b \tilde{b}_n = \delta^c \tilde{c}_n - \tilde{q}, \quad \delta^c \equiv - \left(\frac{h_{cc} c_n}{h_c} - \frac{h_{bc} c_n}{h_b} \right) > 0, \quad \delta^b \equiv - \left(\frac{h_{bb} b_n}{h_b} - \frac{h_{cb} b_n}{h_c} \right) > 0. \tag{C5}$$

Log-linearising the first-order condition for labour supply in eq. (C3) yields:

$$\tilde{l}_n = \frac{\varepsilon_n}{\zeta_n} \tilde{h}_n - \varepsilon_n \tilde{t} - \varepsilon_n \phi(q) \tilde{q}, \tag{C6}$$

where $\varepsilon_n \equiv (u_{ll} l_n / u_l - u_{hl} l_n / u_h)^{-1}$ is the Frisch elasticity of labour supply, $\zeta_n \equiv (u_{hh} h_n / u_h - u_{lh} h_n / u_l)^{-1}$ captures income effects in labour supply, and from Definition 1 follows that $\phi(q) \equiv p'(q)(1+q)/p(q)$ with Gorman polar utility.

Log-linearising utility $u(h_n, l_n)$ while keeping utility fixed − note that environmental quality $D(\cdot)$ is given for each individual − gives:

$$u_h h_n \tilde{h}_n + u_l l_n \tilde{l}_n = 0. \tag{C7}$$

Substituting the first-order condition for labour supply from eq. (C3) in eq. (C7) yields:

$$\tilde{h}_n = \frac{(1-t)nl_n}{p(q)h_n} \tilde{l}_n = \frac{(1-t)nl_n}{((1-t)nl_n + s - \phi(q))} \tilde{l}_n = \theta_n \tilde{l}_n, \quad \theta_n \equiv \frac{(1-t)nl_n}{(1-t)nl_n + s - \phi(q)}. \tag{C8}$$

Solving the system of equations in eqs. (C4), (C5), (C6), (C8) for the relative changes in $\tilde{c}_n$, $\tilde{b}_n$, $\tilde{l}_n$, and $\tilde{h}_n$ gives:

$$\tilde{c}_n = -\frac{\delta^b}{\delta^c \gamma_b + \delta^b \gamma_c} \theta_n \left(\frac{\tilde{t} + \phi(q)\tilde{q}}{\frac{1}{\varepsilon_n} - \frac{\theta_n}{\zeta_n}} \right) + \frac{\gamma_b}{\delta^c \gamma_b + \delta^b \gamma_c} \tilde{q},$$

$$\tilde{b}_n = -\frac{\delta^c}{\delta^c \gamma_b + \delta^b \gamma_c} \theta_n \left(\frac{\tilde{t} + \phi(q)\tilde{q}}{\frac{1}{\varepsilon_n} - \frac{\theta_n}{\zeta_n}} \right) - \frac{\gamma_c}{\delta^c \gamma_b + \delta^b \gamma_c} \tilde{q}, \tag{C9}$$

$$\tilde{l}_n = -\frac{\tilde{t} + \phi(q)\tilde{q}}{\frac{1}{\varepsilon_n} - \frac{\theta_n}{\zeta_n}}, \quad \tilde{h}_n = -\theta_n \left(\frac{\tilde{t} + \phi(q)\tilde{q}}{\frac{1}{\varepsilon_n} - \frac{\theta_n}{\zeta_n}} \right).$$

From eq. (C9) follow the compensated elasticities of labour supply and demand for polluting goods with respect to the tax instruments:

$$\varepsilon_{lt} \equiv \frac{\partial l_n^c}{\partial t} \frac{1-t}{l_n} = -\frac{1}{\frac{1}{\varepsilon_n} - \frac{\theta_n}{\zeta_n}}, \quad \varepsilon_{lq} \equiv \frac{\partial l_n^c}{\partial q} \frac{1+q}{l_n} = -\frac{\phi(q)}{\frac{1}{\varepsilon_n} - \frac{\theta_n}{\zeta_n}}. \tag{C10}$$

$$\varepsilon_{bt} \equiv \frac{\partial b_n^c}{\partial t} \frac{1-t}{b_n} = -\frac{\delta^c \theta_n}{\left(\delta^c \gamma_b + \delta^b \gamma_c \right) \left(\frac{1}{\varepsilon_n} - \frac{\theta_n}{\zeta_n} \right)}, \quad \varepsilon_{bq} \equiv \frac{\partial b_n^c}{\partial q} \frac{1+q}{b_n} = -\frac{\delta^c \theta_n \phi(q)}{\left(\delta^c \gamma_b + \delta^b \gamma_c \right) \left(\frac{1}{\varepsilon_n} - \frac{\theta_n}{\zeta_n} \right)} - \frac{\gamma_c}{\delta^c \gamma_b + \delta^b \gamma_c}. \tag{C11}$$

Consequently, we derive $\varepsilon_{lq} = \phi(q)\varepsilon_{lt}$ and $\phi(q)\varepsilon_{bt} - \varepsilon_{bq} = \gamma_c/(\delta^c \gamma_b + \delta^b \gamma_c) > 0$. Moreover, in the absence of income effects, we have $\zeta_n = \infty$, so that $\varepsilon_{lt} = -\varepsilon_n$, $\varepsilon_{lq} = -\phi(q)\varepsilon_n$.

C.3. Solving for the policy optimum with Gorman polar preferences

Substituting eq. (C3) for Gorman polar form preferences into eq. (A11) yields:

$$\int_0^\infty \left[(1-g_n^*) \left(\frac{\alpha(q)}{(1+q)} + \frac{\phi(q)}{(1+q)}((1-t)nl_n + s) \right) + \frac{t}{(1-t)} \frac{(1-t)}{(1+q)} nl_n\varepsilon_{lq} + \frac{(q-q^P)}{(1+q)} \frac{b_n}{nl_n} nl_n\varepsilon_{bq} \right] f(n)\mathrm{d}n = 0. \tag{C17}$$

Since $\alpha(q)/(1+q) + \phi(q)s/(1+q)$ does not depend on n we have, using the first-order condition for s in eq. (A7):

$$\int_0^\infty \left[(1-g_n^*) \left(\frac{\alpha(q) + \phi(q)s}{(1+q)} \right) \right] f(n)\mathrm{d}n = 0, \tag{C18}$$

Hence, we can write eq. (C18) as:

$$\int_0^\infty \left[(1-g_n^*)nl_n + \frac{t}{(1-t)} \frac{\varepsilon_{lq}}{\phi(q)} nl_n + \frac{(q-q^P)}{(1+q)} \frac{\beta_n\varepsilon_{bq}}{\phi(q)} nl_n \right] f(n)\mathrm{d}n = 0. \tag{C19}$$

Subtract (A10) from eq. (C19) to find:

$$\int_0^\infty \left[\frac{t}{(1-t)} \left(\frac{\varepsilon_{lq}}{\phi(q)} - \varepsilon_{lt} \right) nl_n + \frac{(q-q^P)}{(1+q)} \beta_n \left(\frac{\varepsilon_{bq}}{\phi(q)} - \varepsilon_{bt} \right) nl_n \right] f(n)\mathrm{d}n = 0. \tag{C20}$$

For the general class of weakly separable utility functions with Gorman polar preferences we have shown in eq. (C10) that $\varepsilon_{lq}/\phi(q) = \varepsilon_{lt}$. Moreover, we derived in eq. (C11) that $\phi(q)\varepsilon_{bt} > \varepsilon_{bq}$. Consequently, the first term in brackets is zero and the second term is negative. Hence, optimal policy entails Pigouvian corrective taxes: $q = q^P$. Substitution of $q = q^P$ in Proposition 1 yields Proposition 2.

Appendix D. Proof of Corollary 1

If utility is $u(h(c_n, b_n), 1 - l_n) = h(c_n, b_n) - l_n^{1+1/\vartheta}/(1 + 1/\vartheta)$ and $h(c_n, b_n)$ is of the Gorman (1961) polar form, private marginal utility of income η_n is constant. Consequently, labour supply is given by $l_n = [(1-t)n/p(q)]^\vartheta$, the uncompensated wage elasticity of labour supply with respect to the tax rate equals ϑ, and income effects in labour supply are absent. Using quasi-linearity of the utility function (i.e. $u_{lh} = u_{hh} = 0$), the compensated wage elasticity equals the Frisch elasticity $\varepsilon_n = (u_{ll}l_n/u_l)^{-1} = \vartheta$ and the income-effect term is $\zeta_n = \infty$. Thus, eq. (C10) gives $\varepsilon_{lt} = -\vartheta$ and $\varepsilon_{lq} = -\vartheta\phi(q)$. The optimal income tax from eq. (14) thus boils down to $t/(1-t) = \xi_l/\vartheta$. The Diamond social welfare weights are equal to the standard social welfare weights if income effects in labour supply are absent and pollution taxes are Pigouvian (i.e. $\partial l_n/\partial s = 0$ and $q = q^P$), so that $g_n^* = g_n = \omega_n\eta_n/\lambda = \omega_n p(q)/\lambda$. Substitution of this in eq. (15) yields $p(q) \int_0^\infty \omega_n f(n)\mathrm{d}n = p(q) = \lambda$, since we have normalised the average Pareto weights to 1. If environmental damages are independent from ability, $D_b(b, n) = D'(b)$, eq. (13) boils down to $q = q^P \equiv \int_0^\infty g_n D_b(b, n)/\eta_n f(n)\mathrm{d}n = (D'(b)/\lambda) \int_0^\infty \omega_n f(n)\mathrm{d}n = D'(b)/p(q)$. Note that if the government is utilitarian, i.e. $\omega_n = 1$ for all n, then we have that the social welfare weights are all constant and equal to: $g_n^* = p(q)/\lambda$. Hence, the Feldstein characteristic of labour income is zero ($\xi_l = 0$) and optimal income taxes are zero, too ($t = 0$).

Appendix E. Proof of Proposition 3

We consider the following marginal tax reform. We raise the pollution tax q and the transfer s, and we lower the income tax t so that net income of each household remains constant. We thus consider a tax reform so that $\mathrm{d}s = nl_n\mathrm{d}t + b_n\mathrm{d}q$. Substituting demand for polluting goods in eq. (12) gives:

$$\mathrm{d}s = \left(\frac{\phi(q)(1-t)}{1+q} + \frac{\mathrm{d}t}{\mathrm{d}q} \right) nl_n\mathrm{d}q + \left(\frac{\alpha(q) + \phi(q)s}{1+q} \right) \mathrm{d}q. \tag{E1}$$

Suppose, then, that we implement the following tax reform, which satisfies eq. (E1):

$$\mathrm{d}s = \left(\frac{\alpha(q) + \phi(q)s}{1+q} \right) \mathrm{d}q \quad \text{and} \quad \mathrm{d}t = -\frac{\phi(q)(1-t)}{1+q} \mathrm{d}q. \tag{E2}$$

This tax reform is feasible for all individuals n, since the marginal propensity to consume polluting goods out of net labour income $\phi(q)$ and the intercept of the Engel curve $\alpha(q)$ are the same for everyone. This tax reform leaves non-environmental utility $u(\cdot)$ and net incomes unchanged. Hence, all behavioural responses are compensated changes and there are no income effects. The only effect of the tax reform on private utility is via the change in environmental welfare.

Totally differentiating indirect utility yields (using Roy's identity): $dv_n = -\eta n l_n dt - \eta b_n dq + \eta ds - D_b db = -D_b db$. Totally differentiate total demand for polluting goods to find:

$$db = \int_0^\infty db_n f(n)dn = \int_0^\infty \left(\frac{\partial b_n}{\partial t} + \frac{\partial b_n}{\partial q} \frac{dq}{dt} \right) f(n)dn. \tag{E3}$$

What is the effect of such a tax reform on social welfare? There are two effects: first environmental welfare changes (non-environmental welfare is unaffected) and the government budget changes when the tax system is not optimised. We can use the compensated elasticities here, since non-environmental utility does not change in the reform. The effect on environmental welfare is the sum of the utility changes over the entire population − deflated by the marginal utility of public resources λ:

$$\frac{d\Lambda}{\lambda} = \frac{1}{\lambda} \int_0^\infty \omega_n dv_n f(n)dn = -\int_0^\infty \omega_n \frac{D_b(\cdot)}{\lambda} db f(n)dn = -\int_0^\infty g_n \frac{D_b(\cdot)}{\eta_n} f(n)dn \int_0^\infty \left(\frac{\partial b_n}{\partial t} + \frac{\partial b_n}{\partial q} \frac{dq}{dt} \right) f(n)dn. \tag{E4}$$

Next, substitute $q^P \equiv \int_0^\infty g_n \frac{D_b(b, n)}{\eta_n} f(n)dn$ to find:

$$\frac{1}{\lambda} \frac{d\Lambda}{dt} = -q^P \int_0^\infty \left(\frac{b_n}{1-t} \frac{\partial b_n}{\partial t} \frac{1-t}{b_n} + \frac{b_n}{1+q} \frac{\partial b_n}{\partial q} \frac{1+q}{b_n} \frac{dq}{dt} \right) f(n)dn. \tag{E5}$$

Then, employ $dq/dt = -(1+q)/(\phi(q)(1-t))$ to find the welfare effect of a change in environmental quality:

$$\frac{1}{\lambda} \frac{d\Lambda}{dt} = -\frac{q^P}{1+q} \int_0^\infty \beta_n n l_n \left(\varepsilon_{bt} - \frac{\varepsilon_{bq}}{\phi(q)} \right) f(n)dn. \tag{E6}$$

If R denotes total tax revenues, the effects on the government budget are found by totally differentiating the government budget constraint:

$$\frac{dR}{dt} = \int_0^\infty \left[n l_n + t n \frac{\partial l_n}{\partial t} + t n \frac{\partial l_n}{\partial q} \frac{dq}{dt} + b_n \frac{dq}{dt} + q \frac{\partial b_n}{\partial q} \frac{dq}{dt} + q \frac{\partial b_n}{\partial t} - \frac{ds}{dt} \right] f(n)dn. \tag{E7}$$

Next, substitute the tax reform $dq/dt = -(1+q)/(\phi(q)(1-t))$ and $ds/dt = -(\alpha(q) + \phi(q)s)/(\phi(q)(1-t))$ and rearrange to get:

$$\frac{dR}{dt} = \int_0^\infty n l_n \left[\frac{t}{(1-t)} \left(\varepsilon_{lt} - \frac{\varepsilon_{lq}}{\phi(q)} \right) + \frac{q}{(1+q)} \beta_n \left(\varepsilon_{bt} - \frac{\varepsilon_{bq}}{\phi(q)} \right) \right] b_n f(n)dn. \tag{E8}$$

The total welfare effect of the tax reform is then equal to eq. (E8) plus eq. (E6). Rewriting while using $dq/dt = -(1+q)/(\phi(q)(1-t))$ gives:

$$\frac{1}{\lambda} \frac{d\Lambda}{dq} + \frac{dR}{dq} = -\int_0^\infty \left[\frac{t}{(1-t)} \frac{\phi(q)}{\beta_n} \left(\varepsilon_{lt} - \frac{\varepsilon_{lq}}{\phi(q)} \right) + \frac{(q - q^P)}{(1+q)} \left(\phi(q)\varepsilon_{bt} - \varepsilon_{bq} \right) \right] f(n)dn. \tag{E9}$$

Clearly, if preferences are weakly separable and belong to the Gorman polar form, we have from eq. (C10) that $\varepsilon_{lt} = \varepsilon_{lq}/\phi(q)$, hence the total effect of the tax reform on social welfare equals:

$$\frac{1}{\lambda}\frac{\mathrm{d}\Lambda}{\mathrm{d}q} + \frac{\mathrm{d}R}{\mathrm{d}q} = \int_0^\infty b_n \left[\frac{(q^P - q)}{(1+q)}(\phi(q)\varepsilon_{bt} - \varepsilon_{bq})\right]\mathrm{f}(n)\mathrm{d}n. \tag{E10}$$

We derived in eq. (C11) that $\phi(q)\varepsilon_{bt} > \varepsilon_{bq}$. This corresponds to eq. (17) of Proposition 3.

Appendix F. Further calibration details

Details of our calibration are summarised in Table F.1. We assume a (un)compensated wage elasticity of labour supply equal to $\vartheta = 0.25$, which is consistent with the estimates reported in Blundell and MaCurdy (1999) and Meghir and Phillips (2010). We also perform a sensitivity analysis with $\vartheta = 0.35$.

We assume exponentially declining Pareto weights: $\omega_n = \psi_0 n^{-\psi}$, with $\psi_0 = 27.68 \times 10^6$ such that $E[\omega_n] = 1$. ψ corresponds to the coefficient of relative inequality aversion (cf. Atkinson, 1970).[43] We assume a rather redistributive government by setting $\psi = 4$. We also consider a sensitivity exercise with a lower degree of inequality aversion of $\psi = 2$. The public revenue requirement is $R = 50$, but we also consider a lower requirement of $R = 30$ as a sensitivity exercise.

We use the lognormal (or Galton) distribution to describe the distribution of earning abilities of individuals. Hence, $\ln(n)$ is normally distributed with mean $\ln(\overline{n})$ and standard deviation $\varpi > 0$.[44] We set $\varpi = 0.4$ (cf., Stern, 1976). Mean labour productivity is calibrated so that $E[n] = 100$ without government intervention: $\overline{n} = 100/\exp(0.5\varpi^2) = 92.3$. This yields $E[\ln(n)] = 4.525$. As a sensitivity check, we also examine the effects of a higher dispersion of incomes by setting $\varpi = 0.8$.

Finally, let $D(b, n) = \Xi_0 b(E[n]/n)^{\Xi_1}$ be the function describing pollution damages, where the coefficient Ξ_1 indicates the extent to which pollution damages vary with the ability of individuals. For our benchmark we suppose $\Xi_0 = 0.1$ and that the pollution weights are the same for all individuals, i.e. $\Xi_1 = 0$. As sensitivity exercises, we also explore the case where pollution damages are double the size, i.e. $\Xi_0 = 0.1$, and where the incidence of pollution falls disproportionally on the poorest individuals, i.e. $\Xi_1 = 1$. The Pigouvian pollution tax in eq. (7) or eq. (13) boils down to $q^P = \Xi_0 \int_0^\infty \omega_n (E[n]/n)^{\Xi_1} \mathrm{f}(n)\mathrm{d}n/\lambda$, which simplifies to $q^P = \Xi_0/\lambda$ if $\Xi_1 = 0$.

Table F.1
Parameters calibration (robustness checks in parentheses)

Parameters PIGLOG utility	$\chi_0 = 0.315, \chi_1 = 0.05$
Parameters LES utility	$a = 0.0146, \overline{b} = 7.098$
Wage elasticity of labour supply	$\vartheta = 0.25$ $(\vartheta = 0.35)$
Parameters social welfare weights	$\psi = 4$ $(\psi = 2)$, $\psi_0 = 27.68 \times 10^6$
Pre-tax income distribution	$\varpi = 0.4$ $(\varpi = 0.8)$, $\overline{n} = 92.3$
Public revenue requirement	$R = 50$ $(R = 30)$
Marginal pollution damages	$\Xi_0 = 0.1$ $(\Xi_0 = 0.2)$
Rate at which pollution damages decline with ability	$\Xi_1 = 0$ $(\Xi_1 = 1)$

Appendix G. Proof optimal taxation with production externalities

The derivatives of indirect utility to the tax instruments are: $\partial v_n/\partial s = \eta_n, \partial v_n/\partial t = -\eta_n A(b)nl_n, \partial v_n/\partial q = -\eta_n b_n$, and $\partial v_n/\partial b = \eta_n(1-t)A'(b)nl_n$. The Lagrangian Λ for the maximisation of social welfare is modified to:

$$\max_{\{t,q,s,b\}} \Lambda \equiv \int_0^\infty \omega_n \mathrm{v}(t, q, s, b, n) + \lambda[tA(b)nl_n + qb_n - s - R]\mathrm{f}(n)\mathrm{d}n + \mu\left[b - \int_0^\infty b_n \mathrm{f}(n)\mathrm{d}n\right]. \tag{G1}$$

The following necessary first-order conditions have to hold:

$$\frac{\partial \Lambda}{\partial s} = \int_0^\infty \left[\omega_n \eta_n - \lambda + \lambda tA(b)n\frac{\partial l_n}{\partial s} + (\lambda q - \mu)\frac{\partial b_n}{\partial s}\right]\mathrm{f}(n)\mathrm{d}n = 0, \tag{G2}$$

---

[43] A more conventional social welfare function weights by utility rather than ability, so that $\omega_n = v_n^{-\psi}$. Given that the private marginal utility of income is constant, assuming exogenous social welfare weights is without loss of generality.

[44] More realistic would be to append a Pareto tail to the log-normal distribution to capture the upper end of the earnings distribution (e.g. Atkinson et al., 2011). However, we abstract from this extension. Since we analyse linear income taxes, doing so would not bring any essential insights for our purposes.

$$\frac{\partial \Lambda}{\partial t} = \int_0^\infty \left[-\omega_n \eta_n A(b) n l_n + \lambda A(b) n l_n + \lambda t A(b) n \frac{\partial l_n}{\partial t} + (\lambda q - \mu) \frac{\partial b_n}{\partial t} \right] f(n) dn = 0, \tag{G3}$$

$$\frac{\partial \Lambda}{\partial q} = \int_0^\infty \left[-\omega_n \eta_n b_n + \lambda b_n + \lambda t A(b) n \frac{\partial l_n}{\partial q} + (\lambda q - \mu) \frac{\partial b_n}{\partial q} \right] f(n) dn = 0, \tag{G4}$$

$$\frac{\partial \Lambda}{\partial b} = \int_0^\infty \left[\omega_n \eta_n (1-t) A'(b) n l_n + \lambda t A'(b) n l_n + \mu + \lambda t A(b) n \frac{\partial l_n}{\partial b} + (\lambda q - \mu) \frac{\partial b_n}{\partial b} \right] f(n) dn = 0, \tag{F5}$$

where we have used Roy's identity in each line to substitute for the derivatives of indirect utility.

Using $\frac{\partial l_n}{\partial b} = \frac{\partial l_n}{\partial w} A'(b) = -\frac{\partial l_n}{\partial t} \frac{1-t}{w} A'(b)$, $\frac{\partial b_n}{\partial b} = \frac{\partial b_n}{\partial w} A'(b) = -\frac{\partial b_n}{\partial t} \frac{1-t}{w} A'(b)$, where the second steps follow from $\frac{\partial l_n}{\partial w} = -\frac{\partial l_n}{\partial t} \frac{1-t}{w}$, and $\frac{\partial b_n}{\partial w} = -\frac{\partial b_n}{\partial t} \frac{1-t}{w}$, the first-order condition for optimal pollution b in eq. (G5), can be rewritten as:

$$\int_0^\infty \left[\omega_n \eta_n n l_n + \lambda \frac{t}{1-t} n l_n + \frac{\mu}{(1-t) A'(b)} - \lambda t n \frac{\partial l_n}{\partial t} - \frac{(\lambda q - \mu)}{w} \frac{\partial b_n}{\partial t} \right] f(n) dn = 0. \tag{G6}$$

Next, substitute the first-order condition for t in eq. (G3) to derive that the Pigouvian pollution tax is given by the sum of marginal pollution damages in production:

$$q^P \equiv \frac{\mu}{\lambda} = -A'(b) \int_0^\infty n l_n f(n) dn = -\frac{A'(b)}{A(b)} Y. \tag{G7}$$

We can simplify the first-order conditions for s, t and q, i.e. eqs. (G2), (G3) and (G4) following the same steps as in Appendix A, to replicate the optimal tax formulae of Proposition 1 with a social marginal utility of income of $g_n^* \equiv g_n + t A(b) n \frac{\partial l_n}{\partial s} + (q - q^P) \frac{\partial b_n}{\partial s}$ and an expenditure share of $\beta_n \equiv \frac{(1+q) b_n}{(1-t) A(b) n l_n}$.

References

Abrell, J., Rausch, S., Schwarz, G.A., 2017. How robust is the uniform emissions pricing rule to social equity concerns? J. Environ. Econ. Manag. (forthcoming).
Adler, M., Anthoff, D., Bosetti, V., Garner, G., Keller, K., Treich, N., 2016. Priority for the Worse off and the Social Cost of Carbon. CESifo, Munich. Working Paper No. 6032.
Advani, A., Johnson, P., Leicester, A., Stoye, G., 2013. Household Energy Use in Britain: A Distributional Analysis. Institute for Fiscal Studies, London. IFS Report, No. R85.
Aigner, R., 2014. Environmental taxation and redistribution concerns. FinanzArchiv 70 (29), 249–277.
Allen, F., 1982. Optimal linear income taxation with general equilibrium effects on wages. J. Publ. Econ. 17 (2), 135–143.
Anthoff, D., Hepburn, C., Tol, R.S.J., 2009. Equity weighting and the marginal damage costs of climate change. Ecol. Econ. 68 (3), 836–849.
Anthoff, D., Tol, R.S.J., 2010. On international equity weights and national decision making on climate. J. Environ. Econ. Manag. 60 (1), 14–20.
Atkinson, A.B., 1970. On the measurement of economic inequality. J. Econ. Theor. 2 (3), 244–263.
Atkinson, A.B., Piketty, T., Saez, E., 2011. Top incomes in the long run of history. J. Econ. Lit. 49 (1), 3–71.
Atkinson, A.B., Stiglitz, J.E., 1976. The design of tax structure: direct versus indirect taxation. J. Publ. Econ. 6 (1–2), 55–75.
Aubert, D., Chiroleu-Assouline, M., 2017. Environmental Tax Reform and Income Distribution with Imperfect Heterogeneous Labour Markets. CESifo, Munich. Working Paper No. 6498.
Banks, J. R. Blundell, Lewbel, A., 1997. Quadratic Engel curves and consumer demand. Rev. Econ. Stat. 79 (4), 527–539.
Baumgärtner, S., Drupp, M.A., Meya, J.N., Munz, J.M., Quaas, M.F., 2017. Income inequality and willingness to pay for environmental public goods. J. Environ. Econ. Manag. 85, 35–61.
Berndt, E.R., Wood, D.O., 1975. Technology, prices, and the derived demand for energy. Rev. Econ. Stat. 57 (3), 259–268.
Blundell, R., MaCurdy, T., 1999. Labour supply: a review of alternative approaches, Chapter 27. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labour Economics, vol. 3. Elsevier, New York, pp. 1559–1695.
Boadway, R., Tremblay, J.-F., 2008. Pigouvian taxation in a Ramsey world. Asia-Pacific J. Account. Finance 15 (3), 183–204.
Bohi, D.R., Zimmerman, M.B., 1984. An update on econometric studies of energy demand behavior. Annu. Rev. Energy 9 (1), 105–154.
Boppart, T., 2014. Structural change and the Kaldor facts in a growth model with relative price effects and non-Gorman preferences. Econometrica 82 (6), 2167–2196.
Bovenberg, A.L., 1998. Environmental taxes and the double dividend. Empirica 25 (1), 15–35.
Bovenberg, A.L., 1999. Green tax reforms and the double dividend: an updated reader's guide. Int. Tax Publ. Finance 6 (3), 421–443.
Bovenberg, A.L., De Mooij, R.A., 1994. Environmental levies and distortionary taxation. Am. Econ. Rev. 84 (4), 1085–1089.
Bovenberg, A.L., van der Ploeg, F., 1994. Environmental policy, public finance and the labour market in a second-best world. J. Publ. Econ. 55 (3), 349–390.
Bruce, J.P., Yi, H.-S., Haites, E.F., 1995. Climate Change 1995: Economic and Social Dimensions of Climate Change. IPCC and Cambridge University Press, Cambridge, U.K.

Burtraw, D., Sweeney, R., Walls, M., 2009. The incidence of U.S. climate policy: alternative uses of revenues from a cap-and-trade auction. Natl. Tax J. 62 (3), 497—518.

Chiroleu-Assouline, M., Fodha, M., 2014. From regressive pollution taxes to progressive environmental tax reforms. Eur. Econ. Rev. 69, 126—142.

Comin, D., Lashkari, D., Mestieri, M., 2015. Structural Change with Long-Run Income and Price Effects. mimeo, Dartmouth University.

Corlett, W.J., Hague, D.C., 1953. Complementarity and the excess burden of taxation. Rev. Econ. Stud. 21 (1), 21—30.

Crainger, C.A., Kohlstad, C.D., 2010. Who pays a price on carbon? Environ. Resour. Econ. 46 (3), 359—376.

Crawford, I., Keen, M., Smith, S., 2010. Value added tax and excises, Chapter 4. In: Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G., Poterba, J.M. (Eds.), The Mirrlees Review. Dimensions of Tax Design. Oxford University Press, Oxford, pp. 275—362.

Cremer, H., Gahvari, F., 2001. Second-best taxation of emissions and polluting goods. J. Publ. Econ. 80 (2), 169—197.

Cremer, H., Gahvari, F., Ladoux, N., 1998. Externalities and optimal taxation. J. Publ. Econ. 70 (3), 343—364.

Cremer, H., Gahvari, F., Ladoux, N., 2003. Environmental taxes with heterogeneous consumers: an application to energy consumption in France. J. Publ. Econ. 87 (12), 2791—2815.

Deaton, A.S., 1979. Optimally uniform commodity taxes. Ec. Letters. 2 (4), 357—361.

Deaton, A.S., Muellbauer, J., 1980. An almost ideal demand system. Am. Econ. Rev. 70 (3), 312—326.

De Palma, A., Lindsey, R., 2004. Congestion pricing with heterogeneous travelers: a general equilibrium welfare analysis. Network. Spatial Econ. 4 (2), 135—160.

Diamond, P.A., 1973. Consumption externalities and imperfect corrective pricing. Bell J. Econ. 4 (2), 526—538.

Diamond, P.A., 1975. A many-person Ramsey rule. J. Publ. Econ. 4 (4), 335—342.

Diamond, P.A., Mirrlees, J., 1971a. Optimal taxation and public production, I: production efficiency. Am. Econ. Rev. 61 (1), 8—27.

Diamond, P.A., Mirrlees, J., 1971b. Optimal taxation and public production II: tax rules. Am. Econ. Rev. 61 (3), 261—278.

Feldstein, M.S., 1972. Distributional equity and the optimal structure of public prices. Am. Econ. Rev. 62 (1), 32—36.

Flues, F., Thomas, A., 2015. The Distributional Effects of Energy Taxation, OECD Taxation. OECD Publishing, Paris. Working Papers, No. 23.

Fullerton, D., Metcalf, G.E., 2001. Environmental controls, scarcity rents, and pre-existing distortions. J. Publ. Econ. 80 (2), 249—267.

Fullerton, D., Heutel, G., 2007. The general equilibrium incidence of environmental taxes. J. Publ. Econ. 91 (3—4), 571—591.

Fullerton, D., Monty, H., 2013. Can pollution tax rebates protect low-wage earners? J. Environ. Econ. Manag. 66 (3), 539—553.

Fullerton, D., Heutel, G., Metcalf, G.E., 2012. Does the indexing of government transfers make carbon pricing progressive? Am. J. Agric. Econ. 94 (2), 347—353.

Gahvari, F., 2014. Second-best Pigouvian taxation: a clarification. Environ. Resour. Econ. 59 (4), 525—535.

Goulder, L.H., 1995. Environmental taxation and the double dividend: a reader's guide. Int. Tax Publ. Finance 2 (2), 157—183.

Gorman, W.M., 1961. On a class of preference fields. Metroeconomica 13 (2), 53—56.

Hassett, K.A., Mathur, A., Metcalf, G.E., 2009. The incidence of a U.S. carbon tax: a lifetime and regional analysis. Energy J. 30 (2), 155—178.

Jacobs, B., 2018. The marginal cost of public funds is one at the optimal tax system. Int. Tax Publ. Finance 25 (4), 883—912.

Jacobs, B., Boadway, R., 2014. Optimal linear commodity taxation under optimal non-linear income taxation. J. Publ. Econ. 117 (1), 201—210.

Jacobs, B., de Mooij, R.A., 2011. Pigou Meets Mirrlees: on the Irrelevance of Tax Distortions for the Second-Best Pigouvian Tax. CESifo, Munich. Working Paper No. 3342.

Jacobs, B., de Mooij, R.A., 2015. Pigou meets Mirrlees: on the irrelevance of tax distortions for the second-best Pigouvian tax. J. Environ. Econ. Manag. 71 (May), 90—108.

Kaplow, L., 2012. Optimal control of externalities in the presence of income taxation. Int. Econ. Rev. 53 (2), 487—509.

Klenert, D., Mattauch, L., 2016. How to make a carbon tax reform progressive: the role of subsistence consumption. Econ. Lett. 138, 100—103.

Klenert, D., Schwerhoff, G., Edenhofer, O., Mattauch, L., 2018a. Carbon taxation, inequality and Engel's law: the double dividend of redistribution. Environ. Resour. Econ. 71 (3), 605—624.

Klenert, D., Mattauch, L., Combet, E., Edenhofer, O., Hepburn, C., Rafaty, R., Stern, N., 2018b. Making carbon pricing work for citizens. Nat. Clim. Change 8, 669—677.

Levison, A., 2016. Energy Standards Are More Regressive than Taxes: Theory and Evidence. NBER, Cambridge, Mass. Working Paper No. 22956.

Lewbel, A., Pendakur, K., 2009. Tricks with Hicks: the EASI demand system. Am. Econ. Rev. 99 (3), 827—863.

Magnus, J.R., 1979. Substitution between energy and nonenergy inputs in The Netherlands, 1950-1976. Int. Econ. Rev. 20 (2), 465—484.

Mayeres, I., Proost, S., 2001. Marginal tax reform, externalities and income distribution. J. Publ. Econ. 79 (2), 343—363.

Meghir, C., Browning, M., 1991. The effects of male and female supply on commodity demands. Econometrica 59 (4), 925—951.

Meghir, C., Phillips, D., 2010. Labour supply and taxes, Chapter 3. In: Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G., Poterba, J.M. (Eds.), The Mirrlees Review. Dimensions of Tax Design. Oxford University Press, Oxford, pp. 202—274.

Micheletto, L., 2008. Redistribution and optimal mixed taxation in the presence of consumption externalities. J. Publ. Econ. 92 (10—11), 2262—2274.

Miranda, M.L., Edwards, S.E., Keating, M.H., Paul, C.J., 2011. Making the environmental justice grade: the relative burden of air pollution exposure in the United States. Int. J. Environ. Res. Public Health 8 (6), 1755—1771.

Mirrlees, J.A., 1971. An exploration in the theory of optimum income taxation. Rev. Econ. Stud. 38 (2), 175—208.

Mirrlees, J.A., 1976. Optimal tax theory: a synthesis. J. Publ. Econ. 6 (4), 327—358.

Mirrlees, J.A., 1978. Social benefit-cost analysis and the distribution of income. World Dev. 6 (2), 131—138.

Muellbauer, J., 1976. Community preferences and the representative consumer. Econometrica 44 (5), 525—543.

Nordhaus, W., 2014. Estimates of the social cost of carbon: concepts and results from the DICE-2013R model and alternative approaches. J. Assoc. Environ. Resour. Econ. 1 (1/2), 273—312.

Papageorgiou, C., Saam, M., Schulte, P., 2017. Substitution between clean and dirty energy inputs: a macroeconomic perspective. Rev. Econ. Stat. 99 (2), 281—290.

Parry, I., Mylonas, V., Vernon, N., 2017. Reforming Energy Policy in India: Assessing the Options. IMF, Washington, D.C. IMF Working Paper WP/17/103.

Pindyck, R.S., 1979. Interfuel substitution and the industrial demand for energy: an international comparison. Rev. Econ. Stat. 61 (2), 169—179.

Pirttilä, J., Schöb, R., 1999. Redistribution and internalization: the many-person Ramsey tax rule revisited. Publ. Finance Rev. 27 (5), 541—560.

Pirttilä, J., Suoniemi, I., 2014. Public provision, commodity demand, and hours of work: an empirical analysis. Scand. J. Econ. 116 (4), 1044—1067.

Pizer, W.A., Sexton, S., 2017. Distributional Impacts of Energy Taxes. NBER, Cambridge, Mass. Working Paper No. 23318.

Rausch, S., Metcalf, G.E., Reilly, J.M., 2011. Distributional aspects of carbon pricing: a general equilibrium approach with micro-data for households. Energy Econ. 33 (S1), S20—S33.

Saez, E., 2002. The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. J. Publ. Econ. 83 (2), 217—230.

Saez, E., Stantcheva, S., 2016. Generalized social marginal welfare weights for optimal tax theory. Am. Econ. Rev. 106 (1), 24—45.

Sandmo, A., 1975. Optimal taxation in the presence of externalities. Scand. J. Econ. 77 (1), 86—98.

Sandmo, A., 2000. The Public Economics of the Environment. The Lindahl Lectures 1996. Oxford University Press, Oxford.

Scheuer, F., Werning, I., 2016. Mirrlees Meets Diamond-Mirrlees: Simplifying Non-linear Income Taxation. NBER, Cambridge, Mass. NBER Working Paper No. 22076.

Stern, N.H., 1976. On the specification of models of optimal income taxation. J. Publ. Econ. 6 (1—2), 123—162.

Stiglitz, J.E., 1982. Self-selection and Pareto efficient taxation. J. Publ. Econ. 17 (2), 213—240.

Tovar Reaños, M.A., Wölfing, N.M., 2018. Household energy prices and inequality: evidence from German microdata based on the EASI demand system. Energy Econ. 70, 84—97.

Tsyvinski, A., Sachs, D., Werquin, N., 2016. Non-linear Tax Incidence and Optimal Taxation in General Equilibrium. NBER, Cambridge, Mass. Working Paper No. 22646.

West, S.E., Williams III, R.C., 2004. Estimates from a consumer demand system: implications for the incidence of environmental taxes. J. Environ. Econ. Manag. 47 (3), 535–558.

Williams III, R.C., 2006. An Estimate of the Second-Best Optimal Gasoline Tax, Considering Both Efficiency and Equity. mimeo, University of Maryland.