DEPARTMENT OF ECONOMICS
OxCarre (Oxford Centre for the Analysis of Resource Rich Economies)

Manor Road Building, Manor Road, Oxford OX1 3UQ
Tel: +44(0)1865 281281 Fax: +44(0)1865 281163
reception@economics.ox.ac.uk www.economics.ox.ac.uk

UNIVERSITY OF OXFORD

OxCarre Research Paper 49

–

PRECAUTIONARY CLIMATE CHANGE POLICIES AND OPTIMAL REDISTRIBUTION

Bas Jacobs
Erasmus University Rotterdam

&

Frederick van der Ploeg
OxCarre,
University of Oxford

Direct tel: +44(0) 1865 281281 E-mail: celia.kingham@economics.ox.ac.uk

PRECAUTIONARY CLIMATE CHANGE POLICIES AND OPTIMAL REDISTRIBUTION*

Bas Jacobs, Erasmus University Rotterdam** and Frederick van der Ploeg, University of Oxford***

Abstract

We analyse optimal carbon taxes, optimal redistribution within and between non-overlapping generations, and optimal spending levels on climate abatement and adaptation. A positive probability of unexpected large increases in CO2 emissions results in a lower discount rate for global warming damages. More prudent governments set higher carbon taxes and spend more on abatement and sacrifice intra-generational for inter-generational redistribution. As long as households spend a constant fraction of their income on polluting goods, the carbon tax is not used for redistribution and is set at the modified Pigouvian rate, which is higher than the Pigouvian rate if governments are prudent. However, the carbon tax is set below the modified Pigouvian rate if poor households spend relatively more on polluting goods than rich households (Stone-Geary preferences). Policy simulations give insights into the effects of changes in the probability of climate disaster, degrees of intra- and inter-generational inequality aversion, ease of substitution between clean and dirty goods, elasticity of labour supply, productivity of abatement and adaptation, population growth and economic growth on the rates of discount, inequality, global warming and social welfare.

Keywords: global warming, intra-generational and inter-generational redistribution, equally-distributed-equivalent utility, social discount rate, prudence, carbon tax, income tax, CO2 abatement, climate adaptation, non-homothetic preferences

JEL codes: H21, H23, Q54

26 July 2010

* We are grateful to the very helpful and insightful comments of our discussant Michael Hoel and those of the other participants in an earlier version presented at the Workshop on Climate Change and Distribution, organised by ESOP and Centre for the Study of Mind in Nature in collaboration with the Stanford Center for Ethics in Society, Oslo, 22-23 June 2010.
** Also affiliated with the Tinbergen Institute, Netspar and CESifo. Adress: Erasmus School of Economics, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands. Phone: +31-10-4081481/91. Email: bjacobs@ese.eur.nl .
*** Also affiliated with the University of Amsterdam, Tinbergen Institute, CESifo and CEPR. Address: OxCarre, Department of Economics, Manor Road Building, Oxford OX1 3UQ, United Kingdom. Phone: +44-1865-281285. Email: rick.vanderploeg@economics.ox.ac.uk .

1. Introduction

Climate change has been coined "the greatest externality the world has ever seen" (Stern, 2007). Taxing fossil fuels to internalise this externality seems the obvious way to attack the problem of global warming. The disappointing outcome of the recent Copenhagen Climate Change Conference suggests that governments are reluctant to levy such a carbon tax to curb CO2 emissions and combat climate change. At the heart of this policy inertia may be income distribution issues, since developing countries have hardly contributed to global warming and yet have to put up with most of the costs. The poorest people on the planet will be hardest hit by a tough and ambitious climate change policy as they are the ones that are still consuming dirty goods and cannot afford as easily to switch to cleaner, more expensive goods. Governments of developed economies also find it politically difficult to impose a steep carbon tax to fight climate change as this hurts the lowest income groups in their countries hardest. Indeed, optimal green policies may conflict with left-wing policies. We therefore believe that intra- as well as inter-generational income distribution issues should be at the heart of any discussion of climate change.

This paper takes a first shot at facing the income distributional consequences of climate change both from an intra- and inter-generational perspective.[1] We do this by analysing the optimal carbon tax and optimal CO2 abatement and climate adaptation policies in conjunction with the optimal redistributive income tax and optimal redistribution between non-overlapping generations. It is crucial to analyse the optimal correction for externalities and the optimal provision of public goods *simultaneously* with the optimal distribution of incomes; policy prescriptions will be seriously misguided if the reasons for introducing tax distortions (income distribution, externalities) are left out of the analysis (Kaplow, 2006; Jacobs and de Mooij, 2010).[2] We abstract from the intertemporal choice for the adoption of alternatives to fossil fuels (wind, solar, etcetera) and the green paradoxes that may result from that (e.g., Sinn, 2008; Hoel, 2008; Grafton et al., 2010; van der Ploeg and Withagen, 2010). We address the intricate trade-offs between inter- and intra-generational income redistribution and climate policies by adopting a normative approach, and leave the explicit modelling of political economy aspects for another occasion.

We suppose that the future economy and its citizens are unrelated to those on the planet today, except that we allow society to pass on productive assets and carbon debt to future generations. Future economic prospects are adversely affected by inherited global warming damages. Given the large uncertainties

---

[1] We do this for the global economy and therefore do not deal with the complicated geographical aspects of climate change and the required negotiations between sovereign states.

[2] Kaplow (2006) follows Kaplow (2004, 2008) by analysing a change in environmental tax policy alongside a benefit-absorbing change in the income-tax schedule. This combined policy leaves the welfare distribution unaffected and does not affect labour supply incentives either, since preferences are assumed homogeneous and weakly separable (Laroque, 2005). Hence, the Pigouvian rule can be used to judge the desirability of pollution taxes.

associated with future global warming damages and the possibility of disastrous losses of welfare associated with tipping points and irreversible climate change, Stern (2007) has argued that it is prudent to appeal to the precautionary or better-safe-than-sorry principle and thus to use a lower discount rate for global warming damages. This has led to an intense academic debate (e.g., Nordhaus, 2007; Weitzman, 2007). Some argue that this calls for a lower discount rate or even, in the spirit of the ethical views of Ramsey (1928), a zero discount rate for evaluating global warming damages.[3] The Stern Report uses a social rate of discount of 1.4 percent rather than a market rate of discount of 6 percent per annum (Stern, 2007). One way of justifying this is by noting that under uncertainty the certainty-equivalent discount rate declines over time to a lower bound (Weitzman, 1998). Taking account of uncertain probabilities of future climate change brings the discount rate implied by the market more in line with the one used by the Stern Report (Weitzman, 2009).[4] Others argue that it is consistent to use a lower discount rate for increasingly scarce environmental capital than for faster growing aggregate consumption (Sterner and Persson, 2008). We rationalise using a relatively low rate to discount the costs of future climate change based on climate uncertainty and convex marginal global warming damages.

Most of the focus in the literature on cautious discounting of global warming damages has been on partial-equilibrium analysis and applies only to marginal climate change policies. The partial equilibrium theory of the optimal discount rate which is used for the cost-benefit analysis of marginal investment projects 'unpacks' the social rate of discount rate into (i) a pure rate of time preference (positive); (ii) a wealth effect consisting of the product of the trend rate of growth of consumption and the coefficient of relative risk aversion or inequality aversion (positive) and a term depending on the variance of output and the degree of risk aversion (negative); (iii) a prudence term which makes policy makers more patient and accumulate more to have a buffer against unexpected extreme negative outcomes (negative); and (iv) an ambiguity term consisting of an ambiguity prudence and a pessimism effect (positive/negative) (Gollier and Gierlinger, 2008).[5] If the elastity of intertemporal substitution (inverse of the coefficient of relative intertemporal inequality aversion) differs from the coefficient of relative intratemporal risk aversion, there will be an extra negative term under (ii) which increases in the variance of output (Traeger, 2008). We abstract from determinant (iv) of the social rate of discount. We also focus on future climate uncertainty,

[3] The Stern Review has been discussed extensively with some arguing out that the social discount rate should be consistent with aggregate preferences as revealed by markets (e.g., Nordhaus, 2007). One way of justifying a lower discount rate for global warming damages than for the private part of welfare is within the context of a capital asset pricing model if the return on investment projects to combat climate change (possibly induced by a higher carbon tax) are correlated more with the green than with the private part of welfare (Weitzman, 2007).

[4] With fat tails there is even the possibility that the integral of discounted global warming damages does not converge, which results in the so-called 'dismal theorem' (Weitzman, 2009).

[5] Increasing intra-generational inequality or increasing inequality across countries over time also leads to a lower discount rate and a lower cost of carbon (Emmerling, 2010).

so abstract from uncertainty about the future economy. It is then prudent to discount future social marginal damages of global warming less heavily than suggested by the pure rate of time preference.

In contrast to the partial-equilibrium framework for analysing *marginal* climate change policies, we offer a general-equilibrium analysis for analysing *non-marginal* climate change policies. We thus provide a stochastic general-equilibrium model with endogenous labour supply and choice of polluting and non-polluting goods and with prudent social preferences to rationalise cautious discounting of global warming damages, and use this to explore the consequences for optimal climate and redistribution policies. By assigning a small probability to an unexpected large increase in CO2 emissions[6], the relevant rate for discounting future marginal global warming damages will be endogenous and lower than the one used for the future private component of marginal social welfare. The government maximises equally-distributed-equivalent private utility (private welfare corrected for concern about avoiding intra-generational inequality) minus the social damages of global warming. We generalise the analysis of optimal environmental taxation and optimal income taxation by Jacobs and de Mooij (2010)[7] to a dynamic setting. We characterise our optimal policies for the case where preferences are separable in real consumption and labour supply, but the appendix presents our results for the general case with income effects in labour supply and non-homothetic preferences.

With homothetic preferences over clean and dirty goods, the optimal carbon tax follows a modified Pigouvian rule. Due to future environmental uncertainties the modified Pigouvian tax is set above the normal Pigouvian tax for precautionary reasons. However, no corrections are made for distributional concerns. However, with Stone-Geary preferences poor households have a subsistence need to consume dirty commodities. We show that then the optimal carbon tax is driven below the modified Pigouvian level and environmental quality is sacrificed for income redistribution. The Samuelson-rule for optimal spending on climate adaptation equates marginal social benefits – converted into resource units by dividing by the marginal social cost of equality – to the marginal cost of public funds, which equals one. More unequal societies who care about fighting inequality give less priority to climate policy. Intuitively, if there is substantial inequality, resources are more valuable and the social willingness to pay for a cleaner environment diminishes: 'Erst das Fressen, dann die Moral'. Of course, a higher carbon debt leads

---

[6] We abstract from fat tails, the possibility that the integral of discounted global warming damages does not converge, and the resulting 'dismal theorem' (Weitzman, 2009).

[7] This study combines the principles of optimal environmental taxation (Sandmo, 1975) and the optimal redistributive tax system (Mirrlees, 1971). It criticises the double dividend literature, which suggested that greener preferences need not necessarily lead to both an increase in environmental quality and in employment (Bovenberg and de Mooij, 1994; Bovenberg and van der Ploeg, 1994), for their ad hoc second-best assumption of the unavailability of lump-sum taxes. It argues that the appropriate marginal cost of public funds is unity.

to more spending on climate adaptation. The optimal income tax rate will be higher if labour supply is less elastic (the Ramsey motive) and the desire to redistribute income is larger, i.e., when inequality in earning abilities is more pronounced. Future climate uncertainty combined with convex marginal damages of global warming damages induces cautious discounting of global warming damages and thus results in policy makers paying relatively more attention to fighting climate change and pursuing inter-generational distribution than to intra-generational income redistribution. Our analysis thus demonstrates that it is indeed difficult to reconcile left-wing policies (intra-generational redistribution) with green policies (high carbon taxes and CO2 abatement) that promote inter-generational equity.

Section 2 sets up the model and section 3 uses it to derive the cautious social discount rate and the optimal carbon tax, spending on abatement and adaptation, and the optimal linear income tax. Section 4 analyses non-homothetic preferences to ensure that dirty goods are necessities and clean goods are luxuries, which make it harder to pursue an ambitious climate policy as redistribution is pertinent. Section 5 numerically investigates the effects of changes in the probability of climate disasters, the degree of intra-and inter-generational inequality aversion, the ease of substitution of dirty for clean goods, the wage elasticity of labour supply, and the rates of population and productivity growth on the rate that is used to discount global warming damages, climate and redistribution policies, inequality, global warming and social welfare. Section 6 concludes and suggests directions for further research.

2. The Model

The economy lasts two periods. The variables describing the current economy are indicated by lower-case italics, whereas the variables describing the future, say a century from now, are indicated by upper-case italics. Functions are indicated by roman letters. In each period there is a heterogeneous population of individuals who differ in their (exogenous) earnings ability. Today's and future) earnings ability (or labour productivity per hour worked) are denoted by $n \in [0, \infty)$ and $N \in [0, \infty)$, respectively. The probability density function for the distribution of productivities (equal to before-tax wages) across individuals is f(n). Population size is normalised to one, so $\int_0^\infty \mathrm{f}(n)\mathrm{d}n = 1.$ Average productivity of today's individuals is given by $\mathbf{E}[n] = \int_0^\infty n\mathrm{f}(n)\mathrm{d}n,$ where $\mathbf{E}$[.] denotes the expectations operator.

Future individuals are the offspring of today's individuals. We suppose that there is no link between the current and future generations except through the stocks of public assets and carbon debt left to future generations. We thus consider a model with non-overlapping generations and abstract from dynasties or overlapping generations. The distribution of productivities of future individuals is given by the probability density functions F(N). The future economy may differ from the current economy in two respects. First, there may be more individuals in the future than today. So we have $\int_0^\infty \mathrm{F}(N)\mathrm{d}N = 1+\alpha$, where α denotes growth of the population. Second, we suppose $N = (1+\beta)\, n$ and thus that the future economy is on average more productive than today's economy by a factor $1+\beta$, so that

$\mathbf{E}[N] = \int_0^\infty N\mathrm{F}(N)\mathrm{d}N \,/\, (1+\alpha) = (1+\beta)\mathrm{E}[n]$.

2.1. Private behaviour

Households derive utility from real consumption u_n (utility of consuming a basket of clean and dirty goods) and disutility from labour supply l_n. Without loss of generality, wages per unit of productivity, the price of clean products and the before-tax price of dirty products are normalised to one.

We suppose quasi-linear preferences: $v_n = u_n - \dfrac{l_n^{1+1/\varepsilon}}{1+1/\varepsilon}$. The (un)compensated wage elasticity of labour supply, $\varepsilon > 0$, is constant. The household budget constraint is given by $pu_n = (1-t)nl_n + s$, where t denotes the income tax rate on labour income, s the income transfer (unconditional on productivity n), and p the ideal price index for real consumption. The wage per unit of skill has been normalised to one. Choosing labour supply l_n and real consumption u_n to maximise utility subject to the budget constraint, yields the following expressions for indirect utility, labour supply l_n and real consumption of individual n:

(1) $$v_n = \frac{1}{\varepsilon+1}\left(\frac{(1-t)n}{p}\right)^{\varepsilon+1} + \frac{s}{p}, \quad l_n = \left(\frac{(1-t)n}{p}\right)^{\varepsilon} \text{ and } u_n = \left(\frac{(1-t)n}{p}\right)^{\varepsilon+1} + \frac{s}{p},$$

Labour supply only depends on the substitution effect, not the income effect. Real consumption is determined residually from the household budget constraint.

Real consumption u_n is a homothetic sub-utility function over consumption of clean goods c_n, and dirty ('bad') goods b_n: $u_n = \left(c_n^{\frac{\sigma-1}{\sigma}} + b_n^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$ if $\sigma \neq 1$, $u_n = \ln(c_n) + \ln(b_n)$ if $\sigma = 1$. The household budget constraint is $pu_n = c_n + (1+q)b_n$, where q denotes the carbon tax rate. The price of clean products and the before-tax price of dirty products have been normalised to one. Due to homothetic preferences, the price index (or the unit-expenditure function) of the basket of clean and dirty goods can be written as $p = \mathrm{p}(q)$.

With homothetic preferences, the elasticity of substitution (σ > 0) between the two goods is constant and the budget shares are independent of individual productivities. The corresponding unit-expenditure function and demands for dirty and clean goods are given by:

$$\text{(2)} \qquad \mathrm{p}(q) = \left[1 + (1+q)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \quad \Rightarrow \quad b_n = \frac{\mathrm{p}(q)u_n}{(1+q)^{\sigma} + 1 + q} \text{ and } c_n = \frac{\mathrm{p}(q)u_n}{1+(1+q)^{1-\sigma}},$$

Since $b_n = (1+q)^{-\sigma} c_n$, a higher dirt tax q induces substitution away from dirty to clean goods. Cobb-Douglas preferences imply σ = 1 and $\mathrm{p}(q) = \sqrt{1+q}$, so half the budget is allocated to clean goods and half to dirty goods. In general, a carbon tax only induces a smaller share of the budget spent on dirty goods if the substitution effect outweighs the income effect (i.e., σ > 1).

Future preferences are the same as today's preferences, hence we have:

$$\text{(3)} \qquad L_N = \left(\frac{(1-T)N}{\mathrm{p}(Q)}\right)^{\varepsilon}, \quad U_N = \left(\frac{(1-T)N}{\mathrm{p}(Q)}\right)^{\varepsilon+1} + \frac{S}{\mathrm{p}(Q)}, \; V_N = \frac{1}{\varepsilon+1}\left(\frac{(1-T)N}{\mathrm{p}(Q)}\right)^{\varepsilon+1} + \frac{S}{\mathrm{p}(Q)},$$
$$B_N = \frac{\mathrm{p}(Q)U_N}{(1+Q)^{\sigma} + 1 + q} \text{ and } C_N = \frac{\mathrm{p}(Q)U_N}{1+(1+Q)^{1-\sigma}}.$$

where $\partial V_N / \partial S = 1/P > 0, \partial V_N / \partial T = -NL_N / P < 0$ and $\partial V_N / \partial Q = -B_N / P < 0.$

2.2. Global warming damages and future climate uncertainty

Total CO2 emissions (normalised to consumption of dirty goods) correspond to the total consumption of dirty goods, $b \equiv \int_0^{\infty} b_n \mathrm{f}(n)\mathrm{d}n,$ which is calculated as:

$$\text{(4)} \qquad b = \int_0^{\infty} \mathrm{p}'(q)u_n \mathrm{f}(n)\mathrm{d}n = \int_0^{\infty} \frac{\mathrm{p}'(q)}{\mathrm{p}(q)}\left[(1-t)n\mathrm{v}'\left(\frac{(1-t)n}{\mathrm{p}(q)}\right) + s\right]\mathrm{f}(n)\mathrm{d}n \equiv \mathrm{b}(q,t,s).$$

The partial derivatives of the carbon pollution function defined in (4) can be written as follows:

$$\text{(4')} \qquad \mathrm{b}_q = -\left[\sigma + (1-\nu)\sigma)\right]\frac{b}{1+q} - \varepsilon\left(\frac{\nu}{1+q}\right)^2 (1-t)w < 0, \quad \mathrm{b}_t = -(1+\varepsilon)\nu\frac{w}{1+q} < 0 \text{ and } \mathrm{b}_s = \frac{\nu}{1+q} > 0,$$

where $0 < \nu \equiv (1+q)\mathrm{p}'(q) / p < 1$ indicates the budget share of dirty goods and w today's gross before-tax wage income (defined in equation (9) below). A higher dirt tax rate q has three effects. First, given real household consumption or utility u_n, a higher dirt tax rate induces substitution away from dirty towards clean goods (as $p''(q) > 0$, especially if σ is high). This is the substitution effect. Second, higher consumer prices depress real household income and thus reduce spending including that on dirty goods (especially if v is high). This is the income effect. Third, higher consumer prices depress the consumer wage which curbs labour supply (especially if ε is high) and thus cuts household income and spending including that on dirty goods. All three effects imply that a higher dirt tax rate depresses aggregate carbon pollution. A higher tax rate on labour income t reduces after-tax labour income, both directly and indirectly via the fall in labour supply. This depresses consumer expenditure including that of dirty goods; and therefore leads to lower CO2 emissions. A higher income transfer s boosts expenditure including dirty goods and thus boosts CO2 emissions. A higher average productivity boosts labour supply and induces more emissions.

Global warming is to an extent anthropogenic, so we suppose that CO2 emissions are given by $e = \mathrm{e}(b, g), \mathrm{e}_b > 0, \mathrm{e}_g < 0,$ where g stands for today's spending on CO2 abatement. We suppose that damages to today's social welfare from climate change today are $\mathrm{x}(e)$, $\mathrm{x}' > 0$, $\mathrm{x}'' > 0$, $\mathrm{x}''' \geq 0$ and to future social welfare are $\mathrm{X}(\mathrm{e}(\kappa b, g), G)$, $\mathrm{X}_e > 0$, $\mathrm{X}_G < 0$, $\mathrm{X}_{ee} > 0$, $\mathrm{X}_{GG} > 0$, $\mathrm{X}_{eee} > 0$, where G indicates future spending on climate adaptation and κ denotes an exogenous shock to future global warming damages. The shock κ is known to the future government, but not to today's government. We normalise so that $\mathbf{E}[\kappa] = 1$. The possibility of an extreme worsening of the atmospheric CO2 concentration is modelled by assuming that κ takes on the value $1 - \pi K/(1 - \pi)$ (close to one) with probability $1 - \pi$ and the value $1 + K >> 1$ with a near-zero probability π. Section 3 shows that climate uncertainty together with convex marginal damages of global warming ($\mathrm{X}_{eee} > 0$) gives cautious discounting of future marginal damages.

Most global-warming damages occur many decades later, but present generations still value good stewardship of our planet and care about future climate change. We suppose that the future has resolved the problem of CO2 emissions and has managed to find carbon-free substitutes for fossil fuels. In that case, the optimal future carbon tax is zero, $Q = 0$. We suppose that CO2 abatement and climate adaptation measures consist of clean goods and thus have a price of one. Curbing CO2 emissions can be achieved either via a higher dirt or higher income tax rate. Both can make the income distribution more unequal and lead to tradeoffs between climate and redistribution.

2.3. Inequality aversion and equally-distributed-equivalent private utility

Before we specify social welfare, we define the concept of equally-distributed-equivalent private utility:

$$(5) \qquad v^{EDE} \equiv \Psi^{-1}\left(\int_0^\infty \Psi\left(v_n\right)\mathrm{f}(n)\mathrm{d}n\right) \text{ and } V^{EDE} \equiv \Psi^{-1}\left(\int_0^\infty \Psi\left(V_N\right)\mathrm{F}(N)\mathrm{d}N\right), \quad \Psi' > 0, \Psi'' \leq 0, \Psi''' \geq 0,$$

where the concave function $\Psi(.)$ captures society's intra-generational inequality aversion. We use $\Psi(v_n) = v_n^{1-\psi} / (1-\psi)$ for $\psi \neq 1$ and $\Psi(v_n) = \ln(v_n)$ for $\psi = 1$, which is based on Atkinson's entropy measure of income inequality aversion with ψ indicating the degree of intra-generational inequality aversion (Atkinson, 1970). The case Ψ' =1 (or ψ = 0) corresponds to utilitarian preferences relevant and zero intra-generational inequality aversion. In general, $\Psi' = v_n^{-\psi}$ so that the social weight given to individuals with low utility is higher than those with high utility, especially if ψ is high. However, if $\Psi(.)$ is concave, $\Psi(\mathbf{E}[v_n])$ exceeds $\mathbf{E}[\Psi(v_n)]$ so $v^{EDE} = \Psi^{-1}(\mathbf{E}[\Psi(v_n)])$ must fall short of $\mathbf{E}[v_n]$. This cost of inequality reflects the social planner's aversion to intra-generational inequality. The case $\psi \to \infty$ corresponds to Rawlsian max-min social welfare.

Upon substitution of the indirect utilities from (1) and (3) into (5) and integration across the individuals today and in the future, we obtain $v^{EDE} = \mathrm{v}^{EDE}(q,t,s)$ and $V^{EDE} = \mathrm{V}^{EDE}(T,S).$ Atkinson's measure of inequality is given by $(\mathbf{E}[V_N]-V^{EDE})/\mathbf{E}[V_N]$ and tells us how much utility society is disposed to forsake in order to have equally distributed utilities. This is higher the bigger the dispersion and skewness of the distribution of productivities, the higher the degree of inequality aversion, and the more redistributive policies are. It is useful to have available for future reference the following partial derivatives:

$$(5') \qquad \begin{gathered} \mathrm{v}_q^{EDE} = -\frac{v}{1+q}\left[\frac{s\upsilon}{\mathrm{p}(q)} + \frac{(1-t)(1-\delta)\upsilon}{\mathrm{p}(q)}\right] < 0, \quad \mathrm{v}_t^{EDE} = -\frac{w(1-\delta)\upsilon}{\mathrm{p}(q)} < 0, \quad \mathrm{v}_s^{EDE} = \frac{\upsilon}{\mathrm{p}(q)} > 0, \\ \mathrm{V}_T^{EDE} = -\frac{\Delta}{\mathrm{p}(0)} < 0 \text{ and } \mathrm{V}_S^{EDE} = \frac{\Upsilon}{\mathrm{p}(0)} > 0, \end{gathered}$$

where the marginal social costs of inequality υ and Υ and Feldstein's distributional characteristics for gross wage income, δ and Δ, are defined by

$$(6) \qquad \begin{gathered} \upsilon \equiv \frac{\mathbf{E}\left[\Psi'(v_n)\right]}{\Psi'(v^{EDE})} = \upsilon(\bar{q},\bar{t},\bar{s}) \geq 1, \quad 0 < \Upsilon \equiv \frac{\mathbf{E}\left[\Psi'(V_N)\right]}{\Psi'(V^{EDE})} = \Upsilon(\bar{T},\bar{S}) \geq 1, \\ 0 < \delta \equiv -\frac{\mathrm{cov}\left[\Psi'(v_n), nl_n\right]}{\mathbf{E}\left[\Psi'(v_n)\right]\mathbf{E}\left[nl_n\right]} = 1 - \frac{\mathbf{E}\left[\Psi'(v_n)nl_n\right]}{\mathbf{E}\left[\Psi'(v_n)\right]\mathbf{E}\left[nl_n\right]} = \delta(\bar{q},\bar{t},\bar{s}) \leq 1 \text{ and} \\ 0 < \Delta \equiv -\frac{\mathrm{cov}\left[\Psi'(V_N), NL_N\right]}{\mathbf{E}\left[\Psi'(V_N)\right]\mathbf{E}\left[NL_N\right]} = 1 - \frac{\mathbf{E}\left[\Psi'(V_N)NL_N\right]}{\mathbf{E}\left[\Psi'(V_N)\right]\mathbf{E}\left[NL_N\right]} = \Delta(\bar{T},\bar{S}) \leq 1. \end{gathered}$$

The marginal social costs of inequality (MSCIs) are normalised by the marginal social utility of equally-distributed-equivalent utility to convert from marginal social welfare to resource units. In the absence of intra-generational inequality aversion ($\Psi' = 1$), the MSCIs equal one. Intra-generational equality aversion generally implies that $\Psi'(v^{EDE}) < \mathbf{E}[\Psi'(v_n)]$ especially if the distribution of productivities is very skewed and intra-generational inequality aversion is strong enough. In that case, inequality aversion pushes the MSCIs, υ and Υ, above one. Since a higher income transfer benefits the poor relatively more and a higher income tax rate particularly hits the richest, we conjecture that the MSCIs decrease in these policies. A higher carbon tax rate hits labour supply of the most able people the most and is *ceteris paribus* also a progressive measure, hence we conjecture that today's MSCI also decreases in q.

The distributional characteristics δ and Δ are zero under utilitarian social welfare ($\Psi' = 1$), since private marginal utility of income is constant. With inequality aversion ($\Psi'' < 0$), they correspond to normalised covariances and, as welfare weights fall with income, these covariances are negative and therefore the characteristics δ and Δ are positive. For a Rawlsian social welfare function with $\psi \rightarrow \infty$, δ and Δ equal one, but else they are less than one. These characteristics are strictly positive even if there is no income transfer ($S = 0$).[8] The characteristics δ and Δ give the social value – expressed in monetary terms as a fraction of average taxable wage income – of one extra resource unit of intra-generational redistribution. They are therefore smaller if intra-generational redistribution is already strong, that is if the income tax is progressive (indicated by a high income tax rate and income transfer), and the carbon tax rate is high.

Most studies on optimal taxation employ a more conventional welfare approach which corresponds to using the social welfare functions $\int_0^\infty \Psi\left(v_n\right)\mathrm{f}(n)\mathrm{d}n$ and $\int_0^\infty \Psi\left(V_N\right)\mathrm{F}(N)\mathrm{d}N$ instead of $\mathrm{v}^{EDE}\left(q,t,s\right)$ and $\mathrm{V}^{EDE}(T,S)$, respectively.[9] The expressions for δ and Δ in (6) and the resulting optimal carbon and income tax formulae and the expression for the optimal spending on CO2 abatement and climate adaptation measures today and the future that will be derived in section 3 are then identical, except that the definitions of the MSCIs are no longer normalised by the marginal social utility of EDE utility:

---

[8] For example, for the Galton-distribution we obtain $\Delta = \exp\left[\frac{1}{2}(\varepsilon+1)^2\theta^2(1-\psi)^2\right] / \exp\left[\frac{1}{2}(\varepsilon+1)^2\theta^2(1+\psi)^2\right]$ if $S = 0$, which is less than one provided that $\psi < 1$ and the standard deviation of ln(n), θ, is positive.

[9] Equivalent income is a money metric value that, unlike utility, is not affected by monotonic transformations of utility. Within the context of an optimal linear income tax model, the optimal tax rate is more sensitive to inequality aversion if equally-distributed-equivalent utility is used rather than ordinary utility. With Cobb-Douglas and CES utility, the optimal tax is the same under the two approaches if $\psi = 1$, and the case for high marginal rates does not depend on a low elasticity of substitution between consumption and leisure (Creedy, 2003).

(6′) $$\upsilon \equiv \mathbf{E}\left[\Psi'(v_N)\right] = \upsilon(\bar{q}, \bar{t}, \bar{s}) \text{ and } \Upsilon \equiv \mathbf{E}\left[\Psi'(V_N)\right] = \Upsilon(\bar{T}, \bar{S}).$$

2.4. Social welfare and government behaviour

Social welfare is the present discounted value of equally-distributed-equivalent private utilities times the size of the population, minus global warming damages:

(7) $$\omega\left(q,t,s,g,T,S,G\right) \equiv \left(\frac{1}{1-\zeta}\right)\mathrm{v}^{EDE}(q,t,s)^{1-\zeta} - \mathrm{x}\left(\mathrm{e}\left(\mathrm{b}(q,t,s),g\right)\right) + \left(\frac{1}{1-\zeta}\right)\left(\frac{1+\alpha}{1+\rho}\right)\mathrm{V}^{EDE}(T,S)^{1-\zeta} - \left(\frac{1}{1+\rho}\right)\mathrm{E}\left[\mathrm{X}\left(\mathrm{e}\left(\kappa \mathrm{b}(q,t,s),g\right),G\right)\right],$$

where $0 < \rho < 1$ indicates the pure rate of time preference and $\zeta \geq 0$ denotes the coefficient of relative inter-generational inequality aversion. By completely separating environmental welfare from (distribution-equivalent) non-environmental welfare, we can fully trace the trade-off between environmental quality and distributional concerns in an analytical fashion. Moreover, this assumption can be defended by presuming that future climate damages are assessed by scientific experts, such as the Intergovernmental Panel on Climate Change, rather than being the socially weighted sum of each individual's willingness to pay to avoid future climate damages.

Present and future generations of private individuals are disconnected, but today's government can bring forward economic assets to the future. These assets can be physical capital, public infrastructure, stock of R&D, building dykes and other water defences, fighting desertification, etcetera and are denoted by a. The initial stock is $\bar{a} > 0$. The present budget constraint states that the excess of income and dirt tax revenue over the cost of income transfers and CO2 abatement is the increase in society's stock of assets:

(8) $$a = \bar{a} + \int_0^\infty (tnl_n + qb_n)\mathrm{f}(n)\mathrm{d}n - s - g.$$

The future government inherits this stock of assets. Let r denote the return on these assets (the marginal product of these assets minus a provision for physical depreciation). The future budget constraint then says that inherited assets plus labour and dirt taxes must be sufficient to pay for the transfers to all future citizens plus future spending on climate adaptation to limit global warming:

(8′) $$(1+r)a + (1+\alpha)\int_0^\infty TNL_N \mathrm{f}(N)\mathrm{d}N = (1+\alpha)S + G.$$

The present-value budget constraint says that today's surplus plus interest should equal the future deficit:

$$(8'') \qquad (1+r)\left(\bar{a} + t\mathrm{w}(q,t) + q\mathrm{b}(q,t,s) - s - g\right) = (1+\alpha)S + G - T\mathrm{W}(T),$$

where aggregate present and future gross wage income are given by, respectively,

$$(9) \qquad \mathrm{w}(q,t) \equiv \int_0^\infty n\mathrm{v}'\left((1-t)n/\mathrm{p}(q)\right)\mathrm{f}(n)\mathrm{d}n \;\text{ and }\; \mathrm{W}(T) \equiv \int_0^\infty N\mathrm{v}'\left((1-T)N/\mathrm{p}(0)\right)\mathrm{F}(N)\mathrm{d}N.$$

with partial derivatives $\mathrm{w}_q = -\varepsilon\nu\frac{w}{1+q} \le 0, \quad \mathrm{w}_t = -\varepsilon\frac{w}{1-t} \le 0$ and $\mathrm{W}' = -\varepsilon\frac{W}{1-T} \le 0.$ With inelastic labour supply ($\varepsilon = 0$), today's gross wage bill equals $\mathbf{E}[n]$ as both the wage and today's population are normalised to one. The future gross wage bill is bigger as there will be more workers and each of them will be more productive, $\mathrm{W}(T) = (1+\alpha)(1+\beta)\mathbf{E}[n]$. With elastic labour supply, a higher carbon tax rate boosts consumer prices, depresses the real wage, curbs labour supply and thus lowers wage income. A higher income tax rate then also depresses labour supply and wage income.

3. Cautious discounting of global warming damages

3.1. Future government

The future government chooses the income tax rate T, income transfer S and spending on climate adaptation G to maximise future welfare, $\max_{T,S,G} \Omega(T,S,G) = (1+\alpha)\mathrm{V}^{EDE}(T,S)^{1-\zeta}/(1-\zeta) - \mathrm{X}(\mathrm{e}(\kappa b, g), G),$

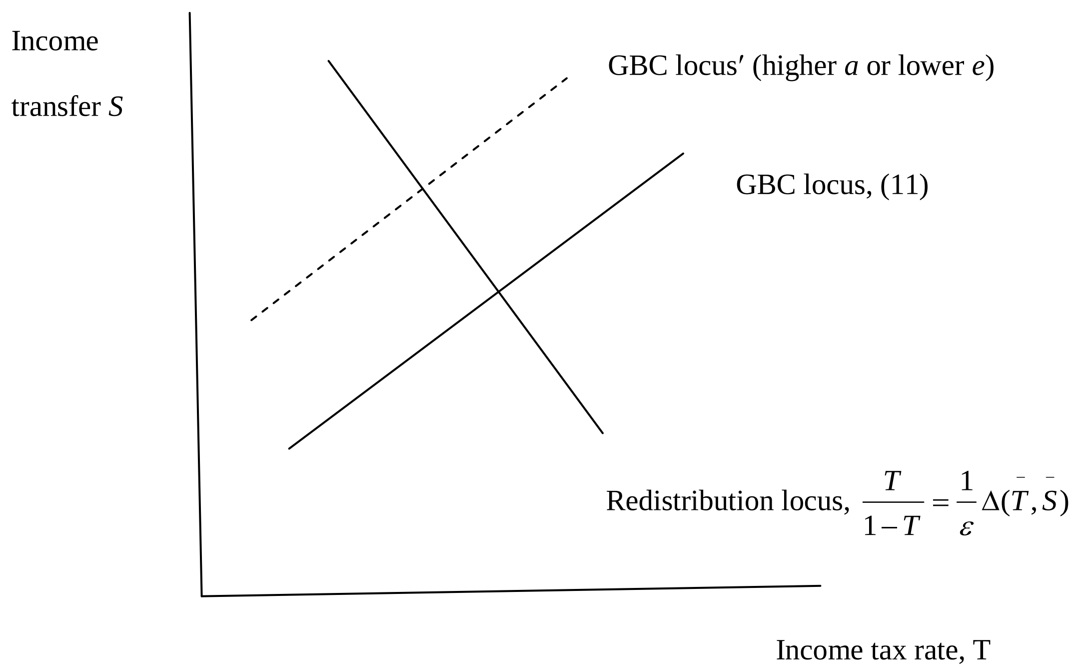
subject to the budget constraint $(1+r)a + T\mathrm{W}(T) = (1+\alpha)S + G,$ conditional on inherited productive assets, inherited carbon debt and the realised shock to climate damages. Denoting by Λ the social cost of society's assets (the Lagrange multiplier of the future budget constraint), we obtain the optimality conditions for the future government (see appendix for the derivations):

$$(10\text{a}) \qquad \frac{T}{1-T} = \frac{1}{\varepsilon}\Delta(\overset{-}{T},\overset{-}{S}) \quad \Rightarrow \quad T = \mathrm{T}(\overset{-}{S}),$$

$$(10\text{b}) \qquad \frac{-\mathrm{X}_G\left(\mathrm{e}(\kappa \overset{-}{b}, g), \overset{+}{G}\right)}{\Lambda} = \frac{-\mathrm{p}(0)\mathrm{X}_G\left(\mathrm{e}(\kappa \overset{-}{b}, g), \overset{+}{G}\right)}{\Upsilon(\overset{-}{T},\overset{-}{S})\mathrm{V}^{EDE}(\overset{-}{T},\overset{+}{S})^{-\zeta}} = 1 \quad \Rightarrow \quad G = \mathrm{G}\left(\mathrm{e}(\kappa \overset{+}{b}, g), \overset{+}{T}, \overset{+/-}{S}\right).$$

Equation (10a) indicates that the optimal income tax rate is high if labour supply is relatively inelastic (low ε), productivities are fairly equally distributed and society cares a lot about avoiding intra-generational income inequality (high Δ, high ψ). If the marginal value of redistribution Δ given in (6) decreases in *T* and *S*, the efficiency condition for the optimal tax rate implies a negative relationship between the optimal income tax rate and the optimal income transfer. This intra-generational redistribution trade-off is plotted in fig. 1 as a downwards-sloping locus, which is closer to the origin if there is little intra-generational inequality aversion (low ψ) and labour supply is elastic (high ε).

Figure 1: More inherited assets and a lower carbon debt makes future redistribution easier



The modified Samuelson-rule (10b) sets the marginal benefit of climate adaptation measures (converted into resource units by dividing by the marginal social cost of productive assets, $\Lambda = \Upsilon V^{EDE^{-\zeta}} / \mathrm{p}(0)$) to the marginal rate of transformation (i.e., one). This implies a unit marginal cost of public funds (cf., Jacobs and de Mooij, 2010). Intuitively, the government equalises the marginal cost of funds for all tax instruments. Since it has a non-distorting lump-sum tax with a unit marginal cost of funds at its disposal, the marginal cost of funds for the distorting tax should be equal to one as well at the optimum. Optimal spending on climate adaptation decreases in the marginal social cost of inequality (MSCI), so that a greater need for intra-generational redistribution towards the poor (higher ψ and θ, so higher Υ) and a less ambitious intra-generational redistribution policy (lower income tax rate *T* and lower transfers *S*) imply less spending on climate adaptation. Of course, without intra-generational aversion to inequality

MSCI = 1 and optimal *G* does not depend on *T* and *S*. However, with intra-generational equality aversion MSCI > 1 and thus (10b) indicates that optimal *G* is lower. The intuition is that a society that has a lot of intra-generational inequality which it cares about must give priority to its redistribution objective and cannot afford to be as ambitious about fighting global warming. Finally, with a lot of inter-generational inequality aversion (ζ large and positive), there is an additional positive effect of a higher income transfer on the optimal level of spending on climate adaptation. However, there is an offsetting negative effect of a higher income tax rate on spending on climate adaptation. So, if inter-generational solidarity is less important than intra-generational solidarity, a higher income tax rate goes hand in hand with lower spending on climate adaptation. If abatement/adaptation affects social damages of global warming in a non-separable fashion from emissions, there is a positive direct effect of emissions on abatement and adaptation (on top of the indirect effects operating via the Lagrange multiplier).

Substitution of (10b) into the future government budget constraint yields:

$$(1+r)a + TW(T) = (1+\alpha)S + \mathrm{G}(\mathrm{e}(\kappa b, g), T, S), \tag{11}$$

which corresponds to the upwards-sloping GBC locus in fig. 1, so a higher labour income tax boosts tax revenue by more than it increases spending on adaptation. Intersection of the redistribution locus (10a) and the GBC loci (11) gives the optimal income tax and income transfer, and thus spending on adaptation:

$$\begin{gathered} T = \mathrm{T}(\overset{-}{a}, \mathrm{e}(\kappa \overset{+}{b}, g)), \quad S = \mathrm{S}(\overset{+}{a}, \mathrm{e}(\kappa \overset{-}{b}, g)) \text{ and} \\ G = \mathrm{G}(\mathrm{e}(\kappa \overset{+}{b}, g), \mathrm{T}(\overset{-}{a}, \mathrm{e}(\kappa \overset{+}{\overset{+}{b}}, g)), \mathrm{S}(\overset{+}{a}, \mathrm{e}(\kappa \overset{+/-}{\overset{-}{b}}, g))) \equiv \mathrm{G}^*(\overset{(+)}{a}, \mathrm{e}(\kappa \overset{+}{b}, g)). \end{gathered} \tag{12}$$

A higher inherited stock of assets and a lower carbon debt induced by lower past emissions (necessitating less future spending on climate adaptation if X(.) is non-separable) shift up the GBC locus which lowers the optimal income tax rate and boosts the optimal income transfer. This makes the future economy richer and easier to pursue more aggressive intra-generational income redistribution.

Substituting the optimal policies (12) into future social welfare, we obtain:

$$\begin{gathered} \Omega(a, \mathrm{e}(\kappa b, g)) \equiv (1+\alpha)\mathrm{V}^{EDE}(\mathrm{T}(a, \mathrm{e}(\kappa b, g)), \mathrm{S}(a, \mathrm{e}(\kappa b, g))^{1-\zeta} / (1-\zeta) - \mathrm{X}(\mathrm{e}(\kappa b, g), \mathrm{G}(a, \mathrm{e}(\kappa b, g)) \\ \Omega_a = (1+r)\Lambda > 0, \ \Omega_e = -\mathrm{X}_e < 0, \end{gathered} \tag{12′}$$

where derivatives follow from the Envelope Theorem. Leaving more assets to future generations boosts future welfare, since it permits a higher income transfer, a lower income tax rate and more spending on climate adaptation. A large carbon debt and future climate disasters lower future social welfare.

3.2. Today's government

Today's government does not know what the shock will be to the future climate and thus has to maximise social welfare taking account of this uncertainty. It therefore maximises expected social welfare

$$(13)\quad \max_{q,t,s,g} \omega\left(q,t,s,g\right) = \frac{1}{1-\zeta}\mathrm{v}^{EDE}(q,t,s)^{1-\zeta} - \mathrm{x}(\mathrm{e}(\mathrm{b}(q,t,s),g)) + \frac{1}{1+\rho}\mathbf{E}\left[\Omega\left(a,\mathrm{e}(\kappa\mathrm{b}(q,t,s),g)\right)\right],$$

subject to today's budget constraint $a = \overline{a} + t\mathrm{w}(q,t) + q\mathrm{b}(q,t,s) - s - g$, aggregate CO2 emissions (4), v^{EDE} from (5), gross wage income (9) and future social welfare (12′), where λ is today's social cost of productive assets (Lagrange multiplier of today's budget constraint). We obtain the following proposition.

Proposition 1: The optimal precautionary climate change and income redistribution policies follow from:

$$(14a)\qquad \gamma^b \equiv \frac{\mathbf{E}\left[\kappa \mathrm{e}_{\kappa b}(\kappa b,g)\mathrm{X}_e\left(\mathrm{e}(\kappa b,g),G\right)\right]}{\mathrm{e}_b(b,g)\mathrm{X}_e\left(\mathrm{e}(b,g),G\right)} \geq 1, \quad \gamma^g \equiv \frac{\mathbf{E}\left[-\mathrm{e}_g(\kappa b,g)\mathrm{X}_e\left(\mathrm{e}(\kappa b,g),G\right)\right]}{-\mathrm{e}_g(b,g)\mathrm{X}_e\left(\mathrm{e}(b,g),G\right)} \geq 1,$$

$$(14b)\qquad q = \left[\mathrm{x}'\left(\mathrm{e}(b,g)\right) + \left(\frac{\gamma^b}{1+\rho}\right)\mathrm{X}_e\left(\mathrm{e}(b,g),G\right)\right]\frac{\mathrm{e}_b(b,g)}{\lambda} \equiv q^P,$$

$$(14c)\qquad \frac{-\mathrm{e}_b(b,g)}{\mathrm{e}_g(b,g)} = q \quad \Rightarrow \quad g = \mathrm{g}\left(\overset{+}{b},\overset{+}{q}\right),$$

$$(14d)\qquad \lambda = -\left[\mathrm{x}'\left(\mathrm{e}(b,g)\right) + \left(\frac{\gamma^g}{1+\rho}\right)\mathrm{X}_e\left(\mathrm{e}(b,g),G\right)\right]\mathrm{e}_g(e,g) = -\left(\frac{1+r}{1+\rho}\right)\mathbf{E}\left[\mathrm{X}_G\left(\mathrm{e}(\kappa b,g),G\right)\right],$$

$$(14e)\qquad t = \frac{\delta(q,t,s)}{\delta(q,t,s)+\varepsilon}, \quad T = \frac{\Delta(T,S)}{\Delta(T,S)+\varepsilon},$$

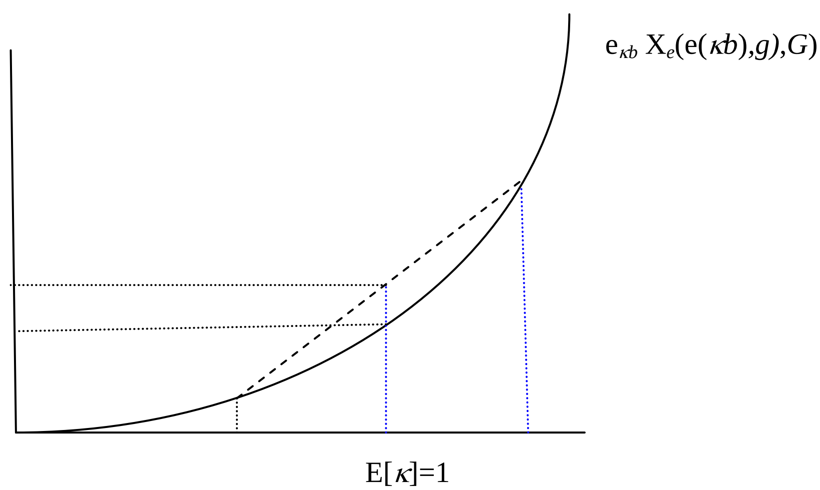
$$(14f)\qquad \left(\frac{1+r}{1+\rho}\right)\left[\frac{\mathbf{E}[\Upsilon(T,S)\mathrm{V}^{EDE}(T,S)^{-\zeta}]}{\mathrm{p}(0)}\right] = \lambda = \frac{\upsilon(t,q,s)\mathrm{v}^{EDE}(q,t,s)^{-\zeta}}{\mathrm{p}(q)},$$

where γ^b and γ^g denote the prudence factors (i.e., the additional weights for evaluating future marginal damages of CO2 emissions and benefits of CO2 abatement, respectively).

Proof: Appendix 2 establishes this proposition as a special case of proposition 3.

Without shocks to climate change (14a) indicates that $\gamma^b = \gamma^g = 1$ so the same discount rate is used to discount marginal changes in EDE-utility as the green part of social welfare, namely the pure rate of time preference ρ. With a small probability of future climate change, (14a) implies that a prudent policy maker reduces the effective rate used to discount marginal global warming damages so that future marginal global warming damages are weighed more heavily (cf., Stern, 2007; Weitzman, 2010). To see this, define $\gamma \equiv \frac{\mathbf{E}\left[\mathrm{e}_{\kappa b}(\kappa b, g)\mathrm{X}_e\left(\mathrm{e}(\kappa b, g), G\right)\right]}{\mathrm{e}_b(b, g)\mathrm{X}_e\left(\mathrm{e}(b, g), G\right)}$ and $\pi \equiv \frac{\mathrm{cov}\left[\kappa, \mathrm{e}_{\kappa b}(\kappa b, g)\mathrm{X}_e\left(\mathrm{e}(\kappa b, g), G\right)\right]}{\mathbf{E}[\mathrm{e}_{\kappa b}(\kappa b, g)\mathrm{X}_e\left(\mathrm{e}(\kappa b, g), G\right)]\mathbf{E}[\kappa]} > 0$. Fig. 2 shows that with uncertainty about future shocks to the atmospheric CO2 concentration κ, we have $\gamma^b > 1$ if the marginal damage function is convex, that is if the third derivative of the damage function with respect to κb is positive. The risk-premium on global warming damages, π, corresponds to the correlation coefficient between climate shocks and marginal damages from global warming which must be positive.

Figure 2: Climate uncertainty and prudence



Hence, we must have $\gamma^b = \gamma(1+\pi) > 1$. Now define $\rho^* \equiv \frac{1+\rho-(1+\pi)\gamma}{(1+\pi)\gamma} < \rho$ as the cautious rate that is used to discount future marginal global warming damages, so that $\frac{\gamma^b}{1+\rho} = \frac{1}{1+\rho^*}$. The extent to which ρ^* is reduced below ρ depends on how bad the size of a future climate disaster might be, how unlikely it is too happen and how prudent the government is (i.e. how convex the marginal damages function is). Note that the cautious discount rate ρ^* might even be negative.

If $\gamma^b = 1$ and $\rho = \rho^*$, the optimal carbon tax rate (14b) is the discounted value of future marginal global warming damages divided by the shadow cost of accumulating public assets, λ. It corresponds to the

Pigouvian tax rate q^P, which is higher if marginal global warming damages are high and the government employs a low discount rate. However, with climatic uncertainty, $\gamma^b > 1$ and ρ^* is reduced below ρ. In that case, (14b) indicates that it is prudent to weigh future marginal global warming damages more heavily and that is optimal to set a higher carbon tax rate. Equation (14c) indicates that the cost of abating one unit of CO2emissions must equal the social price of carbon. This implies that optimal spending on CO2 abatement is high if the carbon tax rate is high and typically also if CO2 emissions are high.

The first part of (14d) is the Samuelson-rule for today's optimal spending on CO2 abatement g. It states that the marginal benefit of CO2 abatement today and the cautiously discounted benefit in the future in resource units (i.e., after dividing by λ) must equal the marginal cost of public funds. The cost of public funds must be one, since at the margin the government has access to lump-sum taxes (cf., Jacobs and de Mooij, 2010). The second part of (14d) is the Samuelson-rule for future climate abatement spending G. Together they imply that the marginal return on CO2 abatement must equal that on climate adaptation properly accounting for prudent discounting of future marginal benefits of abatement. Optimal spending on abatement and adaptation thus go up and down together. If the government is impatient and uses a higher rate of pure time preference than the return on productive assets ($\rho > r$), (14d) implies that there is a bias towards future climate adaptation, away from present CO2 abatement. Conversely, using an almost zero rate of pure time preference (cf., Stern, 2007) leads to a shift from climate adaptation towards CO2 abatement. If there is future climate uncertainty and marginal global warming damages are convex, the government discounts future benefits of CO2 abatement less heavily, $\gamma^g > 1$, which also induces a shift towards from adaptation towards CO2 abatement. We also note from (14c) and (14d) that a higher carbon tax goes hand in hand with higher spending on both CO2 abatement and climate adaptation.

The optimal labour income tax rates (14e) increase in the Feldstein's distributional characteristics and decrease in the wage elasticity of labour supply, so they strike a balance between redistribution and efficient revenue-raising objectives.[10] The labour income tax rate is high if there is a big need and wish to redistribute incomes (high δ or high Δ), but low if it is difficult to raise revenue due to the adverse effect on labour supply, i.e., if labour supply is relatively elastic (low ε). With utilitarian social welfare ($\Psi' = 1$, $\psi = 0$) we note from (6) that $\delta = \Delta = 0$ and thus that the optimal income tax rates are zero. This results from our assumption of quasi-linear preferences and a constant marginal utility of money. With the

[10] Allowing for income effects in labour supply requires one to interpret ε as the compensated wage elasticity of labour supply and would imply that the denominator in the expressions for t and T would be reduced by the income elasticity of labour supply and thus the optimal tax rates are boosted. A higher tax rate reduces income and now boosts labour supply so that it is less costly to raises the tax rate on labour income. Also, the social valuations of income in the Feldstein-distributional characteristics will need to be corrected for income effects.

opposite extreme of Rawlsian max-min social welfare ($\psi \to \infty$), we have $\delta = \Delta = 1$, so that the optimal income tax rates equal the revenue-maximising rates $t = T = 1/(1+\varepsilon)$. With positive but finite aversion to inequality ($\Psi'' < 0$, $\psi > 0$), we have $0 < \delta < 1$ and $0 < \Delta < 1$ so that the optimal income tax rate trades off maximising revenue against redistribution motives. In that case, the optimal income tax rate are on the upwards part of the Laffer curve with t and T strictly less than $1/(1+\varepsilon)$.

Finally, (14f) is the intertemporal efficiency condition.[11] It states that today's shadow cost of government assets today must equal the discounted value of the expected future shadow cost government assets plus the return on government assets. An impatient government with a high rate of pure time preference ($\rho > r$) lowers the cost of assets today and raises it tomorrow, and therefore leaves less productive assets for the future. Conversely, an almost zero rate of pure time preference (cf., Stern, 2007) raises the costs of assets today relative to the cost in the future and thus leaves more assets for the future.

In general, the optimal carbon tax rate is positive and set at the prudence-adjusted Pigouvian level, irrespective of the redistributive preferences of the government, so that $q > 0$ and $\upsilon > \Upsilon > 1$ from (14f).[12] The larger is future uncertainty regarding climate change, the larger is γ^b and the higher is the carbon tax. If there is no future uncertainty regarding climate change and $\gamma^b = 1$, the optimal carbon tax is set at the Pigouvian level; otherwise, it is adjusted upwards for precautionary reasons. The carbon tax is neither used for redistributive reasons nor to reduce labour-tax distortions. Due to homothetic preferences individuals spend a proportional fraction of their income on dirty goods. Therefore, the linear income tax can redistribute equally well as the carbon tax, but avoids distortions in the consumption choices of households. Similarly, the carbon tax is not used to indirectly tax leisure-complements, which could alleviate distortions of the labour tax on labour supply (cf., Corlett and Hague, 1953), since both clean and dirty commodities are equally complementary to leisure (Jacobs and De Mooij, 2010). Thus, the carbon tax is used exclusively for efficiency reasons so as to internalise the environmental externality. The policy challenge is thus to reconcile Pigouvian objectives with intra-and inter-generational redistribution considerations.

---

[11] Note that equation (14f) corresponds to a stochastic version of the Keynes-Ramsey rule which says that the expected growth in the marginal utility of equally-distributed-equivalent utility times the MSCI equals the expected fall in the CPI times the ratio of the pure rate of time preference to the return on productive assets.

[12] With no intra-generational income inequality aversion ($\psi \to 0$), υ and Υ approach one. With also zero inter-generational solidarity ($\zeta \to 0$) and $r = \rho$, (14f) becomes $\mathrm{p}(q) = \mathrm{p}(0)$ and thus today's optimal carbon tax is zero given that in the future there are no CO2 emissions. Hence, a government with neither intra-generational nor inter-generational inequality aversion conducts climate policy primarily via CO2 abatement and climate adaptation.

4. Effects of subsistence needs for dirty goods on climate and redistribution policies

The analysis of sections 2 and 3 is not wholly satisfactory for studying the trade-off between climate change and redistribution objectives. Our chosen functional forms imply that the wage elasticity of labour supply and the elasticity of substitution between clean and dirty goods are constant and thus independent of individual productivities. Hence, the optimal labour income tax rate depends on changes in the Feldstein's inequality characteristic, not on changes in these two elasticities. Furthermore, the optimal labour income tax rate is only targeted at redistribution, not at fighting CO2 pollution. Conversely, the optimal carbon tax rate is targeted at CO2 damages, not at redistribution. This results from our assumption of quasi-linear and homothetic preferences, which results in constant expenditure shares of polluting goods. The income tax rate does not affect the allocation between clean and dirty goods and the dirt tax is superfluous as a distributional device. However, in practise poor people both within rich countries and people living in developing countries seem to consume a large fraction of dirty goods. They are therefore the ones most likely to be hurt by a steep carbon tax. At the same time, their price elasticities for dirty goods are relatively low and for clean goods high, which suggests that there are stronger Ramsey motives for taxing consumption of dirty than of clean products by the poor in addition to the Pigouvian motive to tax consumption of dirty goods. This puts even more demands on the income tax system to redistribute towards the poor. To make the trade-off between redistribution and climate change objectives more realistic and thus more relevant from a policy point of view, we now discuss the effects of non-constant expenditure shares for dirty commodities for optimal climate and redistribution policies.

We will assume Stone-Geary preferences so that one can aggregate across individuals:

(15) $$v_n(c_n, b_n, l_n) = 2\left(b_n - \bar{b}\right)^{\frac{1}{2}} c_n^{\frac{1}{2}} - \frac{l_n^{1+1/\varepsilon}}{1+1/\varepsilon},$$

where the subsistence level for dirty goods is denoted by $\bar{b} > 0$. This yields the following expressions for consumption of dirty goods, clean goods and the basket of both goods, and for labour supply:

(16) $$b_n = \frac{1}{2}\bar{b} + \frac{1}{2(1+q)}\left[(1-t)nl_n + s\right], \quad c_n = \frac{1}{2}\left[(1-t)nl_n + s - (1+q)\bar{b}\right],$$
$$u_n = \frac{(1-t)nl_n + s - (1+q)\bar{b}}{\sqrt{1+q}} \text{ and } l_n = \left(\frac{(1-t)n}{\sqrt{1+q}}\right)^{\varepsilon}.$$

Hence, individual n's labour supply, l_n, consumption of clean and dirty goods, c_n and b_n, and indirect utility v_n can expressed as functions of q, t, s and n. Labour supply has wage elasticity equal to ε as before, so the elasticity with respect to the price of dirty goods equals $\varepsilon/2$. These elasticities imply that clean and dirty goods are still equally complementary to leisure so that there is still no Corlett-Hague (1953) motive to tax leisure complements with the dirt tax. Consequently, this preference structure allows us to focus exclusively on the distributional consequences of non-constant expenditure shares. Indirect utility, aggregate carbon pollution and aggregate before-tax wage income are given by:

(17)
$$v_n = \left((1-t)nl_n + 2s - 2(1+q)\bar{b}\right)^{\frac{1}{2}} - \frac{\left(\frac{(1-t)n}{\sqrt{1+q}}\right)^{\varepsilon+1}}{1+1/\varepsilon},$$
$$\mathrm{b}(q,t,s) = \frac{1}{2}\bar{b} + \frac{1}{2(1+q)}\left[(1-t)\mathrm{w}(q,t) + s\right] \text{ and } \mathrm{w}(q,t) = \int_0^\infty n\left(\frac{(1-t)n}{\sqrt{1+q}}\right)^{\varepsilon} \mathrm{f}(n)\mathrm{d}n$$

Clean goods are luxuries with an income elasticity greater than one ($[(1-t)nl_n + s]/[(1-t)nl_n + s - (1+q)\bar{b}\,] > 1$). Dirty goods are necessities, since their income elasticity is less than unity ($[(1-t)nl_n + s]/[(1-t)nl_n + s + (1+q)\bar{b}\,] < 1$). The elasticity of dirty goods with respect to $1+q$ equals the income elasticity while the elasticity with respect to $1-t$ equals $(1-t)nl_n/[(1-t)nl_n + s + (1+q)\bar{b}\,]$. Note that these elasticities now depend on skills, which captures that the least able and poorest members of society are hurt most by the dirt tax.

Stone-Geary preferences imply that all goods must be substitutes. Hence, we do not put leisure in the Stone-Geary form because then a revenue-neutral tax reform will have zero effect on labour supply. This is true for any weakly separable specification of preferences. The following proposition characterises the optimal climate change and redistribution policies.

Proposition 2: With Stone-Geary preferences, the optimal policies follow from (14a), (14c), (14d), (14f), (10a) and the new expressions for today's optimal carbon tax rate and labour income tax rate:[13]

(14b′)
$$q = q^P + \tilde{q} < q^P, \quad \tilde{q} \equiv q - q^P = (1+q)\left(\frac{2\bar{\mu}\delta^b - \delta}{\bar{\theta} - \bar{\mu}}\right) < 0,$$

(14e′)
$$\frac{t}{1-t} = \frac{\delta}{\varepsilon} - \frac{\bar{\theta}}{2}\frac{\tilde{q}}{1+q} = \frac{\delta}{\varepsilon} + \frac{\delta - 2\delta^b\bar{\mu}}{2\left(\bar{\theta} - \bar{\mu}\right)} > \frac{\delta}{\varepsilon},$$

---

[13] We suppose that the least productive people have sufficient income to purchase their basic needs of dirty consumption goods, so that supernumerary income $(1-t)nl_n + s - (1+q)\bar{b}$ is never negative.

where $\bar{\mu} \equiv \int_0^\infty \left(\frac{(1+q)b_n}{(1-t)nl_n} \right) nl_n \mathrm{f}(n)\mathrm{d}n / w$ is the pre-tax-income-weighted share of spending on dirty goods in after-tax labour income (excluding transfer income), $\bar{\theta} \equiv \int_0^\infty \left(\frac{(1-t)nl_n+s}{(1-t)nl_n} \right) nl_n \mathrm{f}(n)\mathrm{d}n / w > 1$ is the pre-tax-income-weighted ratio of total income to labour income, $\delta = \frac{-\mathrm{cov}\left[nl_n, \frac{\left(v^{EDE}\right)^{-\xi} \Psi'(v_n)}{\Psi'\left(v^{EDE}\right)} + p\lambda\tilde{q}\frac{\partial b_n}{\partial s} \right]}{\mathbf{E}\left[nl_n \right]\mathbf{E}\left[\frac{\left(v^{EDE}\right)^{-\xi} \Psi'(v_n)}{\Psi'\left(v^{EDE}\right)} + p\lambda\tilde{q}\frac{\partial b_n}{\partial s} \right]} > 0,$ is the distributional characteristic of the labour tax, $\delta^b = \frac{-\mathrm{cov}\left[b_n, \frac{\left(v^{EDE}\right)^{-\xi}}{\Psi'\left(v^{EDE}\right)} \Psi' + p\lambda\tilde{q}\frac{\partial b_n}{\partial s} \right]}{\mathbf{E}\left[b_n \right]\mathbf{E}\left[\frac{\left(v^{EDE}\right)^{-\xi} \Psi'(v_n)}{\Psi'\left(v^{EDE}\right)} + p\lambda\tilde{q}\frac{\partial b_n}{\partial s} \right]}$ is the distributional characteristic of the carbon tax (supposed to be negative), both corrected for the deviation of the optimal carbon tax rate from its Pigouvian level defined in (14b), $\upsilon \equiv \mathbf{E}\left[\frac{\Psi'(v_n)}{\Psi'\left(v^{EDE}\right)} + p\lambda\left(v^{EDE}\right)^{\xi}\tilde{q}\frac{\partial b_n}{\partial s} \right]$ is today's MSCI to be used in (14f). Note that $\bar{\theta} - \bar{\mu} > 0$ holds.

Proof: Appendix 2 establishes this proposition as a special case of proposition 3.

The optimal carbon tax (14b′) is always set below its Pigouvian level, provided CO2-polluting goods are relatively consumed more by the poor than the rich (i.e., $\tilde{q} < 0$). Setting carbon taxes below the Pigouvian level is attractive, since carbon taxes are a more efficient instrument to reduce intra-generational equality than taxes on labour income. If consumption of dirty goods is mostly accounted for by the poor, the distributional characteristic δ^b is negative and the second part of (14b′) indicates that the government likes to subsidise CO2-polluting commodities for redistributive reasons (as $\delta > 0$ and $\bar{\theta} > \bar{\mu}$). The fact that it is optimal to set the carbon tax below the Pigouvian level implies from equation (14c) that it is also optimal to set a lower level of spending on CO2 abatement for given CO2 emissions.

We see from (14e′) that the optimal income tax rate now also depends directly on environmental considerations. The labour income tax rate is set higher (i.e., $t/(1-t) > \delta/\varepsilon$) both to redistribute from the rich to the poor in today's generations and to curb CO2 emissions. The regressive nature of the dirt tax thus requires an extra effort of the linear income tax to compensate for the adverse effects on inequality. As usual, the optimal income tax rate is inversely proportional to the wage elasticity of labour supply.

If demand for dirty goods is very elastic (high $-\overline{\mu}$ which implies from (A11) a high value of the price elasticity of the demand for dirty goods $-\overline{\varepsilon}_{bq} = [(1+\varepsilon/2)-\overline{\mu}]/2$), the Ramsey logic dictates that the optimal carbon tax must be set relatively low and consequently the labour income tax rate must be set relatively high. Note that if $\overline{b} = 0$, $\tilde{q} = 0$ and (14b′) and (14e′) reduce to (14b) and (14e), respectively.

Summing up, countries with a large share of their population being poor and consuming CO2-polluting goods cannot afford to have a carbon tax as aggressive as the Pigouvian tax as they have to give more attention to intra-generational redistribution objectives and as a result they implement a higher tax rate on labour income. This reminds one that it difficult to act morally if other basic necessities stand in the way. Indeed, 'Erst kommt das Fressen und dann die Moral' as in Bertold Brecht's *The Threepenny Opera*.

5. Policy insights from simulations

With homothetic preferences (14a)-(14f) together with the present-value budget constraint (8″) yield nine equations which can be solved for $(\gamma, q, t, T, s, S, g, G, \lambda)'$. $(b, v^{EDE}, w, a, \mathbf{E}[\Omega], \omega)'$ follow from (4), (5), (9), (11), (12′) and (13). With Stone-Geary preferences (14b) and (14e) are replaced by (14b′) and (14e′).

In the baseline, initial assets are zero ($\overline{a} = 0$), and population and productivity growth are zero ($\alpha = \beta = 0$). The return on productive assets and the pure rate of time preference are set to $r = \rho = 7.39$, which corresponds to 2 percent per year with a horizon of hundred years. Furthermore, in the baseline the coefficient of relative intra-generational inequality aversion is set to four ($\psi = 4$), the coefficient of relative inter-generational inequality aversion equals one-half ($\zeta = 0.5$). The baseline simulation assumes homothetic preferences, where subsistence needs are zero ($\overline{b} = 0$). We use the Galton distribution for productivities of individuals. The logarithm of individual productivity is thus normally distributed where $\ln(\overline{n})$ is the mean and $\varphi > 0$ the standard deviation of $\ln(n)$. Hence, φ is a measure of inequality in the productivities.[14] [15] We follow Stern (1976) by assuming that $\varphi = 0.4$. Mean labour productivity is calibrated output in each period and is normalised to 100 in the absence of government intervention and the other baseline values of the parameters. This yields $\mathbf{E}[\ln(\overline{n})] = 3.5451$. We model possible extreme worsening of the atmospheric CO2 concentration by assuming that κ takes on the value $1 - \pi K/(1 - \pi)$ (close to one) with probability $1 - \pi$ and $1 + K >> 1$ with probability π. The baseline has $\pi = 0.01$ and $K = 1$. Future climate uncertainty is thus characterised by a fat-tail distribution.

[14] The coefficient of variation for the Galton distribution is given by $\text{sd}(n)/\mathbf{E}[n] = \sqrt{\exp(\varphi^2)-1}$ and increases in φ.

[15] More realistic is to add a Pareto distribution for the upper tail of the distribution for pre-tax earnings.

Based on the studies in Blundell and MaCurdy (1999), who find that compensated wage elasticities are about 0.2 for men and 0.8 for women, the wage elasticity of labour supply is set at two-fifths ($\varepsilon = 0.4$). The elasticity of substitution between clean and dirty goods is calibrated so that the uncompensated price-elasticity of dirty goods demand equals 0.7. The (absolute value of) the uncompensated price elasticity of dirty goods demand is $\nu(1+\varepsilon)+(1-\nu)\sigma$, assuming zero non-labour incomes ($\theta = 1$). Setting the expenditure share of dirty commodities to $\nu = 0.25$, we find $\sigma = 0.5$. With no government intervention we find the following values for distribution-equivalent utility: $\mathrm{v}^{EDE}(0,0,0) = \mathrm{V}^{EDE}(0,0) = 10.31.$

We specify global warming damages as $x = \mathrm{e}(b,g) = \mathrm{h}(b) - \mathrm{f}(g),\ \ \mathrm{h}', \mathrm{h}'', \mathrm{h}''', \mathrm{f}\,' > 0, \mathrm{f}\,'' < 0,$ and allow marginal damages to be convex to allow for prudence. Future CO2 damages are $X = \mathrm{h}(\kappa b) - \mathrm{f}(g) - \mathrm{F}(G)$, $\mathrm{F}' > 0$, $\mathrm{F}'' < 0$. (14c) gives $\mathrm{h}'(b)/\mathrm{f}'(g) = q$ which implies that CO2 abatement g increases in the carbon tax but decreases in CO2 emissions. We express global warming damages in units of lost consumption:[16]
$\mathrm{h}(b) = [\eta_1 b^2 + \eta_2 b^{\eta_3}]\mathrm{v}^{EDE}(0,0,0),\ \ \mathrm{f}(g) = \eta_4 g^{\eta_5}\mathrm{v}^{EDE}(0,0,0),\ \ \mathrm{F}(G) = \eta_6 G^{\eta_7}\mathrm{V}^{EDE}(0,0),\ \ \eta_i > 0, \forall i,\ \ \eta_3 \geq 3,$
$\eta_5, \eta_7 < 1.$ For the benchmark we set $\eta_2 = 5.0\ 10^{-7}$ and $\eta_3 = 3$. We set the abatement and adaptation elasticities to $\eta_5 = \eta_7 = 0.3$. We calibrate $\eta_1 = 5.1\ 10^{-8}$ such that – with no government intervention – environmental damages are around 6 percent in normal circumstances and 2.5 percent of equivalently distributed utility under a climate disaster. We calibrate $\eta_4 = \eta_6 = 0.07$ so that in the absence of carbon and income taxes, the government would optimally spend 3.0 percent of output on abatement in the first period and 2.5 percent of output on abatement in the second period. The prudence factors (14a) become $\gamma^b \equiv \mathbf{E}\left[\kappa \mathrm{h}'(\kappa b)\right]/\mathrm{h}'(b) \geq 1$ and $\gamma^g = 1$ and for the baseline values we have $\gamma = 8.0 > 1$ which gives $\rho^* = 0$.

The first column in table 1 gives the baseline results. The other columns show how the baseline is affected by shocks in various parameters. Tougher substitution of dirty for clean goods (lower σ) makes it less distorting to have a higher carbon tax rate, so the income tax rates are a little lower while mean employment and output increase. Private welfare is higher, but emissions and global warming damages are higher also. Much easier substitution of dirty for clean goods (higher value of σ to greater than one) implies that the substitution effect dominates the income effect in consumption, so that compared with the inelastic baseline income tax rates and the carbon tax rate also increase while both consumption of clean and dirty goods increase. More inelastic labour supply (lower ε) makes it less distorting to have higher

---

[16] Nordhaus (2008) uses $v^{EDE}/(1+0.02388\ \text{Temperature}^2)$ and Weitzman (2010) $v^{EDE}/(1+0.02388\ \text{Temperature}^2 + 5.075\ 10^{-6}\ \text{Temperature}^{6.754})$. With this specification the certainty-equivalent, welfare-equivalent consumption level falls if a Pareto fat tail is added to the log-normal distribution of productivities, productivity growth falls, or the elasticity of intertemporal substitution falls (Weitzman, 2010).

income tax rates resulting in more inter- and intra-generational distribution tomorrow but the carbon tax and spending on abatement and adaptation are lower. Hence, there is a shift from inequality-corrected private welfare to green welfare.

Higher intra-generational inequality aversion (higher ψ) induces a shift from redistribution to climate change policies, hence income tax rates rise but the carbon tax, abatement and adaptation fall. Interestingly, mean employment, output and household income fall and thus consumption of both clean and dirty goods falls. Zero inter-generational inequality aversion ($\zeta = 0$) induces higher income taxes today and lower in the future, and a much lower carbon tax and much lower spending on abatement and adaptation. This leads to a bigger carbon debt passed on to future generations, but employment, output and EDE-utility rise both today and in the future. The main effect is thus more global warming.

A higher rate of pure time preference (ρ = 3%/annum) induces a lower income tax rate today and a higher future income tax rate, a higher carbon tax, less abatement and more adaptation, which gives rise to higher CO2 emissions and higher private and green welfare today and lower in the future. A higher return on productive assets (r = 3%/annum) makes it more attractive to save for future generations, so today income taxes rise and transfers fall. A smaller carbon debt is passed on despite a lower carbon tax and lower abatement. EDE-utility falls today and rises tomorrow but global warming damages rise.

Population growth ($\alpha > 0$) induces less redistribution today and more in the future (with a lower income tax today and a higher one in the future) and a more aggressive climate policy (with higher carbon tax, abatement and adaptation). As a result, consumption and EDE-utility rise today and fall in the future. Productivity growth ($\beta > 0$) implies that future generations are richer, so less productive assets and a bigger carbon debt are passed. Interestingly, this is achieved via a lower income tax today and a higher one in the future but via a higher carbon tax and higher spending on abatement to offset the increase in today's consumption and CO2 emissions due to the huge transfer-induced increase in today's income.

A bigger potential climate disaster (higher K) has a huge effect on the prudence factor, so future marginal global warming damages receive a much bigger weight. This leads to a more aggressive climate policy with a much higher carbon tax, higher abatement and higher adaptation. Inequality-corrected private welfare today is thus lower while in the future it is higher but CO2 emissions and global warming are substantially less. A higher probability of climate disaster (higher π) implies lower income tax rates today and in the future leading to bigger mean labour supply but lower mean output today and in the future and thus a smaller carbon debt despite less aggressive climate policies. Finally, more convex marginal damages of CO2 emissions (higher η_3) also induces a bigger prudence factor and thus heavier weighting

Table 1: Sensitivity of optimal climate change and redistribution policies

	Baseline	$\sigma = 0.2$	$\sigma = 1.2$	$\varepsilon = 0.2$	$\psi = 50$	$\zeta = 0$	$\rho =$ 3%/yr
Policy instruments							
Income tax rate period 1, *t*	0.468	0.459	0.439	0.573	0.631	0.476	0.455
Income tax rate period 2, *T*	0.462	0.450	0.418	0.569	0.630	0.409	0.575
Pollution tax rate period 1, *q*	0.258	0.377	1.194	0.112	0.188	0.079	0.300
Lump-sum transfer period 1, *s*	35.316	48.351	325.131	26.775	39.837	32.539	41.758
Lump-sum transfer period 2, *S*	40.307	58.432	491.536	28.528	43.650	78.011	3.473
CO2 abatement period 1, *g*	6.319	5.579	1.112	4.988	5.571	0.907	6.427
Climate adaptation period 2, *G*	5.380	4.750	0.946	4.247	4.743	0.772	1.612
Quantities							
Average clean consumption period 1, *c*	35.222	42.331	318.968	23.756	30.518	35.503	38.419
Average clean consumption period 2, *C*	41.143	55.885	409.139	25.533	34.277	62.953	16.828
Average dirty consumption period 1 = Pollution, b(*q*, *t*, *s*)	31.407	39.706	124.270	22.523	28.003	34.187	33.702
Average ‘dirty’ consumption period 2, *B*	34.586	45.743	171.896	23.563	30.102	59.475	13.796
Average labour supply period 1	1.798	2.183	11.531	1.296	1.573	1.845	1.803
Average labour supply period 2	1.894	2.353	13.615	1.312	1.631	1.967	1.724
Average output period 1	74.096	89.950	475.131	51.628	64.816	76.040	74.286
Average output period 2	78.033	96.939	560.992	52.290	67.218	81.028	71.014
Prices							
Real price index period 1, p(*q*)	4.501	2.820	0.046	4.222	4.367	4.156	4.579
Real price index period 2, p(0)	4.000	2.378	0.031	4.000	4.000	4.000	4.000
Shadow price GBC, λ	0.067	0.073	0.225	0.079	0.073	0.259	0.062
Welfare and inequality							
MSCI period 1,	1.071	1.064	1.052	1.051	1.048	1.078	1.062
MSCI period 2	1.067	1.058	1.041	1.049	1.047	1.038	1.284
Prudence factor, γ	7.992	7.993	7.998	7.988	7.991	7.992	7.992
EDE-utility period 1	12.747	26.918	10520.6	10.006	10.861	13.138	14.142
EDE-utility period 2	15.996	37.406	21923.9	11.101	12.914	26.655	4.291
Environmental damage period 1	-1.095	-0.885	9.156	-1.110	-1.095	-0.494	-1.063
Environmental damage period 2 disaster, X with high κ	-1.171	0.224	77.721	-1.810	-1.452	0.281	-0.513
Environmental damage period 2 normal, X with low κ	-2.295	-2.047	8.149	-2.225	-2.249	-1.168	-1.902
Social welfare	9.461	12.961	230.230	8.495	8.810	16.948	8.871

Table 1: Continued

	r =3%/yr	α=2%/yr	β=2%/yr	K = 4	π = 0.2	η_3 = 4
Policy variables						
Income tax rate period 1, t	0.491	0.441	0.327	0.447	0.288	0.438
Income tax rate period 2, T	0.303	0.493	0.558	0.424	0.281	0.401
Pollution tax rate period 1, q	0.175	0.388	1.441	1.146	0.247	2.042
Lump-sum transfer period 1, s	25.539	49.601	150.746	41.569	21.179	42.740
Lump-sum transfer period 2, S	231.728	25.524	128.202	65.811	25.104	85.218
CO2 abatement period 1, g	5.111	8.030	20.255	8.515	6.262	10.005
Climate adaptation period 2, G	16.235	6.837	17.245	7.250	5.331	8.518
Quantities						
Average clean consumption period 1, c	30.288	41.764	77.205	31.834	34.427	28.198
Average clean consumption period 2, C	145.997	32.076	377.458	56.030	39.974	67.000
Average dirty consumption period 1= Pollution, b(q, t, s)	27.933	35.448	49.416	21.736	30.832	16.167
Average ‘dirty’ consumption period 2, B	129.196	24.997	188.5784	31.0385	33.827	27.998
Average labour supply period 1	1.792	1.796	1.699	1.620	1.995	1.496
Average labour supply period 2	2.100	1.849	4.100	1.947	2.096	1.977
Average output period 1	73.846	74.009	69.986	66.747	72.595	61.649
Average output period 2	86.526	76.199	1417.193	80.223	76.265	81.455
Prices						
Real price index period 1, p(q)	4.344	4.745	6.566	6.075	4.480	7.530
Real price index period 2, p(0)	4.000	4.000	4.000	4.000	4.000	4.000
Shadow price GBC, λ	0.077	0.056	0.030	0.0541	0.067	0.048
Welfare and inequality						
MSCI period 1	1.092	1.0539	1.0160	1.0565	1.051	1.051
MSCI period 2	1.012	0.539	1.2183	1.044	1.047	1.035
Prudence factor, γ	7.991	7.993	7.9946	124.694	7.991	15.998
EDE-utility period 1	10.569	15.496	27.5189	10.331	12.232	8.344
EDE-utility period 2	67.648	5.708	106.606	23.263	15.227	28.642
Environmental damage period 1	-1.065	-1.118	-1.156	-1.3190	-1.100	-1.088
Environmental damage period 2 disaster, X with high κ	-1.942	-0.794	1.506	3.9432	-1.233	2.823
Environmental damage period 2 normal, X with low κ	-2.733	-2.410	-2.871	0.6103	-2.297	-2.474
Social welfare	9.852	14.055	14.447	8.8204	9.298	8.430

of expected future marginal global warming damages. This leads to substantially more aggressive climate policies with a much higher carbon tax and spending on abatement and adaptation and less intra-generational and inter-generational redistribution of incomes with lower income tax rates. Consumption of dirty goods falls while consumption of clean goods, labour supply, output and inequality-corrected private welfare fall today and rise in the future. The net result of this more prudent policy is that global warming damages in case of the bad state – climate disaster – are limited.

6. Concluding remarks

Although we have made bold assumptions (quasi-linear preferences, iso-elastic labour supply, constant elasticity of substitution between consumption of clean and dirty goods, no CO2 emissions in the future), we have been able to analyse intricate tradeoffs between the optimal redistribution of incomes within and between generations, the optimal carbon tax rate, and optimal spending on CO2 abatement and climate adaptation. Cautious discounting of future marginal global warming damages arises from future climatic uncertainty and convex marginal global warming damages. We have used equally-distributed-equivalent utility to value the social costs of inequality. We have also shown how non-homothetic utility and basic needs for dirty goods lowers the carbon tax below the modified Pigouvian rate and introduces an add-on term to the optimal labour income tax rate. Non-homothetic preferences thus imply that the carbon tax and income tax are no longer solely directed at global warming and redistribution, respectively, since fighting climate change must give way to income redistribution with basic needs.

Our most relevant policy messages can be summarised as follows. Firstly, the potential of bigger extreme climate change and a more prudent policy maker lower the rate by which future marginal global warming damages are discounted. Consequently, there will be a more aggressive environmental policy at the expense of less intensive intra-generational distribution of incomes. Secondly, a bigger aversion to intra-generational income inequality implies less ambitious climate policies and thus a higher income tax and a lower carbon tax and spending on CO2 abatement and climate adaptation. More inter-generational inequality aversion leads to a much higher carbon tax and spending on abatement and adaptation so that a lower carbon debt is passed on to future generations and global warming diminishes. However, income taxes today fall and rise in the future. Thirdly, a more elastic labour supply makes it more costly to have a high labour income tax rate and therefore it will be more difficult to pursue intra-and inter- generational income distribution leading to a shift from the green to the inequality-corrected private part of social welfare. Tougher substitution of dirty for clean goods makes it less distorting to have a higher carbon tax rate, but the effects depends on the relative importance of the substitution and income effects. Fourthly, higher impatience induces a lower income tax rate today and a higher future income tax rate, a higher carbon tax, less abatement and more adaptation, which gives rise to higher CO2 emissions and more welfare today than in the future. A higher return on productive assets makes it more attractive to save for future generations. Finally, population growth leads to less redistribution today and more in the future and a more aggressive climate change policy. Productivity growth implies that future generations are richer, so less productive assets and a bigger carbon debt are passed on to the future.

There are several important directions for future research. One way of having a more interesting trade-off between climate change and redistribution objectives is to allow for heterogeneous valuations of marginal damages from global warming. Much of the evidence cited in the Stern Review suggests that it is the poorest citizens on our globe that will be most adversely affected by global warming. With heterogeneous valuations of global warming damages one can allow for this. This implies that the poorest have a higher willingness to pay for an ambitious environmental policy and should therefore pay relatively the most, but the dilemma is that the poorest citizens on the planet are the least able to afford aggressive climate policies. It is therefore crucial to also address the political economy of climate policies. For example, if the income distribution is skewed so that the median income voter is poorer than the average income voter, the median voter will vote for relatively more weight on intra-generational redistribution of incomes and less on inter-generational distribution and climate policies. Finally, to understand the political economy of climate policies one needs to analyse tax reform as tax rates and spending on CO2 abatement and climate adaptation are unlikely to be set optimally in the real world.

References

Atkinson, A.B. (1970). On the measurement of economic inequality, Journal of Economic Theory, 2, 3, 244-263.

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

Bovenberg, A.L. and R.A. de Mooij (1994). Environmental levies and distortionary taxation, American Economic Review, 84, 4, 1085-1089.

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

Creedy, J. (2003). The optimal linear income tax model: utility or equivalent income?, Scottish Journal of Political Economy, 45, 1, 99-110.

Emmerling, J. (2010). Discounting and intragenerational inequity, mimeo., Toulouse School of Economics.

Gollier, C. and J. Gierlinger (2008). Social efficient discounting under ambiguity aversion, mimeo., Toulouse School of Economics.

Grafton, R.Q., T. Kompas and N. Van Long (2010). Biofules subsidies and the Green Paradox, Working Paper No. 2960, CESifo, Munich.

Hoel, M. (2008). Bush meets Hotelling: effects of improved improved renewable energy technology on greenhouse gas emissions, Working Paper No. 2492, CESifo, Munich.

Jacobs, B. and R.A. de Mooij (2010). Pigou meets Mirrlees: on the irrelevance of tax distortions for the second-best Pigouvian tax, mimeo., Erasmus University Rotterdam.

Kaplow, L. (2004). On the (ir)relevance of distribution and labor supply distortions to public goods provision and regulation, Journal of Economic Perspectives, 18, 4, 159-175.

Kaplow, L. (2006). Optimal control of externalities in the presence of income taxation, Working Paper No. 12339, NBER, Cambridge, Mass.

Kaplow, L. (2008). The Theory of Taxation and Public Economics, Princeton University Press, Princeton, N.J., and Oxford.

Laroque, G. (2005). Indirect taxation is superfluous under separability and taste homogeneity: A simple proof, Economics Letters, 87, 1, 141-144.

Nordhaus, W.D. (2007). A review of the Stern Review on the economics of climate change, Journal of Economic Literature, 45, 3, 686-702.

Nordhaus, W.D. (2008). A Question of Balance, Yale University Press, New Haven, Connecticut.

Ploeg, F. van der and C. Withagen (2010). Is there really a Green Paradox?, Research Paper 35, OxCarre, University of Oxford.

Ramsey, F.P. (1928). A mathematical theory of saving, Economic Journal, 38, 543-559.

Sandmo, A. (1975). Optimal taxation in the presence of externalities, Scandinavian Journal of Economics, 77, 1, 86-98.

Sinn, H.-W. (2008). Public policies against global warming, International Tax and Public Finance, 15, 4, 360-394.

Stern, N.H. (1976). On the specification of models of optimum income taxation, Journal of Public Economics, 6, 1-2, 123-162.

Stern, N.H. (2007). The Economics of Climate Change: The Stern Review, Cambridge University Press, Cambridge and New York.

Sterner, T. and U.M. Persson (2008). An even Sterner review: introducing relative prices into the discounting debate, Review of Environmental Economic Policy, 2, 1, 61-76.

Traeger, C. (2008). The social discount rate under intertemporal risk aversion and ambiguity, mimeo., Department of Agricultural and Resource Economics, UC Berkeley.

Weitzman, M. (1998). Why the far distant future should be discounted at its lowest possible rate, Journal of Environmental Economics and Management, 36, 201-208.

Weitzman, M. (2007). The Stern Review of the economics of climate change, Journal of Economic Literature, 45, 3, 703-724.

Weitzman, M. (2009). On modeling and interpreting the economics of catastrophic climate change, Review of Economics and Statistics, XCI, 1, 1-19.

Weitzman, M. (2010). GHG targets as insurance against catastrophic climate damages, Keynote Lecture WCERE 2010 Conference, Montreal.

Appendix: Optimal climate and redistribution policies for the general case

We first derive the optimal climate and redistribution policies for general functional forms for utility functions that are separable in the consumption basket of clean and dirty goods and labour supply.

Proposition 3: The optimal policies for utility functions $v_n\left(c_n, b_n, l_n\right) = u\left(c_n, b_n\right) - h\left(l_n\right)$ follow from:

(A1a) $$\frac{T}{1-T} = \frac{\Delta}{-\bar{\varepsilon}_{LT}}, \quad \delta = \frac{t}{1-t}\left(-\overline{\varepsilon_{lt}}\right) + \frac{\tilde{q}}{1+q}\left(-\overline{\varepsilon_{bt}}\right)$$

(A1b) $$q^P = \left[\mathrm{x}'\left(\mathrm{e}(b,g)\right) + \left(\frac{(1+\pi)\gamma}{1+\rho}\right)\mathrm{X}_e\left(\mathrm{e}(b,g), G\right)\right]\frac{\mathrm{e}_b(b,g)}{\lambda} \text{ and } \bar{\mu}\delta^b = \frac{t}{1-t}\left(-\overline{\varepsilon_{lq}}\right) + \frac{\tilde{q}}{1+q}\left(-\overline{\mu\varepsilon_{bq}}\right),$$

(A1c) $$-\mathrm{X}_G\left(\mathrm{e}(\kappa b, g), G\right) = \Lambda,$$

(A1d) $$\left(\frac{1+r}{1+\rho}\right)\mathbf{E}\left[\mathrm{X}_G\left(\mathrm{e}(b,g),G\right)\right] = \left[\mathrm{x}'\left(\mathrm{e}(b,g)\right) + \left(\frac{\gamma^g}{1+\rho}\right)\mathrm{X}_e\left(\mathrm{e}(b,g),G\right)\right]\mathrm{e}_g(b,g),$$

where $\bar{\varepsilon}_{LT} \equiv \mathbf{E}\left[\varepsilon_{LT} NL_N\right] / \mathrm{E}\left[NL_N\right]$, $\overline{\varepsilon_{lt}} \equiv \mathbf{E}\left[\varepsilon_{lt} nl_n\right] / \mathbf{E}\left[nl_n\right]$ and $\overline{\varepsilon_{bt}} \equiv \mathbf{E}\left[\varepsilon_{bt} nl_n\right] / \mathbf{E}\left[nl_n\right]$ are the income-weighted future labour supply, present labour supply and consumption of dirty goods with respect to 1 – *t*, $\overline{\varepsilon_{bq}} \equiv \mathbf{E}\left[\varepsilon_{bq} nl_n\right] / \mathbf{E}\left[nl_n\right]$ is the income-weighted elasticity of consumption of dirty goods with respect to 1 + *q*, $\bar{\mu} \equiv \mathbf{E}\left[\mu_n nl_n\right] / \mathbf{E}\left[nl_n\right]$ is the income-weighted share of $\mu_n \equiv \frac{(1+q)b_n}{(1-t)nl_n}$, $\Delta \equiv \frac{-\mathrm{cov}\left[NL_N, \eta_N\right]}{\mathbf{E}\left[NL_N\right]\mathbf{E}\left[\eta_N\right]}$ and

$\delta \equiv \frac{-\mathrm{cov}\left[nl_n, \eta_n\right]}{\mathbf{E}\left[nl_n\right]\mathbf{E}\left[\eta_n\right]}$ define the present and future distributional characteristics, respectively,

$\delta^b \equiv \frac{-\mathrm{cov}\left[b_n, \eta_n\right]}{\mathbf{E}\left[b_n\right]\mathbf{E}\left[\eta_n\right]}$ is the distributional characteristic for the dirty commodity, $\tilde{q} \equiv q - q^P$,

$\pi \equiv \frac{\mathrm{cov}\left[\kappa, \mathrm{e}_{\kappa b}(\kappa b, g)\mathrm{X}_e(\mathrm{e}(\kappa b, g), G)\right]}{\mathbf{E}[\mathrm{e}_{\kappa b}(\kappa b, g)\mathrm{X}_e(\mathrm{e}(\kappa b, g), G)]\mathbf{E}[\kappa]}$ and $\gamma \equiv \frac{\mathbf{E}\left[\mathrm{e}_{\kappa b}(\kappa b, g)\mathrm{X}_e(\mathrm{e}(\kappa b, g), G)\right]}{\mathrm{e}_b(b, g)\mathrm{X}_e(\mathrm{e}(b, g), G)} \geq 1.$

Proof: Households maximise utility subject to $c_n + (1+q)b_n = (1-t)nl_n + s$. This yields the first-order conditions $\frac{-v_l}{v_c} = \frac{h'}{u_c} = (1-t)n, \quad \frac{u_b}{u_c} = 1+q.$ Roy's lemma gives the partial derivatives of $\mathrm{v_n}(s,t,q)$, so $\frac{\partial v_n}{\partial s} = 1/p$, $\frac{\partial v_n}{\partial t} = -nl_n/p$ and $\frac{\partial v_n}{\partial q} = -b_n/p$ with $p \equiv \frac{c_n + (1+q)b_n}{\mathrm{u}(c_n, b_n)}$. The Slutsky-equations are:

(A2) $$\frac{\partial l_n}{\partial t} = \frac{\partial l_n^*}{\partial t} - nl_n \frac{\partial l_n}{\partial s}, \quad \frac{\partial b_n}{\partial t} = \frac{\partial b_n^*}{\partial t} - nl_n \frac{\partial b_n}{\partial s}, \quad \frac{\partial l_n}{\partial q} = \frac{\partial l_n^*}{\partial q} - b_n \frac{\partial l_n}{\partial s} \text{ and } \frac{\partial b_n}{\partial q} = \frac{\partial b_n^*}{\partial q} - b_n \frac{\partial b_n}{\partial s},$$

where asterisks denote a compensated change. Analogous expressions can be derived for period two. The Lagrangian for the future government is given by

$$\max_{\{S,T,G\}} \mathrm{L}(T,S,G,\Lambda) \equiv (1+\alpha)\frac{\mathrm{V}^{EDE}(T,S)^{1-\zeta}}{1-\zeta} - \mathrm{X}\left(\mathrm{e}(\kappa \mathrm{b}(q,t,s),g),G\right)$$
$$+\Lambda\left[(1+r)a + (1+\alpha)T\mathbf{E}\left[NL_N\right] - (1+\alpha)S - G\right],$$

where $V^{EDE} \equiv \Psi^{-1}\left(\mathbf{E}\left[\Psi\left(V_N(T,S)\right)\right]\right)$ and Λ is the Lagrange multiplier. Using $\frac{\partial V^{EDE}}{\partial V_N} = \frac{\Psi'(V_N)}{\Psi'(V^{EDE})}\mathrm{F}(N)$, we obtain the first-order conditions:

(A3a) $$\frac{\partial \mathrm{L}}{\partial S} = \mathbf{E}\left[\frac{(1+\alpha)(V^{EDE})^{-\zeta}}{\Psi'(V^{EDE})}\frac{\Psi'(V_n)}{\mathrm{p}(0)} - \Lambda(1+\alpha) + \Lambda(1+\alpha)TN\frac{\partial L_N}{\partial S}\right] = 0,$$

(A3b) $$\frac{\partial \mathrm{L}}{\partial T} = \mathbf{E}\left[-\frac{(1+\alpha)(V^{EDE})^{-\zeta}}{\Psi'(V^{EDE})}\frac{\Psi'(V_n)}{\mathrm{p}(0)}NL_N + \Lambda(1+\alpha)NL_N + \Lambda(1+\alpha)NT\frac{\partial L_n}{\partial T}\right] = 0,$$

(A3c) $$\frac{\partial \mathrm{L}}{\partial G} = -\mathrm{X}_G\left(\mathrm{e}(\kappa b, g), G\right) - \Lambda = 0,$$

where we have used Roy's lemma in (A3a) and (A3b). Define the social marginal value of one unit of income including the income effects on the tax base as $\eta_N \equiv \frac{(V^{EDE})^{-\xi}}{\Psi'(V^{EDE})}\frac{\Psi'(V_N)}{\mathrm{p}(0)} + \Lambda TN\frac{\partial L_N}{\partial S}$. Thus, we find from (A3a) that $\mathbf{E}\left[\eta_N\right] = \Lambda.$ The first-order condition for T, (A3b), is rewritten by substituting the Slutsky equation for labour supply, the definitions for η_N and Δ to find the first part of (A1a). Equation (A3c) implies that G follows from (A1c). Optimal policies are written as $\mathrm{T}\left(a, \mathrm{e}(\kappa b, g)\right)$, $\mathrm{S}\left(a, \mathrm{e}(\kappa b, g)\right)$ and $\mathrm{G}\left(a, \mathrm{e}(\kappa b, g)\right)$, and future welfare as

(A4) $$\Omega(a,e) \equiv (1+\alpha)\frac{\mathrm{V}^{EDE}\left(\mathrm{T}\left(a,\mathrm{e}(\kappa b,g)\right), \mathrm{S}\left(a,\mathrm{e}(\kappa b,g)\right)\right)^{1-\zeta}}{1-\zeta} - \mathrm{X}\left(\mathrm{e}(\kappa \mathrm{b}(q,t,s),g), \mathrm{G}\left(a,\mathrm{e}(\kappa b,g)\right)\right)$$

with $\Omega_a\left(a,\mathrm{e}(\kappa b,g)\right)=\Lambda\left(1+r\right)$ and $\Omega_e\left(a,\mathrm{e}(\kappa b,g)\right)=-\mathrm{X}_e\left(\mathrm{e}(\kappa b,g),G\right)$. From the Envelope Theorem we know that a marginal change in the initial conditions does not affect welfare through a resulting change in $\mathrm{T}\left(a,\mathrm{e}(\kappa b,g)\right)$, $\mathrm{S}\left(a,\mathrm{e}(\kappa b,g)\right)$ and $\mathrm{G}\left(a,\mathrm{e}(\kappa b,g)\right)$, since future policies are set optimally.

Let $v^{EDE}\equiv\Psi^{-1}\left(\mathrm{E}\left[\Psi\left(v_n(t,s)\right)\right]\right)$, so that the Lagrangian for today's maximisation problem is defined by

(A5) $$\max_{\{s,t,q,g,a\}}\mathrm{L}(s,t,q,g,a,\lambda)\equiv\frac{\mathrm{v}^{EDE}\left(s,t,q\right)^{1-\zeta}}{1-\zeta}-\mathrm{x}\left(\mathrm{e}(\mathrm{b}(q,t,s),g)\right)+\frac{\mathbf{E}\left[\Omega\left(a,\mathrm{e}(\kappa\mathrm{b}(q,t,s),g)\right)\right]}{1+\rho}$$
$$+\lambda\left[\left(\bar{a}+\mathbf{E}\left[\left(tnl_n+qb_n\right)\right]-s-g-a\right)\right],$$

where λ denotes the Lagrange multiplier. First-order conditions for an optimum are given by:

(A6a) $$\frac{\partial\mathrm{L}}{\partial s}=\mathbf{E}\left[\frac{\left(v^{EDE}\right)^{-\zeta}}{\Psi'\left(v^{EDE}\right)}\frac{\Psi'\left(v_n\right)}{p}-\lambda+\lambda tn\frac{\partial l_n}{\partial s}+\lambda\tilde{q}\frac{\partial b_n}{\partial s}\right]=0,$$

(A6b) $$\frac{\partial\mathrm{L}}{\partial t}=\mathbf{E}\left[-\frac{\left(v^{EDE}\right)^{-\zeta}}{\Psi'\left(v^{EDE}\right)}\frac{\Psi'\left(v_n\right)}{p}nl_n+\lambda nl_n+\lambda nt\frac{\partial l_n}{\partial t}+\lambda\tilde{q}\frac{\partial b_n}{\partial t}\right]=0,$$

(A6c) $$\frac{\partial\mathrm{L}}{\partial q}=\mathbf{E}\left[-\frac{\left(v^{EDE}\right)^{-\zeta}}{\Psi'\left(v^{EDE}\right)}\frac{\Psi'\left(v_n\right)}{p}b_n+\lambda b_n+\lambda tn\frac{\partial l_n}{\partial q}+\lambda\tilde{q}\frac{\partial b_n}{\partial q}\right]=0,$$

(A6d) $$\frac{\partial\mathrm{L}}{\partial g}=-\left[\mathrm{x}'(\mathrm{e}(b,g))+\left(\frac{\gamma^g}{1+\rho}\right)\mathrm{X}_e\left(\mathrm{e}(b,g),G\right)\right]\mathrm{e}_g(b,g)-\lambda=0,$$

(A6e) $$\frac{\partial\mathrm{L}}{\partial a}=\left(\frac{1+r}{1+\rho}\right)\mathbf{E}\left[\Lambda\right]-\lambda=0,$$

where we substituted Ω_a and Ω_e in all conditions, Roy's lemma in A(6a)-(A6c), and $\mathrm{b}_s=\mathbf{E}\left[\frac{\partial b_n}{\partial s}\right]$, $\mathrm{b}_t=\mathbf{E}\left[\frac{\partial b_n}{\partial t}\right]$ and $\mathrm{b}_q=\mathbf{E}\left[\frac{\partial b_n}{\partial q}\right]$ in (A6a)-(A6c), respectively, and we have used $\tilde{q}\equiv q-q^P$ with

$$q^P\equiv\left[\mathrm{x}'(e(b,g))+\left(\frac{1}{1+\rho}\right)\frac{\mathbf{E}\left[\kappa\mathrm{e}_{\kappa b}(\kappa b,g)\mathrm{X}_e(\mathrm{e}(\kappa b,g),G)\right]}{\mathrm{e}_b(b,g)}\right]\frac{e_b}{\lambda}=\left[\mathrm{x}'(e(b,g))+\left(\frac{\gamma^b}{1+\rho}\right)\mathrm{X}_e(\mathrm{e}(b,g),G)\right]\frac{e_b}{\lambda},$$

where $\gamma^g\equiv\frac{\mathbf{E}\left[-\mathrm{e}_g(\kappa b,g)\mathrm{X}_e(\mathrm{e}(\kappa b,g),G)\right]}{-\mathrm{e}_g(b,g)\mathrm{X}_e(\mathrm{e}(b,g),G)}\geq1$ and $\gamma^b\equiv\frac{\mathbf{E}\left[\kappa\mathrm{e}_{\kappa b}(\kappa b,g)\mathrm{X}_e(\mathrm{e}(\kappa b,g),G)\right]}{\mathrm{e}_b(b,g)\mathrm{X}_e(\mathrm{e}(b,g),G)}$. Using E[$\kappa x$] = cov[$\kappa$,$x$] + E[$x$]E[$\kappa$], we can write $\gamma^b=\gamma(1+\pi)$ and thus we can rewrite q^P as (A1b).

Define $\eta_n\equiv\frac{\left(v^{EDE}\right)^{-\xi}}{\Psi'\left(v^{EDE}\right)}\frac{\Psi'\left(v_n\right)}{p}+\lambda tn\frac{\partial l_n}{\partial s}+\lambda\tilde{q}\frac{\partial b_n}{\partial s}$. From (A6a) we get $\mathrm{E}\left[\eta_n\right]=\lambda$. After substituting the Slutsky equations, the definitions for η_n, δ, and $\mu_n\equiv\frac{(1+q)b_n}{(1-t)nl_n}$ into equation (A6b), we obtain the second part of (A1a). After substituting the Slutsky equations, the definitions for η_n and μ_n into equation (A6c), we get $\mathbf{E}\left[(1-\eta_n)\frac{(1+q)b_n}{(1-t)} - \frac{t}{1-t} nl_n\varepsilon_{lq} - \frac{\tilde{q}}{1+q}\mu_n nl_n \varepsilon_{bq} \right] = 0.$ It thus follows that the optimal pollution tax is determined by the second part of (A1b). Equation (A6e) yields $\left(\frac{1+r}{1+\rho}\right)\mathbf{E}\left[\Lambda\right] = \lambda.$ Using equation (A1), equations (A6d) and (A6e) give rise to:

(A8) $$-\left[\mathrm{x}_e(\mathrm{e}(b,g)) + \left(\frac{\gamma^g}{1+\rho}\right)\mathrm{X}_e\left(\mathrm{e}(b,g),G\right)\right]\mathrm{e}_g(b,g) = \lambda \text{ and } \left(\frac{1+r}{1+\rho}\right)\mathbf{E}\left[\Lambda\right] = \lambda.$$

Using $-\mathrm{X}_G\left(\mathrm{e}(b,g),G\right) = \Lambda$ and the equations of (A8), we obtain (A1d). Q.E.D.

Case: Homothetic preferences (proposition 1)

Expenditure shares are independent of ability: $\nu \equiv \frac{(1+q)b_n}{(1-t)nl_n + s} = \frac{1}{1+(1+q)^{\sigma-1}}$ and $1-\nu \equiv \frac{c_n}{(1-t)nl_n + s} = \frac{1}{1+(1+q)^{1-\sigma}}.$ Marginal utility of income equals $\eta \equiv \frac{\partial v_n}{\partial c_n} = 1/\mathrm{p}(q).$ To obtain the elasticities needed for proposition 3, we linearise the model around an initial equilibrium and use tilde's (~) to denote relative changes, e.g., $\tilde{x} \equiv \mathrm{d}x/x$, except for the tax rates, i.e., $\tilde{t} \equiv dt/(1-t)$ and $\tilde{q} \equiv \mathrm{d}q/(1+q)$. Linearising the utility function at constant utility ($\mathrm{d}v_n = 0$) yields $(1-\nu)\tilde{c}_n + \nu\tilde{b}_n = \tilde{l}_n$ with $\nu \equiv \frac{(1+q)b_n}{(1-t)nl_n + s}.$ Linearising the first-order condition for labour supply gives $\tilde{l}_n = \varepsilon(-\tilde{t} + \tilde{u}_c).$ Using that u_n is linear homogeneous in c_n and b_n (i.e., $u_{cc}c_n = -u_{cb}b_n$), we get $\tilde{u}_c = \frac{\nu}{\sigma}(\tilde{b}_n - \tilde{c}_n).$ Use $u_b b_n / u = \nu$ so that $\tilde{u}_c = -\nu\tilde{q}.$ Linearising the marginal rate of substitution for c_n and b_n yields $\tilde{b}_n = -\sigma\tilde{q} + \tilde{c}_n$, where $\sigma \equiv \left(\frac{\mathrm{u}_{cb}(c_n,b_n)u_n}{\mathrm{u}_b(c_n,b_n)\mathrm{u}_c(c_n,b_n)}\right)^{-1} > 0$ denotes the constant elasticity of substitution in u_n.

Thus, relative changes in labour supply and dirty goods consumption are $\tilde{l}_n = -\varepsilon(\tilde{t} + \nu\tilde{q})$ and $\tilde{b}_n = -\varepsilon\tilde{t} - (\nu(\varepsilon - \sigma) + \sigma)\tilde{q}.$ The elasticities are thus constant across individuals:

(A9) $$\varepsilon_{lt} \equiv \frac{\partial l_n}{\partial t}\frac{1-t}{l_n} = -\varepsilon < 0,\ \ \varepsilon_{lq} \equiv \frac{\partial l_n}{\partial q}\frac{1+q}{l_n} = -\varepsilon\nu < 0,\ \ \varepsilon_{bt} \equiv \frac{\partial b_n}{\partial t}\frac{1-t}{b_n} = -\varepsilon < 0$$
$$\text{and } \varepsilon_{bq} \equiv \frac{\partial b_n}{\partial q}\frac{1+q}{b_n} = -\left(\nu\varepsilon + (1-\nu)\sigma\right) < 0.$$

There are no income effects in labour supply, $\frac{\partial L_N}{\partial S} = 0,$ and the income-weighted compensated labour supply elasticity is constant, $-\bar{\varepsilon}_{lt} = \varepsilon$, so we confirm the second part of (14e). From (A3a) we get $\Upsilon \equiv \frac{E[\Psi'(V_N)]}{\Psi'(V^{EDE})} = \mathrm{p}(0)\Lambda\left(V^{EDE}\right)^{\xi}$. With $\Delta = \frac{-\mathrm{cov}[NL_N, \Psi'(V_N)]}{\mathbf{E}[NL_N]\mathbf{E}[\Psi'(V_N)]}$, equation (A1c) then reduces to (10b) or $-\mathrm{X}_G(\mathrm{e}(\kappa b, g), G) = \Lambda = \Upsilon\left(V^{EDE}\right)^{-\xi} / \mathrm{p}(0)$. We find the optimal income tax after substitution of the elasticities (A9) into (A1a) to get $\delta = \frac{t}{1-t}\varepsilon + \frac{\tilde{q}}{1+q}\varepsilon\bar{\mu}$. We can use $b_n = \nu\frac{(1-t)nl_n + s}{1+q}$ and (A6a) to get $\mathbf{E}\left[(1-\eta_n)\frac{(1+q)b_n}{1-t}\right] = \mathbf{E}\left[(1-\eta_n)\nu nl_n\right]$. Substitution in (A1b) and using the elasticities (A10) yields $\delta = \frac{t}{1-t}\varepsilon + \frac{\tilde{q}}{1+q}\frac{\bar{\mu}}{\nu}\left(\nu\varepsilon + (1-\nu)\sigma\right)$. We solve (A1a) and (A1b) to find $\tilde{q} = 0,$ hence $q = q^P$ as given in the first part of (A1b) and $\frac{t}{1-t} = \frac{\delta}{\varepsilon}$. This establishes (14b) and the first part of (14e). Income effects in labour supply are absent ($\frac{\partial l_n}{\partial s} = 0$) and the pollution tax is set at the Pigouvian level ($\tilde{q} = 0$), so $\eta_n \equiv \frac{\left(v^{EDE}\right)^{-\zeta}}{\Psi'(v^{EDE})}\frac{\Psi'(v_n)}{p}$ and $\delta \equiv \frac{-\mathrm{cov}[nl_n, \Psi'(v_n)]}{\mathbf{E}[nl_n]\mathbf{E}[\Psi'(v_n)]}$. Consequently, $\upsilon \equiv \frac{\mathbf{E}[\Psi'(v_n)]}{\Psi'(v^{EDE})} = \lambda p\left(v^{EDE}\right)^{\xi}$ and we obtain $\left(\frac{1+r}{1+\rho}\right)\frac{\mathbf{E}[(V^{EDE})^{-\zeta}\Upsilon]}{\mathrm{p}(0)} = \frac{\left(v^{EDE}\right)^{-\zeta}\upsilon}{\mathrm{p}(q)}$ or (14f). Equation (14d) follows from rewriting (A1d). The second part of (14d) follows from (10b) and (14f). Finally, (14c) follows directly from (14b) and (14d). This establishes proposition 1.

Case: Stone-Geary preferences (proposition 2)

With Stone –Geary preferences (15), we have real price index $p \equiv \mathbf{p}(q) = \sqrt{(1+q)}$ and (16)-(17). The expenditure shares are not constant, $\nu_n \equiv \frac{(1+q)b_n}{(1-t)nl_n + s} = \frac{1}{2}\left(1 + \frac{(1+q)\bar{b}}{(1-t)nl_n + s}\right)$, but marginal utility of income $\eta \equiv \frac{\partial v_n}{\partial c_n} = 1/\mathrm{p}(q)$ is constant. To obtain the elasticities, we proceed as before. The linearised utility function is given by $(1-\nu_n)\tilde{c}_n + \nu_n\tilde{b}_n = \tilde{l}_n$. The linearised first-order condition for labour supply is $\tilde{l}_n = -\varepsilon\tilde{t} - \frac{1}{2}\varepsilon\tilde{q}$. The linearised marginal rate of substitution for c_n and b_n is $\tilde{c}_n = \tilde{q} + \left(\frac{b_n}{b_n - \bar{b}}\right)\tilde{b}_n$. Thus, we obtain $\tilde{b}_n = -\frac{\varepsilon}{(1-\nu_n)\left(\frac{b_n}{b_n-\bar{b}}\right)+\nu_n}\tilde{t} - \frac{\frac{1}{2}\varepsilon + (1-\nu_n)}{(1-\nu_n)\left(\frac{b_n}{b_n-\bar{b}}\right)+\nu_n}\tilde{q}$ and the elasticities:

(A10) $$\varepsilon_{lt} \equiv \frac{\partial l_n}{\partial t}\frac{1-t}{l_n} = -\varepsilon < 0, \quad \varepsilon_{lq} \equiv \frac{\partial l_n}{\partial q}\frac{1+q}{l_n} = -\frac{1}{2}\varepsilon < 0,$$
$$\varepsilon_{bt} \equiv \frac{\partial b_n}{\partial t}\frac{1-t}{b_n} = -\frac{\varepsilon}{(1-\nu_n)\left(\frac{b_n}{b_n-\bar{b}}\right)+\nu_n} < 0 \text{ and } \varepsilon_{bq} \equiv \frac{\partial b_n}{\partial q}\frac{1+q}{b_n} = -\frac{\frac{1}{2}\varepsilon+(1-\nu_n)}{(1-\nu_n)\left(\frac{b_n}{b_n-\bar{b}}\right)+\nu_n} < 0.$$

Note that the consumption elasticities depend on ability. (If $\bar{b}=0,$ we have the same elasticities as with homothetic preferences if $\nu = 1/2$ and $\sigma = 1$.) The income-weighted elasticities are:

(A11) $$\overline{\varepsilon_{bt}} = -\frac{\mathbf{E}\left[\frac{(1+q)b_n}{(1-t)nl_n}\left(\frac{\varepsilon}{(1-\nu_n)\mu_n+\nu_n}\right)nl_n\right]}{\mathbf{E}[nl_n]} = -\frac{1}{2}\varepsilon\bar{\theta} < 0,$$
$$\overline{\varepsilon_{bq}} = -\frac{\mathbf{E}\left[\frac{(1+q)b_n}{(1-t)nl_n}\left(\frac{\frac{1}{2}\varepsilon+1-\nu_n}{(1-\nu_n)\mu_n+\nu_n}\right)nl_n\right]}{\mathbf{E}[nl_n]} = -\frac{\left(1+\frac{1}{2}\varepsilon\right)\bar{\theta}}{2}+\frac{\bar{\mu}}{2} < 0,$$

where $\mu_n \equiv \frac{(1+q)b_n}{(1-t)nl_n}$ and $\theta_n \equiv \frac{(1-t)nl_n+s}{(1-t)nl_n} > 1$. Again, $\frac{\partial L_N}{\partial S}=0$ and $-\bar{\varepsilon}_{lt}=\varepsilon$, so we get (10a). Again, from $\Upsilon \equiv \frac{E[\Psi'(V_N)]}{\Psi'(V^{EDE})} = \mathrm{p}(0)\Lambda\left(V^{EDE}\right)^{\xi}$ and $\Delta = \frac{-\operatorname{cov}[NL_N,\Psi'(V_N)]}{\mathbf{E}[NL_N]\mathbf{E}[\Psi'(V_N)]}$ and (A1c), we get (10b). After substitution of the elasticities (A10) and (A11) into (A1a), we get $\delta = \frac{t}{1-t}\varepsilon + \frac{\tilde{q}}{1+q}\varepsilon\frac{1}{2}\bar{\theta}.$ Substituting (A10) and (A11) in (A1b) gives the expression $2\bar{\mu}\delta^b = \frac{t}{1-t}\varepsilon + \frac{\tilde{q}}{1+q}\left(\left(1+\frac{1}{2}\varepsilon\right)\bar{\theta}-\bar{\mu}\right).$ We can use two expressions to solve for optimal income and pollution taxes, that is

(A12) $$\frac{t}{1-t} = \frac{\delta}{\varepsilon} + \frac{(\delta - 2\delta_b\bar{\mu})\bar{\theta}}{2(\bar{\theta}-\bar{\mu})} \text{ and } \frac{\tilde{q}}{1+q} = \frac{2\bar{\mu}\delta^b - \delta}{\bar{\theta}-\bar{\mu}}.$$

These expressions correspond to (14′) and (14e′), respectively. Note that $\bar{\theta} = -2\bar{\varepsilon}_{bt} > 0$ and $\bar{\mu} = 2\bar{\varepsilon}_{bq} + (2+\varepsilon)\bar{\varepsilon}_{bt}/\varepsilon < 0,$ so that $\bar{\theta}-\bar{\mu} = -2\bar{\varepsilon}_{bq} - \bar{\varepsilon}_{bt} > 0.$ Income effects in labour supply are absent ($\frac{\partial l_n}{\partial s}=0$). However, $\tilde{q} \neq 0$ so we get $\eta_n \equiv \frac{\left(v^{EDE}\right)^{-\xi}}{\Psi'\left(v^{EDE}\right)}\frac{\Psi'(v_n)}{p} + \lambda\tilde{q}\frac{\partial b_n}{\partial s}.$ Hence, we obtain:

(A13) $$\delta = \frac{-\operatorname{cov}\left[nl_n, \frac{\left(v^{EDE}\right)^{-\xi}\Psi'(v_n)}{\Psi'\left(v^{EDE}\right)} + p\lambda\tilde{q}\frac{\partial b_n}{\partial s}\right]}{\mathbf{E}[nl_n]\mathbf{E}\left[\frac{\left(v^{EDE}\right)^{-\xi}\Psi'(v_n)}{\Psi'\left(v^{EDE}\right)} + p\lambda\tilde{q}\frac{\partial b_n}{\partial s}\right]} \text{ and } \delta^b = \frac{-\operatorname{cov}\left[b_n, \frac{\left(v^{EDE}\right)^{-\xi}}{\Psi'\left(v^{EDE}\right)}\Psi' + p\lambda\tilde{q}\frac{\partial b_n}{\partial s}\right]}{\mathbf{E}[b_n]\mathbf{E}\left[\frac{\left(v^{EDE}\right)^{-\xi}\Psi'(v_n)}{\Psi'\left(v^{EDE}\right)} + p\lambda\tilde{q}\frac{\partial b_n}{\partial s}\right]}.$$

We redefine $\upsilon \equiv \mathbf{E}\left[\frac{\Psi'(v_n)}{\Psi'\left(v^{EDE}\right)} + p\lambda\left(v^{EDE}\right)^{\xi}\tilde{q}\frac{\partial b_n}{\partial s}\right] = p\lambda\left(v^{EDE}\right)^{\xi}$ and thus obtain (14f). This establishes proposition 2.