Optimal taxation, environment quality, socially responsible firms and investors

by Thomas I. Renström*, Luca Spataro‡, Laura Marsiliani**

Abstract

We characterize the optimal pollution-, capital- and labour-tax structure in a continuous-time model in the presence of pollution (resulting from production), both in the first- and second-best, allowing investors to be driven by social responsibility objectives. The social responsibility objective takes the form of warm-glow, as in Andreoni (1990) and Dam (2011), inducing firms to reduce pollution through increased abatement activity. Among the results, the second-best pollution tax displays an additivity property and the Chamley-Judd zero capital-income tax can be violated under warm-glow preferences. We also show that first- and second-best pollution taxes are positive, under warm-glow preferences, and, under mild assumptions, the latter yield lower first-best pollution taxes and lower pollution intensity.

JEL Classification: D21, D53, G11, H21, H23, M14, Q58.

Keywords: Socially responsible investment, corporate social responsibility, environmental quality, optimal taxation, pollution.

* Durham University Business School, Durham (UK). Email: t.i.renstrom@durham.ac.uk.

‡Corresponding author. Dipartimento di Economia e Management, University of Pisa (Italy). Email: luca.spataro@unipi.it, tel: +39 050 2216217

** Durham University Business School, Durham (UK). Email: laura.marsiliani@durham.ac.uk.

1. Introduction

In the present paper we ask how optimal pollution taxes, as well as capital and labour taxes are affected by socially responsible objectives of investors, both in the first- and second-best. More precisely, to the extent to which socially responsible investors manage to induce firms to reduce pollution, what is then the structure of optimal taxes and, in particular, pollution taxes?

The issue of environmental quality has been increasingly debated in the last decades. Many international summits (from Kyoto in 1997 to Paris 2015), the diffusion of credit rating agencies, shareholder activism, mobilization of NGOs and social media prove this growth of environmental concerns (Ballestrero et al. 2015). Socially responsible investment (SRI) has been argued to be a possible instrument to improve environmental quality through a market mechanism.

According to Eurosif, SRI "is a long-term oriented investment approach, which integrates ESG [i.e. Environmental, Social, Governance] factors in the research, analysis and selection process of securities within an investment portfolio. It combines fundamental analysis and engagement with an evaluation of ESG factors in order to better capture long term returns for investors, and to benefit society by influencing the behaviour of companies." (Eurosif 2016, p. 9). Hence, SRI is a process of identifying and investing in companies that meet certain standards of Corporate Social Responsibility (CSR) [1] through such activities and strategies as positive or negative screening, shareholder advocacy, impact and community investing (for more details see GSIA, 2016).

As a matter of fact, according to the latest Global Sustainable Investment Alliance report (GSIA, 2016) global SRI assets amounted to $22.89 trillion at the start of 2016 (an increase of

[1] For a recent review of economic literature on CSR, see Brekke and Pekovic (2018).

25% since 2014), and represented 26% of total assets managed in the world, with $8.72 trillion in total assets managed in the US (10% in 2001), $12.04 trillion in Europe, $6.7 billion in Canada, $515.7 billion in Australia and New Zealand, $52.1 billion in Asia.

In light of this recent trend and for its wide potential implications concerning, for example, the design of pollution abatement policies and of fiscal incentives for green initiatives[2], several scholars have started analysing the phenomenon from an economic perspective. However, while the literature on pollution taxes has been flourishing[3], the economic literature on SRI and on its consequences on taxation is still embryonic and results are mixed.

As for the existing economic literature on SRI[4], Hainkel et al. (2001), in a one-period model, show that negative screening on polluting firms by fund managers can induce these firms to adopt cleaner technologies, in that otherwise they would incur higher costs of capital. The positive effects of financial markets are also stressed by Dam (2011), who argues that SRI creates a role for the stock market to deal with intergenerational environmental externalities, consisting in the fact that short-lived individuals fail to account for the long-term effects of pollution. The author shows that, although socially responsible investors are short-lived, the

[2] For example, in 1995 the Dutch government has launched the Green Funds Scheme, a tax incentive scheme for investors into green initiatives. In the U.S. there are examples of tax-credit bonds, (bond investors receive tax credits instead of interest payments so issuers do not have to pay interest on their green bond issuances) or tax-exempt bonds. More in general, developed countries have committed under the international negotiations to mobilising USD 100 billion per year by 2020 for climate change mitigation and adaptation in developing countries (see OECD 2015). However, how to count and track climate finance is still an open question.

[3] The seminal work is Sandmo (1975). See also Cremer et al. (2001) and the survey by Bovenberg and Goulder (2002). More recent works on this subject are Bontems and Bourgeon (2005), Goulder and Parry (2008), Gahvari (2014), Jacobs and De Mooij (2015), Kampas and Horan (2016), Belfiori (2018), Aronsson and Sjögren, (2018).

[4] For a survey on the topic, see Renneboog et al. (2008).

forward-looking nature of stock prices, reflecting the warm-glow motive, can help to mitigate the conflict between current and future generations.

Dam and Scholtens (2015) develop a model that links SRI and CSR, showing that responsible firms have higher returns on assets, although the overall effect on stock market returns depends on the relative strength of supply and demand side effects. On the other hand, Barnea et al. (2005) argue that negative screening reduces the incentives of polluting firms to invest, but also the total level of investment in the economy. Dam and Heijdra (2011) analyse the effects of SRI and public abatement on environmental quality in a growth model where investors feel partly responsible for environmental pollution when holding firm equity (due to a warm-glow mechanism as in Andreoni 1990). In this scenario, the authors show that SRI behaviour by households partially offsets the positive effects on environmental quality of public abatement policies.

Finally, according to Vanwalleghem (2017), SRI may have a mixed effect on firms' incentives to remove negative externalities. Whereas SRI screening incentivizes the removal of externalities (as predicted by Heinkel et al. 2001 and confirmed by the empirical work of Hong and Kacperczyk 2009), SRI trading can disincentivise it when traders disagree on the externality removal's cash flow effects.[5]

While providing interesting results, the above-mentioned literature has not analysed the optimal structure of taxes in presence of both socially responsible investors and firms.

Hence, in this paper we specify a continuous-time model, where pollution is a by-product of production, but firms can engage in abatement, reducing net pollution. We model investors' social-responsibility objective through a warm-glow mechanism as in Andreoni

[5] Other works focusing on socially responsible firms and financial markets are Graff Zivin and Small (2005) and Baron (2007). However, these are partial-equilibrium and static models, in which social responsibility is concerned with charitable giving and not with abatement of externalities or public bads.

(1990) and Dam (2011). Through investors' portfolio choice, firms are induced to engage in socially responsible activities (abatement). By allowing for different specifications of the warm-glow function[6], we show the circumstances under which the well-established zero capital-income tax result, Chamley (1986) and Judd (1985) can be violated.

In this paper we show that pollution tax is positive for any specification of the warm-glow and, in the second-best, displays an additive property.

Taxes on production-factor incomes, in the first-best, are either zero or negative (negative taxation arising in case the perceived pollution content of firms is negatively related to the total scale of economic activity). Hence, in the latter case a violation of the zero-capital income tax result occurs. Finally, for illustrative purposes we provide an example with specific utility function, to show that that the presence of warm-glow implies lower pollution taxes, lower pollution intensity and lower installed capital in comparison to an economy without warm-glow preferences.

[6] While there is increasing evidence of the very existence of warm-glow preferences (see Andreoni et al. 2017), the exact shape is far from being clear. Although some recent works have produced axiomatizations of the warm glow that can help to characterize its shape (see Evren and Minardi 2017 and the literature therein), we believe that the argument of Bernehim and Rangel (2005) is still valid, as they write: "Unfortunately, in the context of warm glow giving and public goods, these processes are not yet well understood. The warm glow model remains a "black box", and one can interpret it as a reduced form for a variety of mechanisms with starkly differing welfare implications" (p. 63). Moreover, Diamond (2006), while recognizing that warm glow may improve our description of how individuals behave, argues for ignoring the effects of warm–glow utility on social welfare when considering optimal tax structures. The reason being that warm glow amounts to preferences over how the public good is produced (see also Andreoni 2006). In the light of these issues, in this paper we undertake the following choices: given that in our model the presence of warm-glow has testable implications, in that it produces an extra-cost of capital to polluting firms that could be empirically estimated, we analyse optimal taxes under different specifications of the warm-glow, leaving the choice of the exact shape to future research. Second, following Allgood (2009), in the last section we also provide results for the case in which the warm-glow is set arbitrarily close to zero.

The paper is organized as follows: in section 2 we specify the model and characterize the decentralized equilibrium; in section 3 we present the Ramsey problem of optimal taxation and provide the solutions; in section 4 we discuss the results and in section 5 we illustrate the role of the warm-glow in affecting the first-best pollution tax. Section 5 concludes.

2. The model setup

In this section, we specify the benchmark model. The model contains H identical households and J identical firms. We assume that an infinitely lived consumer-investor in each period is endowed with a unit of time that can be allocated either to leisure or to work. Moreover, individuals are endowed with an instantaneous utility function $u(c(t), l(t), p(t), Q(t))$, where $c(t)$, is consumption for that individual at period t, $l(t)$ is labour supply, $p(t)$ is an index of the responsibility that the individual feels for the pollution caused by firms that it holds shares in (*warm-glow*) and $Q(t)$, is the environmental quality. This utility is assumed to be increasing in $c(t)$ and $Q(t)$, decreasing in $l(t)$ and $p(t)$ and strictly concave. Hence, in each period an individual chooses consumption, labour supply and saving allocation.

As for firms, we assume perfectly competitive markets and constant return to scale technology. As a consequence, we can retain the "standard" second-best framework, in the sense that there are no rents.

Finally, we assume the government finances an exogenous stream of per-capita expenditure g by issuing debt (which is the only clean asset in the market) and levying taxes. To retain the second-best, we levy taxes on the choices made by the families, i.e. savings, labour supply and by firms (pollution). Consequently, we introduce a capital-income tax, a labour income tax and a tax on pollution.

2.1. Households

The lifetime utility function of an individual household, at period 0, is:

$$U(0) = \int_0^\infty e^{-\rho t} u(c(t), l(t), p(t), Q(t)) dt \quad (1)$$

with $u_c, u_Q > 0, u_l, u_p < 0, u_{cc}, u_{ll}, u_{pp}, u_{QQ} < 0$ and $\rho > 0$ the intertemporal discount rate. Population size, H, is assumed to be constant. Individual household takes the path of Q as given. In line with Dam and Scholtens (2015), the warm-glow $p(t)$ is assumed to be a function of the individual's portfolio invested in polluting firms:

$$p(t) \equiv \sum_{j=1}^{J} \frac{e^j(t)}{\bar{E}^j} \bar{p}^j(t) \quad (2)$$

where $e^j(t)$ is the number of shares of firm j owned by the individual, $\bar{E}^j$ is number of total shares of firm j, assumed to be constant, $\bar{p}^j(t)$ is the "pollution content" of firm j as perceived by the individual. The idea is that the household feels responsible for the pollution associated with its share ownership, even though the pollution is not directly felt. We allow the latter to be a function of other potentially relevant variables:

$$\bar{p}^j(t) \equiv \bar{p}(x^j(t); X(t), F(t), Q(t)) \quad (3)$$

where $x^j(t)$ is the flow of pollution produced by the jth firm, $X(t) \equiv \sum_{j=1}^{J} x^j(t)$ the aggregate flow of pollution, $F(t)$ is the aggregate gross production of the homogenous good. We assume that $\bar{p}^j$ is linear in x^j (as any non-linearity can be captured by u). Notice that $x^j(t)$ is controlled by the firm j, i.e. each firm can affect its "rating" through its decision, while aggregate variables are taken as given by each firm. While our analysis and results are quite general, in order to better characterize their economic content, we can focus on some specifications of the perceived pollution content function $\bar{p}^j(t)$:

Assumption H1: *The warm-glow perceived pollution content function $\bar{p}^j(t)$ assumes one of the following forms:*

$$\bar{p}^j(t) = \gamma \cdot x^j(t) \quad \text{(S.1)}$$

$$\bar{p}^j(t) = \gamma \cdot \frac{x^j(t)}{Q(t)} \quad \text{(S.2)}$$

$$\bar{p}^j(t) = \gamma \cdot \frac{x^j(t)}{F(t)} \quad \text{(S.3)}$$

$$\bar{p}^j(t) = \gamma \cdot \frac{x^j(t)}{X(t)} \quad \text{(S.4)}$$

with $\gamma > 0$. The first and the second specification for $\bar{p}^j(t)$ are in line with the existing literature (Dam 2011 and Dam and Heijdra 2011), while the others, to the best of our knowledge, are new and are meant to represent a situation in which the pollution content of firms, as perceived by the individual for warm-glow purposes, is relative to either aggregate economic activity (gross GDP) ($\bar{p}^j(t) = \gamma \cdot \frac{x^j(t)}{F(t)}$), or aggregate total pollution ($\bar{p}^j(t) = \gamma \cdot \frac{x^j(t)}{X(t)}$).

In particular, Assumption S.3 can be interpreted by stating that, ceteris paribus (i.e. for a given flow of pollution produced by owned firms), the bigger the economy's scale or dimension (proxied by total GDP), the lower the pain the individual will suffer from (indirectly) being responsible for producing that amount of pollution. Put it differently, a given flow of pollution will produce a different perceived damage to the individual and the latter perceived damage will be higher the smaller the economy is: the individual feels less responsible for generating a certain pollution flow the higher the level of GDP generated by those polluting firms.[7]

At each instant of time t, individual's wealth is

$$a(t) \equiv b(t) + \sum_{j=1}^{J} e^j(t)\, P_e^j(t) \tag{4}$$

where $b(t)$ is per-capita public debt, $P_e^j(t)$ the stock market price of shares. By defining $\omega^j(t) \equiv \frac{e^j(t)P_e^j(t)}{a(t)}$ as the portfolio share invested in firm j, and $V^j(t) \equiv \bar{E}^j P_e^j(t)$ as the stock market value of firm j, we have:

$$p(t) \equiv \sum_{j=1}^{J} \frac{\omega^j(t)a(t)}{V^j(t)} \bar{p}^j(t) \tag{5}$$

and the individual budget constraint reads as[8]:

$$\dot{a}(t) = \sum_{j=1}^{J} \omega^j(t)\bar{r}_e^j(t)\, a(t) + \left[1 - \sum_{j=1}^{J} \omega^j(t)\right]\bar{r}(t)a(t) + \bar{w}(t)l(t) - c(t) - z(t) \tag{6}$$

---

[7] As already mentioned in footnote 6, we decided to analyse different specifications of the warm-glow preferences, given that there is little or no empirical evidence on the exact shape of the warm-glow function. A recent empirical work on this issue is Carpenter (2018) (see also the empirical literature mentioned therein). However, the latter studies focus on the warm-glow concerning private donation to a public good, and do not consider the issue of environmental externalities and capital markets, as in our case. In any case, we follow their suggestion to assume the warm-glow as a concave function (in fact, after equation (1) we assume that $u_p < 0$ and $u_{pp} < 0$). Finally, linearity of $\bar{p}(\cdot)$ in $x^j(\cdot)$, besides being in line with the existing literature, allows also for aggregation. See also our comments after eq. (30).

[8] We follow Merton (1971).

where $\bar{r}_e^j \equiv r_e^j\big(1-\tau^a(t)\big)$ is net-of-tax return on share j, $\bar{r}(t) \equiv r(t)\big(1-\tau^a(t)\big)$ is net-of-tax interest rate on public debt, $\bar{w}(t) \equiv w(t)\Big(1-\tau^l(t)\Big)$ is the net-of-tax wage, $z(t)$ a lump sum tax (which we will set to zero in the second-best analysis) and $\tau^a(t), \tau^l(t)$ are the tax rates on capital income and labour income, respectively. Returns on shares of firm j are:

$$r_e^j(t) \equiv \frac{\dot{V}^j(t)}{V^j(t)} e^j(t) + \frac{d^j(t)}{V^j(t)} \tag{7}$$

where $\frac{d^j(t)}{V^j(t)}$ is the dividend pay-out ratio and $d^j(t)$ total dividend payments by firm j.

The individual's problem is to maximize (1) w.r.t. $c(t), l(t), \omega^j(t)$ subject to (4) and (6). The associated current value Hamiltonian is:

$$\Lambda(t) = u(t) + q(t)\dot{a}(t) \tag{8}$$

with $q(t)$ the shadow price of wealth. FOCs yield:

$$u_c(t) - q(t) = 0 \tag{9}$$

$$u_l(t) + q(t)\bar{w}(t) = 0 \tag{10}$$

$$u_p(t)\frac{\bar{p}^j(t)}{V^j(t)}a(t) + q(t)a(t)[1-\tau^a(t)][r_e^j(t) - r(t)] = 0 \tag{11}$$

$$q(t)[[1-\tau^a(t)]\big[\textstyle\sum_{j=1}^J \omega^j(t) r_e^j + \big(1-\sum_{j=1}^J \omega^j(t)\big) r(t)\big] + u_p(t)\sum_{j=1}^J \frac{\omega^j(t)\bar{p}^j(t)}{V^j(t)} = \rho q(t) - \dot{q}(t) \tag{12}$$

Note that eq. (9) and eq. (11) provide:

$$-\frac{u_p(t)}{u_c(t)}\bar{p}^j(t) = [1-\tau^a(t)]V^j(t)[r_e^j(t) - r(t)] \tag{13}$$

Equations (9)-(10) provide the usual optimality conditions for consumption and labour supply; eq. (13) is the no-arbitrage condition for shares and bonds. Intuitively, since $u_p < 0, u_c > 0$ and $\bar{p}(\cdot) > 0$, the equation states that the individual investor demands a higher rate of return on shares than on clean bonds, because the former produces an undesirable side effect in the form of perceived pollution: in fact, the left hand side measures the marginal willingness to pay for avoiding the responsibility of owning shares in a polluting firm times the (perceived) pollution-content of the firm, while the right hand side is the marginal post-tax benefit of owning shares in the same firm measured as the excess return compared to government bonds. Notice that the return on assets in production is greater than the return on government bonds and the difference is proportional to the pollution content by the firm, thus there is a pollution premium, a compensation required by the household for holding "dirty assets". Exploiting (7), (13) becomes:

$$\frac{u_p(t)}{u_c(t)}\frac{\bar{p}^j(t)}{[1-\tau^a(t)]} + \dot{V}^j(t) + d^j(t) - r(t)V^j(t) = 0 \tag{14}$$

Finally, pre-multiplying (11) by $\omega^j(t)$ and summing from j=1 to J and using (12) we have:

$$q(t)[1-\tau^a(t)]r(t) = \rho q(t) - \dot{q}(t). \tag{15}$$

2.2. Firms

We assume that each firm runs its business in a perfectly competitive market, endowed with constant-returns-to-scale production technology that uses capital and labour inputs to produce a homogenous good. We shall also assume that each firm's technologies are the same. Hence, it will be possible to aggregate the firms to obtain a representative firm. The production function for firm j is:

$$y^j(t) = f^j\big(k^j(t), l^j(t)\big) \tag{16}$$

with $k^j(t)$ physical capital input and $l^j(t)$ labour input, respectively. We follow Brock and Taylor (2005) by assuming that, at any time t, every unit of output generates ε units of pollution as a joint product of output and that pollution can be reduced by abatement activity by the firm, $\alpha(t)$. The latter is supposed to be carried out through a CRS technology that is an increasing function of the total scale of firm activity $f(t)$ and of the firm's efforts at abatement, $f^{\alpha}(t)$. If abatement at level $\alpha(t)$ removes $\varepsilon \cdot \alpha(t)$ units of pollution, we have that total emissions (pollution) $x(t)$ by firm j is equal to:

$$x^j(t) = \varepsilon \cdot f^j(t) - \varepsilon \cdot \alpha\big(f^j(t), f^{\alpha j}(t)\big) \tag{17}$$

Defining $\psi^j(t) \equiv \frac{f^{\alpha j}(t)}{f^j(t)}$ as the fraction of output devoted to abatement activity and exploiting CRS, we get:

$$\frac{x^j(t)}{f^j(t)} = \varepsilon \cdot \big[1 - \alpha\big(1, \psi^j(t)\big)\big] = \varepsilon \cdot \big[1 - \alpha\big(\psi^j(t)\big)\big] \tag{18}$$

with α increasing in ψ^j and, thus, eq. (18) gives $\psi^j(t) = \Psi\left(\frac{x^j(t}{f^j(t)}\right)$ [9], with $\Psi' < 0, \Psi'' > 0$. Gross operating profits of the firms are:

$$\pi^j(t) \equiv \left[1 - \Psi\left(\frac{x^j(t}{f^j(t)}\right)\right] f^j\left(k^j(t), l^j(t)\right) - w(t)l^j(t) - \tau^x(t)x^j(t) \tag{19}$$

where $\tau^x(t)$ is the tax on pollution at time t. Given that we assume that the number of shares remains constant and that we abstract from corporate bonds[10], new investments, $i^j(t)$, can only by financed via retained earnings, $Re(t)$, i.e. $\pi^j(t) = d^j(t) + Re^j(t)$. That is, by exploiting the capital accumulation identity:

$$\dot{k}^j(t) = i^j(t) - \delta k^j(t) \tag{20}$$

with δ the (constant) instantaneous depreciation rate, we get:

$$\dot{k}^j(t) = \pi^j(t) - d^j(t) - \delta k^j(t) \tag{21}$$

and, exploiting (19), (21) becomes:

$$\dot{k}^j(t) = \left[1 - \Psi\left(\frac{x^j(t}{f^j(t)}\right)\right] f^j\left(k^j(t), l^j(t)\right) - w(t)l^j(t) - \tau^x(t)x^j(t) - d^j(t) - \delta k^j(t) \tag{22}$$

---

[9] From (18), $\alpha(\psi^j(t)) = 1 - \frac{1}{\varepsilon} \cdot \frac{x^j(t}{f^j(t)}$ gives $\psi^j(t) = \alpha^{-1}\left(1 - \frac{1}{\varepsilon} \cdot \frac{x^j(t}{f^j(t)}\right) \equiv \Psi\left(\frac{x^j(t}{f^j(t)}\right)$.

[10] Corporate bonds would be equivalent to shares in our model, as they would also carry the same pollution premium as shares.

Now, integrating (14) we get:

$$V^j(0) = \int_0^\infty e^{-\int_0^t r(s)ds} \left[d^j(t) + \frac{u_p(t)}{u_c(t)} \frac{\bar{p}^j(t)}{[1-\tau^a(t)]}\right] dt \tag{23}$$

which provides the value of the firm at time 0. Substituting for $d^j(t)$ from (22), (23) reads as:

$$V^j(0) = \int_0^\infty e^{-\int_0^t r(s)ds} \left\{\left[1 - \Psi\left(\frac{x^j(t}{f^j(t)}\right)\right] f^j(k^j(t), l^j(t)) - w(t)l^j(t) - \tau^x(t)x^j(t) - \delta k^j(t) + \frac{u_p(t)}{u_c(t)} \frac{\bar{p}^j(t)}{[1-\tau^a(t)]} - \dot{k}^j(t)\right\} dt \tag{24}$$

Given the assumption of perfect competition, the firm hires labour, $l^j(t)$, on the spot market and remunerates it according to its marginal productivity. In fact, FOCs on (24) w.r.t. $l^j(t)$ and $x^j(t)$ yield, respectively:

$$\left[1 - \Psi\left(\frac{x^j(t}{f^j(t)}\right) + \Psi'\left(\frac{x^j(t}{f^j(t)}\right)\frac{x^j(t)}{f^j(t)}\right] f_l^j(t) - w(t) = 0 \tag{25}$$

$$\frac{u_p(t)}{u_c(t)} \frac{1}{[1-\tau^a(t)]} \frac{\partial \bar{p}^j(t)}{\partial x^j(t)} - \Psi'\left(\frac{x^j(t}{f^j(t)}\right) - \tau^x(t) = 0 \tag{26}$$

The optimality condition for $k^j(t)$, $\frac{dV^j(0)}{dk^j(t)} = \frac{d}{dt}\frac{dV^j(0)}{d\dot{k}^j(t)}$, classical calculus of variation, gives:

$$\int_0^\infty e^{-\int_0^t r(s)ds} \left\{\left[1 - \Psi\left(\frac{x^j(t}{f^j(t)}\right) + \Psi'\left(\frac{x^j(t}{f^j(t)}\right)\frac{x^j(t)}{f^j(t)}\right] f_k^j(t) - \delta\right\} dt = \frac{d}{dt}\int_0^\infty e^{-\int_0^t r(s)ds}\, dt \Rightarrow$$

$$\left[1 - \Psi\left(\frac{x^j(t}{f^j(t)}\right) + \Psi'\left(\frac{x^j(t}{f^j(t)}\right)\frac{x^j(t)}{f^j(t)}\right] f_k^j(t) - \delta = r(t) \tag{27}$$

Finally, it can be shown that, by plugging (25)-(27) into (24) and exploiting CRS in $f^j(t)$, then $maxV^j(0) \equiv \bar{E}^j P_e{}^j(0) = k^j(0)$.

3. The Ramsey problem

We now solve the optimal tax problem (Ramsey problem). In doing so, we adopt the primal approach, consisting of the maximization of a direct social welfare function through the choice of quantities (i.e. allocations; see Atkinson and Stiglitz 1972). For this purpose, we must restrict the set of allocations among which the government can choose to those that can be decentralized as a competitive equilibrium. We now provide the constraints that must be imposed on the government's problem in order to comply with this requirement.

In our framework there is an implementability constraint associated with the individual's intertemporal choice plan. More precisely this constraint is the individual budget constraint with prices substituted for by using the individual's first order conditions, which yields (see Appendix A.1):

$$a(0)q(0) = \int_0^\infty e^{-\rho t}\left[u_p(t)p(t) + u_l(t)l(t) + u_c(t)c(t) + u_c(t)z(t)\right]dt \tag{28}$$

Finally, there are two feasibility constraints, one requires that, under the assumption that firms are equal, private and public consumption plus investment be equal to aggregate output, i.e.[11]

$$\dot{K}(t) = \left[1 - \Psi\left(\frac{X(t)}{F(t)}\right)\right] F\big(K(t), L(t)\big) - c(t)H - \delta K(t) - G(t). \tag{29}$$

---

[11] In fact, aggregating over firms we get:

$\sum_{j=1}^{J}\left[1 - \Psi^j\left(\frac{x^j(t}{f^j(t)}\right)\right] f^j(k^j(t), l^j(t)) = \sum_{j=1}^{J}\left[1 - \Psi\left(\frac{x(t)}{f(t)}\right)\right] f(k(t), l(t)) = \left[1 - \Psi\left(\frac{X(t)}{F(t)}\right)\right] F(K(t), L(t)),$

with $L(t) = l(t)H$.

with $L(t) = l(t)H$. The other one is given by the dynamics of environmental quality, which we assume, as in Dam and Heijdra (2011):

$$\dot{Q}(t) = -\mu Q(t) - \eta X(t) + \phi \tag{30}$$

This formulation captures the idea that there is a natural level of environmental quality, which the environment converges to in absence of pollution, $\frac{\phi}{\mu}$. Moreover, from (30) $X(t)$ can be thought of as the negative of decentralized provision of additions to a public good $Q(t)$ (as in Besley and Gathak 2007).

Notice that in equilibrium $\omega^j(t) \equiv \frac{V^j(t)}{H \cdot a(t)}$; then, by (4):

$$p(t) \equiv \sum_{j=1}^{J} \frac{\omega^j(t)a(t)}{V^j(t)} \bar{p}^j(t) = \sum_{j=1}^{J} \frac{\bar{p}^j\left(x^j(t); F(t), X(t), Q(t)\right)}{H} = \frac{\bar{p}(X(t); F(t), X(t), Q(t))}{H}$$

where the last equality follows from $\bar{p}^j(t)$ being linear in $x^j(t)$. Hence, at the equilibrium:

$\bar{p}(t) = \gamma \cdot X(t)$ in specification (S.1)

$\bar{p}(t) = \gamma \cdot \frac{X(t)}{Q(t)}$ in specification (S.2)

$\bar{p}(t) = \gamma \cdot \frac{X(t)}{F(t)}$ in specification (S.3)

$\bar{p}(t) = \gamma$ in specification (S.4)

Suppose that the tax programme is chosen in period 0, so the problem of the policymaker is to maximize (1) subject to eq. (28) and, $\forall t \geq 0$, (29) and (30). The current value Hamiltonian is:

$$\Lambda(t) = Hu\left(c(t), l(t), \frac{\bar{p}(t)}{H}, Q(t)\right) + \lambda H\left[u_p(t)p(t) + u_l(t)l(t) + u_c(t)c(t) + u_c(t)z(t)\right] +$$

$$q^k(t)\left\{\left[1 - \Psi\left(\frac{X(t)}{F(t)}\right)\right] F(K(t), l(t)H) - c(t)H - \delta K(t) - G(t)\right\} + q^Q(t)\{-\mu Q(t) - \eta X(t) + \phi\} \tag{31}$$

where λ is the multiplier associated with the implementability constraint and $q^k(t)$ and $q^Q(t)$ are the co-states associated with the other constraints and where we have made use of the equilibrium condition $L(t) = l(t)H$.

Preliminarily, notice that, by eq. (3), we allow the warm-glow to be a function of other variables, say S_i, so that the partial derivative of the Hamiltonian with respect to any such variable will be (omitting time subscripts):

$$\frac{d\Lambda}{dS_i} = \frac{\partial\Lambda}{\partial S_i} + \frac{\partial\Lambda}{\partial p}\frac{\partial p}{\partial S_i},$$

with

$$\frac{\partial\Lambda}{\partial p} = Hu_p(1 + \lambda\Delta_p) \tag{32}$$

and $\Delta_p \equiv 1 + \frac{u_{pp}p}{u_p} + \frac{u_{lp}l}{u_p} + \frac{u_{cp}c}{u_p}$, referred to as the "general equilibrium elasticity" of the warm-glow. The first order conditions of the Ramsey problem are (omitting time subscripts):

$$\frac{\partial\Lambda}{\partial c} = 0 \Longrightarrow u_c(1 + \lambda\Delta_c) = q^k \tag{33}$$

$$\frac{\partial\Lambda}{\partial l} = 0 \Longrightarrow u_p(1 + \lambda\Delta_p)\frac{\partial\bar{p}}{\partial F}F_L + u_l(1 + \lambda\Delta_l) + q^k\left[1 - \Psi\left(\frac{X}{F}\right) + \Psi'\left(\frac{X}{F}\right)\frac{X}{F}\right]F_L = 0 \tag{34}$$

$$\frac{\partial\Lambda}{\partial K}: u_p\left(1+\lambda\Delta_p\right)\frac{\partial\bar{p}}{\partial F}F_K + q^k\left\{\left[1-\Psi\left(\frac{X}{F}\right)+\Psi'\left(\frac{X}{F}\right)\frac{X}{F}\right]F_K-\delta\right\} = \rho q^k - \dot{q}^k \tag{35}$$

$$\frac{\partial\Lambda}{\partial X} = 0 \Longrightarrow u_p\left(1+\lambda\Delta_p\right)\frac{\partial\bar{p}}{\partial X} - q^k\Psi'\left(\frac{X}{F}\right) - \eta q^Q = 0 \tag{36}$$

$$\frac{\partial\Lambda}{\partial Q}: u_p\left(1+\lambda\Delta_p\right)\frac{\partial\bar{p}}{\partial Q} + Hu_Q\left(1+\lambda\Delta_Q\right) = (\rho+\mu)q^Q - \dot{q}^Q \tag{37}$$

with $\Delta_c \equiv 1+\frac{u_{cc}c}{u_c}+\frac{u_{lc}l}{u_c}+\frac{u_{pc}p}{u_c}, \Delta_l \equiv 1+\frac{u_{ll}l}{u_l}+\frac{u_{lc}c}{u_l}+\frac{u_{pl}p}{u_l}$, usually referred to as the "general equilibrium elasticities" of consumption and labour, respectively and $\Delta_Q \equiv \frac{u_{pQ}p}{u_Q}+\frac{u_{lQ}l}{u_Q}+\frac{u_{cQ}c}{u_Q}$.

By dividing (34) by (33), exploiting (25) (recognizing that $f_l^j(t) = F_L$) and the equilibrium condition stemming from (9) and (10) (i.e. $u_c\overline{w} = -u_l$) we get:

$$\frac{\tau^l}{1-\tau^l} = \frac{u_p}{u_l}\frac{\partial\bar{p}}{\partial F}F_L + \lambda\frac{u_p}{u_l}\frac{\left(\Delta_p-\Delta_c\right)}{(1+\lambda\Delta_c)}\frac{\partial\bar{p}}{\partial F}F_L + \lambda\frac{(\Delta_l-\Delta_c)}{(1+\lambda\Delta_c)} \tag{38}$$

which provides the implicit expression for the labour-income tax.

As for the capital income tax, plugging (27) into (35) (recognizing that $f_k^j(t) = F_K$) and dividing by (33) we get:

$$\frac{\dot{q}^k}{q^k} = -\frac{u_p\left(1+\lambda\Delta_p\right)}{u_c(1+\lambda\Delta_c)}\frac{\partial\bar{p}}{\partial F}F_K + \rho - r \tag{38'}$$

Next, plugging (9) into (33), time-differentiating and exploiting (15) we obtain

$$\frac{\dot{q}^k}{q^k} = \frac{\dot{q}}{q} + \lambda \frac{\dot{\Delta}_c}{1+\lambda\Delta_c} = \rho - [1 - \tau^a]r + \lambda \frac{\dot{\Delta}_c}{1+\lambda\Delta_c} \tag{38''}$$

Equating (38') and (38'') and simplifying yields:

$$\tau^a = -\frac{u_p}{u_c}\frac{\partial\bar{p}}{\partial F}\frac{F_K}{r} + \frac{u_p}{u_c}\lambda\frac{(\Delta_c - \Delta_p)}{(1+\lambda\Delta_c)}\frac{\partial\bar{p}}{\partial F}\frac{F_K}{r} - \frac{\lambda}{r}\frac{\dot{\Delta}_c}{1+\lambda\Delta_c} \tag{39}$$

Preliminarily, notice that if the implementability constraint does not bind, then its associated multiplier λ is equal to zero. This means that, given that only the resource constraints are binding, the government is able to implement the first best allocation, thus correcting for the presence of market failures (the usual conditions for Pareto optimality apply). On the other hand, if the implementability constraint binds, then λ is different from zero and the Lipsey-Lancaster (1956) second-best theorem applies. In particular, we assume that this case occurs because the Government cannot raise the amount of (per capita) lump-sum taxes that are needed to implement the first best allocation (say, z^*), but only the amount $\bar{z}$ (with $0 \leq \bar{z} < z^*$). Given the second-best suboptimality of this situation, the first derivative of the current value Hamiltonian with respect to z, evaluated at the upper bound $\bar{z}$, must be positive. From (31), $\left.\frac{\partial\Lambda}{\partial z}\right|_{z=\bar{z}} = \lambda H u_c > 0$ only if $\lambda > 0$. For this reason, λ is usually interpreted as a measure of the deadweight loss brought about by distortionary taxation.[12]

As for the capital income tax, eq. (39) shows that, at the steady state (i.e. when all per-capita variables are constant and hence $\dot{\Delta}_c = 0$), it is different from zero only if $\frac{\partial\bar{p}}{\partial F} \neq 0$, both in the first and second best (i.e. when $\lambda = 0$ and $\lambda > 0$, respectively). The reason is that, in this case, the

[12] For the sake of simplicity, we will set $\bar{z} = 0$ in the analysis that follows.

government finds it optimal to hit (or subsidize) capital income to correct for the externality that the capital stock exerts on the individual's warm-glow via total output (i.e. assumption S.3). Along the transition path, on the other hand, capital income tax can be different from zero even if $\frac{\partial \bar{p}}{\partial F} \neq 0$, provided that $\dot{\Delta}_c \neq 0$: if the general elasticity of consumption decreases (increases), future consumption becomes less (more) elastic and then it is second-best optimal to levy positive (negative) taxes on capital income (i.e. on future consumption). As the for labour income tax, eq. (38) applies both along the transition path and the steady state and shows that, in the first best ($\lambda = 0$), the same argument applies as in the previous case: the government hits (or subsidizes) labour income only for corrective purposes (externality in the warm-glow, $\frac{\partial \bar{p}}{\partial F} \neq 0$). However, in the second best ($\lambda > 0$), things are different, in that a non-zero labour income tax may be optimal even in the case of absence of externality, and, as we will show in section 4, the sign of the optimal tax will depend, among other things, on the nature of labour supply (i.e. whether the latter is a normal of inferior good). We provide further comments on such findings in section 4.

As for the Pigouvian tax on pollution *X*, by substituting for q^Q from (36) into (37) and exploiting (33) one gets:

$$\frac{u_p}{u_c}\frac{(1+\lambda\Delta_p)}{(1+\lambda\Delta_c)}\left[\frac{\eta}{\rho+\mu}\frac{\partial\bar{p}}{\partial Q}-\frac{\partial\bar{p}}{\partial X}\right]+\frac{\eta}{\rho+\mu}H\frac{u_Q}{u_c}\frac{(1+\lambda\Delta_Q)}{(1+\lambda\Delta_c)}+\frac{\eta}{\rho+\mu}\frac{\dot{q}^Q}{u_c(1+\lambda\Delta_c)}=-\Psi'\left(\frac{X}{F}\right) \tag{40}$$

Next, exploiting (15) and (38') the following relationship obtains:

$$\frac{1}{[1-\tau^a]}=\frac{\rho-\frac{u_p(1+\lambda\Delta_p)\partial\bar{p}}{u_c(1+\lambda\Delta_c)\partial F}F_K-\frac{\dot{q}^k}{q^k}}{\rho-\frac{\dot{q}}{q}} \tag{40'}$$

Plugging (40') into (26), substituting into (40) for $-\Psi'$ and rearranging terms we can provide the following decomposition of τ^x:

$$\tau^x = \tau^x_{FB} + \tau^x_{FB}(p) + \tau^x_{SB} + \tau^x_{SB}(p) + \tau^x_D(p) \tag{41}$$

with

$$\tau^x_{FB} = \frac{\eta}{\rho+\mu} H \frac{u_Q}{u_c} \tag{41a}$$

$$\tau^x_{FB}(p) = \frac{u_p}{u_c}\left(\frac{\eta}{\rho+\mu}\frac{\partial\bar{p}}{\partial Q} - \frac{u_p}{u_c}\frac{\partial\bar{p}}{\partial X}\frac{\partial\bar{p}}{\partial F}\frac{F_K}{\rho}\right) \tag{41b}$$

$$\tau^x_{SB} = \frac{\eta}{\rho+\mu} H \frac{u_Q}{u_c}\lambda\frac{(\Delta_Q - \Delta_c)}{(1+\lambda\Delta_c)} \tag{41c}$$

$$\tau^x_{SB}(p) = \frac{u_p}{u_c}\lambda\frac{(\Delta_p - \Delta_c)}{(1+\lambda\Delta_c)}\left(\frac{\eta}{\rho+\mu}\frac{\partial\bar{p}}{\partial Q} - \frac{\partial\bar{p}}{\partial X} - \frac{u_p}{u_c}\frac{\partial\bar{p}}{\partial X}\frac{\partial\bar{p}}{\partial F}\frac{F_K}{\rho}\right) \tag{41d}$$

$$\tau^x_D(p) = \frac{u_p}{u_c}\frac{\partial\bar{p}}{\partial X}\frac{1}{\rho-\frac{\dot{q}}{q}}\left\{\left[1 - \frac{u_p(1+\lambda\Delta_p)}{u_c(1+\lambda\Delta_c)}\frac{\partial\bar{p}}{\partial F}\frac{F_K}{\rho}\right]\frac{\dot{q}}{q} - \frac{\dot{q}^k}{q^k}\right\} + \frac{\eta}{\rho+\mu}\frac{\dot{q}^Q}{u_c(1+\lambda\Delta_c)} \tag{41e}$$

where τ^x_{FB}, τ^x_{SB} are the first-best and second-best tax components (i.e. when $\lambda = 0$ and $\lambda > 0$, respectively) in a framework without the warm-glow component ($u_p = 0$) and $\tau^x_{FB}(p) + \tau^x_{SB}(p)$ are the first-best and second-best components of the pollution tax that add to the previous ones in the presence of warm-glow. Notice that in the absence of public goods and externalities, that is with Q not being present in the utility function and no externalities in H.1 (i.e. avoiding (S.2) or (S.3) specifications), the decentralised equilibrium would coincide with the planner's solution (i.e. first welfare theorem applies), with stock prices incorporating both economic and social returns as in Baron (2007). $\tau^x_D(p)$ is the dynamic component of the pollution tax that is only present along the transitional path and is zero along the steady state. For the sake of clarity, in the reminder of the work we will focus on the steady state results[13].

As for the first best, without warm-glow ($u_p = 0$), X is an externality: there is no market for such a commodity, while it affects the public good consumed by individuals (Q), so that its marginal cost should be corrected by Pigouvian taxation (eq. 41a). On the other hand, in the presence of warm-glow ($u_p < 0$), X is a private good representing a (negative) contribution to a public good (Q) and is not an externality because it has a market (stock market). However, its market price is not optimal and should be corrected through the Samuelson's rule (again eq.41a). More precisely, notice that in cases (S.1) and (S.3) the first-best add-on component due to the warm glow is zero (in that in 41b the terms $\frac{\partial \bar{p}}{\partial Q} = \frac{\partial \bar{p}}{\partial F} = 0$), while it is different from zero in case (S.2) and (S.4). In fact, in the latter cases, while X is still a private contribution to a public, there are also externalities at work: either the level of Q or F affect the marginal disutility of the warm-glow. The corrections for such externalities, in case (S.2), is operated through the add-on term $-\frac{u_p}{u_c}\left(\frac{\eta}{\rho+\mu}\frac{\bar{p}}{Q}\right) > 0$, while in case (S.3) is operated through the add-on term $\left(\frac{u_p}{u_c}\right)^2\left(\frac{\bar{p}}{F^2}\frac{F_K}{\rho}\right) > 0$ and by hitting/subsidising both capital and labour income (see next section for further explanation of these findings and the sign of these corrective components).

Finally, equation (41d) is the second-best warm glow add-on component. Its sign is the sign of $\left(\Delta_p - \Delta_c\right)$, and is related to whether consumption or warm glow is a stronger Hicks complement of leisure. Second-best tax theory produces uniform taxation of commodities which are equally strong Hicks complements to leisure. The good which has the stronger

---

[13] It follows from (15), (27), (33)-(37) that when real quantities are constant, the individual co-states q^k, q^Q and q are also constant. We focus on steady state results though there may be occasions, for certain models, where a second-best economy does not converge to a steady state (see Straub and Werning 2018).

complementarity should be taxed at a higher rate (Atkinson and Stiglitz 1972). If $(\Delta_p - \Delta_c) > 0$ it means (in our model) that warm glow as a good has stronger complementarity with leisure than consumption c. This should imply that warm glow as a good should bear a higher tax than ordinary consumption. However, as we do not have a tax instrument on warm glow itself (and the consumption tax has been set to zero), the extra tax on pollution x (41d) plays the role of this missing instrument. Only when we have equal complementarity $(\Delta_c - \Delta_p) = 0$, this component is redundant (as then c and p are taxed at the uniform rate 0).

Table 1 summarizes our findings on τ^x (see Appendix A.4 for details)

Table 1: Summary of the results on the pollution tax τ^x at the steady state

	τ^x	
	No warm-glow (u_p =0)	**With warm-glow ($u_p < 0$)**
First-Best	τ^x_{FB}	$\tau^x_{FB} + \tau^x_{FB}(p)$
Second-Best	$\tau^x_{FB} + \tau^x_{SB}$	$\tau^x_{FB} + \tau^x_{FB}(p) + \tau^x_{SB} + \tau^x_{SB}(p)$

4. Discussion of the results

In this section we analyse and comment on our obtained results on the optimal structure of taxes along the steady state.

We first present the optimal tax structure at the first-best, and we focus on the second-best. We adopt a simplifying assumption on the shape of the utility function, on which we introduce the restriction of partial additivity.

Assumption H2: *The instantaneous utility function assumes the following form*:

$$u(c, l, p, Q) = v(c, l) + h(p, Q)$$

i.e. additive and separable in (c, l) *and* (p, Q).

We now provide the following Lemma:

Lemma 1: Under H2, if leisure is non-inferior, then:

L.1) $\Delta_l - \Delta_c > 0$

L.2) $\Delta_p - \Delta_c > 0$

L.3) $\left(1 + \lambda\Delta_p\right) > 0$

Proof. See Appendix A.2. □

Under the above assumptions, we can now provide the following Proposition characterizing the first-best tax structure.

Proposition 1. *At the steady state, the first-best tax structure is the following:*

a) $\tau^x > 0$*;*

b) $\tau^a, \tau^l = 0$ *under specifications (S.1), (S.2) and (S.4),*

c) $\tau^a, \tau^l < 0$ *under specification (S.3).*

Proof. The result on τ^x descends from eq. (41), whereby at the first-best $\tau^x = \tau^x_{FB} + \tau^x_{FB}(p)$, with $\tau^x_{FB} > 0$ for all specifications of $\bar{p}^j(t)$ and $\tau^x_{FB}(p) > 0$ for specifications (S.2) and (S.3), (zero otherwise[14]). As for τ^a, τ^l, the results descend from the fact that, at the first-best, $\lambda = 0$ and, that, under specifications (S.1), (S.2) and (S.4), $\frac{\partial \bar{p}}{\partial F} = 0$, so that the results sub b) of zero taxes follow from mere observation of (38) and (39). Under specification (S.3), $\frac{\partial \bar{p}}{\partial F} < 0$; moreover, recall that $u_c > 0, u_l, u_p < 0$. Hence, by observation of (38) and (39), $sign(\tau^a) =$

[14] Note that under specification (S.4), $\frac{\partial \bar{p}}{\partial X}$ appearing in the expression for $\tau^x_{FB}(p)$ is zero.

$sign(\tau^l)$. Finally, by (33), $(1+\lambda\Delta_c)>0$ and by Lemma 1, sub L.3), $\left(1+\lambda\Delta_p\right)>0$, so that $\tau^a, \tau^l<0$. □

Under formulation (S.1) and (S.4) there is no correction for warm-glow in the pollution tax (i.e. $\tau_{FB}^x(p)=0$) and we get standard expression for corrective taxation, while the add-on correction is present under specifications (S.2) and (S.3) (see eq. 41b). Furthermore, both taxes on capital and labour income are either zero or, in case the warm-glow depends on the scale of economic activity (S.3), negative (i.e. both inputs should be subsidized to reduce the individual's perceived damage caused by firms).

The economic rationale behind our findings it that, in formulation (S.2), while firms realise the consequence of pollution on its own perceived pollution content, they do not realise that they (in the aggregate) affect the state of the environment (environmental quality). If they did realise they affected the aggregate, they would have an incentive to lower pollution at each date to increase Q (again in order to lower the cost of capital). This needs to be corrected for in the first best, with an extra (positive) component added to the Pigou tax.

As for formulation (S.3), the economic explanation for the result is that each firm realises that its individual pollution affects the perceived pollution content, but does not take into account the effect on aggregate production. If the firm could increase aggregate production, it would do so in order to reduce the perceived pollution content and lower the pollution premium (to lower the cost of capital). In the first best, this needs to be corrected for. Thus, capital and labour are subsidised to increase aggregate production. However, the correction to increase aggregate production, to lower the perceived pollution content, will imply that the abatement incentive for the firm is lowered. Therefore, the new Pigou tax needs to contain the extra (positive) component.

Let us now turn to the second-best tax structure, which we characterize through the following Proposition:

Proposition 2. *At the steady state, the second-best tax structure is the following:*

I) $\tau^x > 0$*;*

II) As for τ^a:

II.A) $\tau^a = 0$ *under specifications (S.1), (S.2) and (S.4),*

II.B) $\tau^a < 0$ *under specification (S.3)*.

III) As for τ^l:

III.A) $\tau^l > 0$ *under specifications (S.1), (S.2) and (S.4),*

III.B) Its sign is ambiguous under specification (S.3).

Proof. See Appendix A.3. □

As a final comment, we notice that the ambiguity of the sign of the labour income tax in specification (S.3) stems from the fact that the second-best component would make it optimal for the policymaker to levy positive taxes on labour income, while the first-best component does exert an opposite effect. We can summarize our results through the following Table (see Appendix A.4 for details).

Table 2: Summary of the results on first-best and second best-tax structure

	$\boldsymbol{\tau^x}$	$\boldsymbol{\tau^a}$		$\boldsymbol{\tau^l}$	
		Specifications (S.1), (S.2), (S.4)	Specification (S.3)	Specifications (S.1), (S.2), (S.4)	Specification (S.3)
First-Best	$\tau^x > 0$	$\tau^a = 0$	$\tau^a < 0$	$\tau^l = 0$	$\tau^l < 0$
Second-Best	$\tau^x > 0$			$\tau^l > 0$	τ^l ambiguous

5. The role of the warm-glow in the first-best allocation

In this section we provide an example to assess the role of the warm-glow in shaping the optimal tax structure in the first-best. To simplify, we assume the following utility function:

$$u(c, l, p, Q) = \log\left(c - \frac{l^{1+\sigma}}{1+\sigma}\right) + \log(Q) + \xi \cdot h(p) \tag{42}$$

with $h' < 0$ and $h'' < 0$. As for the warm-glow, we choose specification (S.1), i.e. $\bar{p}^j(t) = \gamma \cdot x^j(t)$ and, as for technology, the following Cobb-Douglas production function: $F = K^\beta \cdot L^{1-\beta}$, while capital depreciation, δ, is assumed to be zero. We wish to assess how the corrective tax changes as warm-glow preferences become stronger, that is how τ^x varies with ξ.

Recall that, in the first-best, under (S.1), both τ^a and τ^l are zero. By dividing (9) and (10) and exploiting (25), at the steady state we get:

$$-\frac{u_l}{u_c} = l^\sigma = w = \left[1 - \Psi\left(\frac{X}{F}\right) + \Psi'\left(\frac{X}{F}\right)\frac{X}{F}\right] F_L \tag{43}$$

which, by defining $\theta \equiv \frac{X}{F}$, can be written as:

$$H \cdot l^{1+\sigma} = w = [1 - \Psi(\theta) + \Psi'(\theta) \cdot \theta] \cdot (1 - \beta) \cdot F \tag{44a}$$

or

$$L \equiv H \cdot l = H \cdot [1 - \Psi(\theta) + \Psi'(\theta) \cdot \theta]^{\frac{1}{\sigma}} \cdot (1 - \beta)^{\frac{1}{\sigma}} \cdot \left(\frac{K}{L}\right)^{\frac{1}{\sigma}} \tag{44b}$$

Next, rewriting (27) as:

$$[1 - \Psi(\theta) + \Psi'(\theta) \cdot \theta] \cdot \beta \cdot \left(\frac{L}{K}\right)^{1-\beta} = \rho + \delta \quad (45)$$

By exploiting (44b), (45) and the production function we get and implicit equilibrium equation $F = F(\theta)$, and by log-differentiation of (44b), (45) and the production function we get:

$$\frac{F_\theta}{F} = \frac{1 + \beta \cdot \sigma}{(1 - \beta) \cdot \sigma} \cdot \frac{\Psi''(\theta) \cdot \theta}{[1 - \Psi(\theta) + \Psi'(\theta) \cdot \theta]} \geq 0.$$

Taking (26)

$$\tau^x = -\Psi'(\theta) + \frac{\xi h'\left(\frac{\gamma}{H} \cdot X\right) \cdot \gamma}{u_c} \quad (46)$$

[15]

and recognizing that:

$$\frac{1}{u_c} = \frac{1}{H} \cdot \left(H \cdot c - H \cdot \frac{l^{1+\sigma}}{1+\sigma}\right) \quad (47)$$

and exploiting (29) at steady state:

$$c \cdot H = [1 - \Psi(\theta)] \cdot F(\theta) - G \quad (48)$$

to substitute for $c \cdot H$ into (47) and (46), we get:

---

[15] Notice that, under specification (S.1), the equilibrium value of the warm-glow function is $p = \sum_{j=1}^{J} \frac{\bar{p}^j}{H} = \gamma \cdot \frac{X}{H}$, with $\frac{\partial \bar{p}}{\partial X} = \gamma$.

$$\tau^x = -\Psi'(\theta) + \xi \cdot h'\left(\frac{\gamma}{H} \cdot \theta \cdot F(\theta)\right) \cdot \frac{\gamma}{H} \cdot \left[(1-\Psi) \cdot F(\theta) - G - \frac{1-\beta}{1+\sigma} \cdot (1 - \Psi(\theta) + \Psi'(\theta) \cdot \theta) \cdot F(\theta)\right] \tag{49}$$

Moreover, by substituting for Q from (30), evaluated at the steady state, into (41a) and exploiting (47) and (29), yields:

$$\tau^x = \frac{\eta}{\rho+\mu} \cdot H \cdot M(\theta, F(\theta)) \tag{50}$$

with $M\big(\theta, F(\theta)\big) \equiv \frac{u_Q}{u_c} = \mu \cdot \frac{1-\Psi(\theta)-\frac{G}{F(\theta)}-\frac{1-\beta}{1+\sigma}\cdot[1-\Psi(\theta)+\Psi'(\theta)\cdot\theta]}{\frac{\phi}{F(\theta)}-\eta\cdot\theta} > 0.$

Equations (49) and (50) provide the two equilibrium conditions for first-best optimal allocations of θ and τ^x. In order to evaluate the effect of the warm-glow on the steady state, we apply Cramer's rule to get the partial derivatives of θ and τ^x with respect to the parameter ξ. Total differentiation of (49) and (50) yields:

$$d\tau^x = -(\Psi'' + \xi \cdot m) \cdot d\theta + \frac{h\prime(p)}{u_c} \cdot d\xi \tag{51}$$

$$d\tau^x = \frac{\eta}{\rho+\mu} \cdot H \cdot M_\theta \cdot d\theta \tag{52}$$

with

$$m \equiv -\xi h''(p)\left(\frac{\gamma}{H}\right)^2 (F + \theta F_\theta)\left[(1-\Psi)F - G - \frac{1-\beta}{1+\sigma}(1 - \Psi + \Psi'\theta)F\right] + \xi h'(p)\frac{\gamma}{H}\left[1 - \Psi - \frac{1-\beta}{1+\sigma} \cdot (1 - \Psi + \Psi'\theta)F_\theta - \Psi'F - \Psi''F\theta\frac{1-\beta}{1+\sigma}\right] \tag{53}$$

$$M_\theta = \mu \frac{\left[-\Psi' + \frac{G}{F}\frac{F_\theta}{F} - \frac{1-\beta}{1+\sigma}\Psi''\theta\right]\left(\frac{\phi}{F} - \eta\theta\right) + \left[1 - \Psi - \frac{G}{F} - \frac{1-\beta}{1+\sigma}(1 - \Psi + \Psi'\theta)\right]\left(\frac{\phi}{F}\frac{F_\theta}{F} - \eta\right)}{\left[\frac{\phi}{F} - \eta\theta\right]^2} \tag{54}$$

Sufficient for $m > 0$ and $M_\theta > 0$ is $-\Psi' - \Psi''\theta \geq 0$ (which we will assume throughout the example[16]). From (51) and (52) we get:

$$\frac{d\tau^x}{d\xi} = \frac{h'(p)}{u_c\left[1 + (\Psi'' + \xi \cdot m) \cdot \frac{1}{\frac{\eta}{\rho+\mu} \cdot H \cdot M_\theta}\right]} < 0$$

and

$$\frac{d\theta}{d\xi} = \frac{h'(p)}{u_c\left[\frac{\eta}{\rho+\mu} \cdot H \cdot M_\theta + (\Psi'' + \xi \cdot m)\right]} < 0$$

Hence, we have shown that stronger warm-glow preferences reduce the level of the optimal pollution tax and the pollution intensity. Moreover, given that $F_\theta > 0$, stronger warm-glow preferences reduce total production and the level of capital.

6. Conclusions

In this paper we characterize optimal taxes in a continuous-time model in the presence of pollution as a joint product of production. We explicitly allow investors to engage in socially responsible investments through a *warm-glow* mechanism as in Andreoni (1990) and Dam (2011) and firms to engage in corporate socially responsible activities through pollution abatement activity.

[16] it is possible to show that the latter inequality holds under a CES abatement function with low elasticity of substitution.

We show that the first-best tax structure consists in positive taxation of pollution and either zero or negative taxation of production-factor incomes (negative taxation arising in case the perceived pollution content of firms is negatively related to the total scale of economic activity).

As for the second-best structure, we show that the pollution tax has an additivity property, consisting of the first-best component, plus the first best-warm-glow component, plus a second-best component, plus the second-best warm-glow component. Leisure being non-inferior is sufficient for the add-on components to be positive or zero, apart from the second-best warm-glow component, which can take on any sign. Overall, in the second-best the total pollution tax is positive, suggesting that warm-glow is not a substitute for the government.

While the first-best tax rule for the capital income tax also holds in the second-best, it emerges that in general, sufficient for the labour income tax to be positive is that leisure is non inferior (though its sign can be ambiguous, if the perceived pollution content of the firm depends on gross GDP). Finally, in an example we show that the presence of warm-glow preferences yields lower first-best pollution taxes.

Finally, we have focused on an identical household economy, where an individual specific lump-sum tax is unavailable (the Ramsey tradition). We do this to keep the analysis as simple as possible, given that the model is already complex, and it is the first paper dealing with optimal taxation in a dynamic general equilibrium model, in the presence of environmental warm-glow, firms' abatement activities and financial markets. we leave the second-best optimal-tax case with heterogeneous households for future research.

References

Allgood, S. (2009). Charity, impure altruism, and marginal redistributions of income. *National Tax Journal*, 219-235.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, *100*(401), 464-477.

Andreoni, J. (2006). Philanthropy. *Handbook of the economics of giving, altruism and reciprocity*, *2*, 1201-1269.

Andreoni, J., Rao, J. M., Trachtman, H. (2017). Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *Journal of Political Economy*, *125*(3), 625-653.

Aronsson, T., Sjögren, T. (2018). Optimal Taxation, Redistribution, and Environmental Externalities, *International Review of Environmental and Resource Economics*: 11(3) 233-308.

Atkinson, A. B., Stiglitz, J. E. (1972). The structure of indirect taxation and economic efficiency. *Journal of Public economics*, *1*(1), 97-119.

Ballestero, E., Pérez-Gladish, B., Garcia-Bernabeu, A. (2015). The ethical financial question and the MCDM framework. In *Socially Responsible Investment* (pp. 3-22). Springer, Cham.

Barnea, A., Heinkel, R., Krause, A., (2005). Green investors and corporate investment. *Structural Change and Economic Dynamics* 16 (3), 332–346.

Baron, D. P. (2007). Corporate social responsibility and social entrepreneurship. *Journal of Economics & Management Strategy*, *16*(3), 683-717.

Belfiori, M. E. (2018). Climate change and intergenerational equity: Revisiting the uniform taxation principle on carbon energy inputs. *Energy Policy*, *121*, 292-299.

Besley, T., Ghatak, M. (2007). Retailing public goods: The economics of corporate social responsibility. *Journal of public Economics*, *91*(9), 1645-1663.

Bontems, P., Bourgeon, J. M. (2005). Optimal environmental taxation and enforcement policy. *European Economic Review*, *49*(2), 409-435.

Bovenberg, A.L., Goulder, L.H. (2002). Environmental taxation and regulation. In: Auerbach, A.J., Feldstein, M.S. (eds) *Handbook of public economics*, vol 3. North-Holland, Amsterdam, 1471–1545.

Brekke, K. A., Pekovic, S. (2018). Why Are Firms Environmentally Responsible? A Review and Assessment of the Main Mechanisms. *International Review of Environmental and Resource Economics*, *12*(4), 355-398.

Brock, W. A., Taylor, M. S. (2005). Economic Growth and the Environment: A Review of Theory and Empirics, in: Philippe Aghion & Steven Durlauf (ed.), *Handbook of Economic Growth*, edition 1, volume 1, chapter 28, pages 1749-1821 Elsevier.

Carpenter, J. (2018). The Shape of Warm Glow: Field Experimental Evidence from a Fundraiser *IZA Discussion Paper No. 11760.*

Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica: Journal of the Econometric Society*, 607-622.

Cremer, H., Gahvari, F., Ladoux, N. (2001). Second-Best Pollution Taxes and the Structure of Preferences. *Southern Economic Journal, 68*(2), 258-280.

Dam, L. (2011). Socially responsible investment in an environmental overlapping generations model. *Resource and Energy Economics*, *33*(4), 1015-1027.

Dam, L., Heijdra, B. J. (2011). The environmental and macroeconomic effects of socially responsible investment. *Journal of Economic Dynamics and Control*, *35*(9), 1424-1434.

Dam, L., Scholtens, B. (2015). Toward a theory of responsible investing: On the economic foundations of corporate social responsibility. *Resource and Energy Economics*, *41*, 103-121.

Diamond. P (2006). Optimal Tax Treatment of Private Contributions for Public Goods with and without Warm Glow Preferences. *Journal of Public Economics* 90, 897-919.

Eurosif (2016): *European SRI Study.* Available at: http://www.eurosif.org/wp-content/uploads/2016/11/SRI-study-2016-HR.pdf

Evren, Ö., Minardi, S. (2017). Warm-glow Giving and Freedom to be Selfish. *The Economic Journal*, *127*(603), 1381-1409.

Gahvari, F. (2014). Second-best pigouvian taxation: a clarification. *Environmental and Resource Economics*, *59*(4), 525-535.

Graff Zivin, J., Small, A. (2005). A Modigliani-Miller theory of altruistic corporate social responsibility. *The BE Journal of Economic Analysis & Policy*, *5*(1).

Goulder, L. H., Parry, I. W. (2008). Instrument choice in environmental policy. *Review of environmental economics and policy*, *2*(2), 152-174.

GSIA (2016): *Global Sustainable Investment Review.* Available at: http://www.gsi-alliance.org/wp-content/uploads/2017/03/GSIR_Review2016.F.pdf

Heinkel, R., Kraus, A., Zechner, J. (2001). The effect of green investment on corporate behavior. *Journal of financial and quantitative analysis*, *36*(4), 431-449.

Hong, H., Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, *93*(1), 15-36.

Jacobs, B., & De Mooij, R. A. (2015). Pigou meets Mirrlees: On the irrelevance of tax distortions for the second-best Pigouvian tax. *Journal of Environmental Economics and Management*, *71*, 90-108.

Judd, K. L. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of public Economics*, *28*(1), 59-83.

Kampas, A., Horan, R. (2016). Second-best pollution taxes: revisited and revised. *Environmental Economics and Policy Studies*, *18*(4), 577-597.

Lipsey, R. G., Lancaster, K. (1956). The general theory of second best. *The review of economic studies*, *24*(1), 11-32.

Merton, R. C. (1971). Optimum Consumption and Portfolio-Rules in a Continuous-Time Framework. *Journal of Economic Theory*.

OECD (2015). Policy Guidance for investment in clean energy infrastructure: expanding access to clean energy for green growth and development, OECD Publishing.

Renneboog, L., Ter Horst, J., Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance*, *32*(9), 1723-1742.

Sandmo, A. (1975). Optimal taxation in the presence of externalities. *The Swedish Journal of Economics*, 86-98.

Straub, L. and Werning, I. (2018). Positive Long Run Capital Taxation: Chamley-Judd Revisited, mimeo, MIT.

Vanwalleghem, D. (2017). The real effects of sustainable & responsible investing?. *Economics Letters*, *156*, 10-14.

Appendix A.1. The implementability constraint

Time derivative of $q(t)a(t)$ is

$$\frac{d}{dt}[q(t)a(t)] = \dot{q}(t)a(t) + q(t)\dot{a}(t) \qquad \text{(A.1.1)}$$

Exploiting (6) and (15), (A.1.1) reads as:

$$\frac{d}{dt}[q(t)a(t)] = q(t)\{\overline{w}(t)l(t) - c(t) - z(t)\} - u_p(t)\sum_{j=1}^{J}\frac{\omega^j(t)\overline{p}^j(t)}{V^j(t)}a(t) + a(t)\rho q(t) \qquad \text{(A.1.2)}$$

By substituting for $q(t)$ from (9) and (10) and exploiting (4) it follows:

$$\frac{d}{dt}[q(t)a(t)] - \rho q(t)a(t) = -u_p(t)p(t) - u_c(t)c(t) - u_c(t)z(t) - u_l(t)l(t) \qquad \text{(A.1.3)}$$

Multiplying both sides by $e^{-\rho t}$ (A.1.3) can be written as:

$$\frac{d}{dt}[q(t)a(t)e^{-\rho t}] = -e^{-\rho t}[u_p(t)p(t) + u_c(t)c(t) + u_c(t)z(t) + u_l(t)l(t)] \qquad \text{(A.1.4)}$$

and integrating it follows that:

$$a(0)q(0) = \int_0^\infty e^{-\rho t}\big[u_p(t)p(t) + u_l(t)l(t) + u_c(t)c(t) + u_c(t)z(t)\big]dt \qquad \text{(A.1.5)}$$

which is eq. (28) in the text.

Appendix A.2. Proof of Lemma 1

q is the marginal shadow value of assets and is inversely related to assets. We take as normality of, say c, the case when $\frac{dc}{dq} < 0$, as it corresponds to $\frac{dc}{da} > 0$, the "income effect" keeping prices fixed. Normality of leisure is when labour increases in q, that is $\frac{dl}{dq} > 0$.

Differentiating (9) and (10) for partially separable utility (assumption H2) we have:

$$\begin{bmatrix} u_{cc} & u_{cl} \\ u_{lc} & u_{ll} \end{bmatrix} \begin{bmatrix} \frac{\partial c}{\partial q} \\ \frac{\partial l}{\partial q} \end{bmatrix} = \begin{bmatrix} 1 \\ -\overline{w} \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{u_l}{u_c} \end{bmatrix} \qquad \text{(A.2.1)}$$

Cramer's rule provides the following:

$$\frac{\partial l}{\partial q} = \frac{1}{u_l} \frac{\frac{u_{cc}}{u_c} - \frac{u_{lc}}{u_l}}{u_{cc}u_{ll} - u_{cl}u_{lc}} \qquad \text{(A.2.2)}$$

Concavity of u implies

$$u_{cc}u_{ll} - u_{cl}u_{lc} > 0 \qquad \text{(A.2.3)}$$

Then, since $u_l < 0$, $sign\left(\frac{\partial l}{\partial q}\right) = sign\left(-\frac{u_{cc}}{u_c} + \frac{u_{lc}}{u_l}\right)$, that is, leisure is non-inferior if $-\frac{u_{cc}}{u_c} + \frac{u_{lc}}{u_l} \geq 0$.

From definition of Δ_c and Δ_l

$$\Delta_c - \Delta_l = \left(\frac{u_{lc}}{u_l} - \frac{u_{cc}}{u_c}\right) c + \left(\frac{u_{ll}}{u_l} - \frac{u_{lc}}{u_c}\right) l. \tag{A.2.4}$$

From eq. (6) in steady state

$$c = \bar{w}l + \bar{R}a = -\frac{u_l l}{u_c} + \bar{R}a \tag{A.2.5}$$

where $\bar{R}$ is after-tax return on assets. Substituting (A.2.5) into (A.2.4) and rearranging we have:

$$\Delta_c - \Delta_l = \frac{u_l}{u_c^2} l \left[u_{cc} - 2\frac{u_c}{u_l} u_{lc} + \left(\frac{u_c}{u_l}\right)^2\right] + \left(\frac{u_{lc}}{u_l} - \frac{u_{cc}}{u_c}\right) \bar{R}a\,. \tag{A.2.6}$$

The term in squared brackets is a quadratic form of the Hessian of u and is negative under concavity. Since $u_l < 0$, we have $\Delta_c - \Delta_l > 0$ if $-\frac{u_{cc}}{u_c} + \frac{u_{lc}}{u_l} \geq 0$, i.e. if leisure is non-inferior.

From definition of Δ_p and Δ_c,

$$\Delta_p - \Delta_c = \frac{u_{pp}}{u_p} p - \frac{u_{cc}}{u_c} c - \frac{u_{lc}}{u_l} l. \tag{A.2.7}$$

Using (A.2.5), (A.2.7) can be written as

$$\Delta_p - \Delta_c = \frac{u_{pp}}{u_p} p + \frac{u_l}{u_c}\left(\frac{u_{cc}}{u_c} - \frac{u_{lc}}{u_l}\right) l - \frac{u_{cc}}{u_c} \bar{R}a\,. \tag{A.2.8}$$

Then, $-\frac{u_{cc}}{u_c} + \frac{u_{lc}}{u_l} \geq 0 \Rightarrow \Delta_p - \Delta_c > 0$, since $\frac{u_{pp}}{u_p} p > 0$.

Next, from definition of Δ_Q and using the steps above, we have

$$\Delta_Q - \Delta_c = \frac{u_{pQ}}{u_Q} p - 1 + \frac{u_l}{u_c}\left(\frac{u_{cc}}{u_c} - \frac{u_{lc}}{u_l}\right) l - \frac{u_{cc}}{u_c} \bar{R}a\,. \tag{A.2.9}$$

Non-inferiority of leisure implies $\Delta_Q - \Delta_c > \frac{u_{pQ}}{u_Q} p - 1$.

If complete additive separability, $\Delta_Q= 0$, then from the definition of Δ_c we have $\Delta_Q - \Delta_c= -\frac{u_{cc}}{u_c}c - 1$.

Finally, from (33) $(1+\lambda\Delta_c) > 0$, next $(1+\lambda\Delta_c) = \left(1+\lambda\Delta_p\right) - \lambda\left(\Delta_p - \Delta_c\right)$. So, $\Delta_p - \Delta_c > 0$ implies $\left(1+\lambda\Delta_p\right) > 0$.

Appendix A.3. Proof of Proposition 2

Proof: As for the sign of τ^x, by using (15), in steady state we get

$$r = \frac{\rho}{[1-\tau^a]}. \tag{A.3.1}$$

equating (A.3.1) and (27), at the symmetric equilibrium we obtain:

$$\left[1 - \Psi\left(\frac{X}{F}\right) + \Psi'\left(\frac{X}{F}\right)\frac{X}{F}\right]F_K - \delta = \frac{\rho}{[1-\tau^a]}. \tag{A.3.2}$$

From (35), in steady state we can write:

$$\rho = \frac{u_p}{q^k}\left(1+\lambda\Delta_p\right)\frac{\partial\bar{p}}{\partial F}F_K + \left\{\left[1 - \Psi\left(\frac{X}{F}\right) + \Psi'\left(\frac{X}{F}\right)\frac{X}{F}\right]F_K - \delta\right\} \tag{A.3.3}$$

Substituting from eq. (A.3.2), (A.3.3) can be written as:

$$\rho = \frac{u_p}{q^k}\left(1+\lambda\Delta_p\right)\frac{\partial\bar{p}}{\partial F}F_K + \frac{\rho}{[1-\tau^a]}. \tag{A.3.4}$$

and, exploiting (33)

$$\frac{1}{[1-\tau^a]} = 1 - \frac{u_p}{u_c}\frac{\left(1+\lambda\Delta_p\right)}{(1+\lambda\Delta_c)}\frac{\partial\bar{p}}{\partial F}\frac{F_K}{\rho} \tag{A.3.5}$$

Substituting for the LHS of (A.3.5) into (26) provides:

$$\tau^x = \frac{u_p}{u_c}\frac{\partial\bar{p}}{\partial X}\left[1 - \frac{u_p}{u_c}\frac{(1+\lambda\Delta_p)}{(1+\lambda\Delta_c)}\frac{\partial\bar{p}}{\partial F}\frac{F_K}{\rho}\right] - \Psi'\left(\frac{X}{F}\right) \tag{A.3.6}$$

Exploiting (36) to substitute for $\Psi'\left(\frac{X}{F}\right)$ into (A.3.6) we obtain:

$$\tau^x = \frac{u_p}{u_c}\frac{\partial\bar{p}}{\partial X} - \frac{u_p}{u_c}\frac{\partial\bar{p}}{\partial X}\frac{u_p}{u_c}\frac{(1+\lambda\Delta_p)}{(1+\lambda\Delta_c)}\frac{\partial\bar{p}}{\partial F}\frac{F_K}{\rho} - \frac{u_p}{u_c}\frac{(1+\lambda\Delta_p)}{(1+\lambda\Delta_c)}\frac{\partial\bar{p}}{\partial X} + \eta\frac{q^Q}{q^k} \tag{A.3.7}$$

Collecting terms of (A.3.7) it descends:

$$\tau^x = \frac{u_p}{u_c}\frac{\partial\bar{p}}{\partial X}\frac{(1+\lambda\Delta_p)}{(1+\lambda\Delta_c)}\left[\lambda\frac{(\Delta_c-\Delta_p)}{(1+\lambda\Delta_p)} - \frac{u_p}{u_c}\frac{\partial\bar{p}}{\partial F}\frac{F_K}{\rho}\right] + \eta\frac{q^Q}{q^k} \tag{A.3.8}$$

By our assumptions and by Lemma 1, $\frac{u_p}{u_c}\frac{\partial\bar{p}}{\partial X}\frac{(1+\lambda\Delta_p)}{(1+\lambda\Delta_c)}\left[\lambda\frac{(\Delta_c-\Delta_p)}{(1+\lambda\Delta_p)} - \frac{u_p}{u_c}\frac{\partial\bar{p}}{\partial F}\frac{F_K}{\rho}\right] > 0$.

Finally, given that $q^k > 0$ and that $q^Q > 0$ (from eq. (30) and increase of ϕ relaxes the environmental constraint, so that $\frac{\partial\Lambda}{\partial\phi} = q^Q > 0$), it follows that $\tau^x(t) > 0$.

As for τ^a, the argument presented in Proposition 1 applies. Finally, as for τ^l, under specifications (S.1), (S.2) and (S.4), the term $\frac{\partial\bar{p}}{\partial F}$ in eq. (38) is zero, while $(\Delta_l - \Delta_c) > 0$ by Lemma 1, sub L.1), so that $\tau^l > 0$. Under specification (S.3), $\frac{\partial\bar{p}}{\partial F} < 0$, so that the sign of τ^l is ambiguous. □

Appendix A.4

Detailed summary of the results on first-best and second best-tax structure at the steady state

<table>
<tr><td></td><td></td><td>S.1
$(\bar{p} = \gamma \cdot X)$</td><td>S.2
$\left(\bar{p} = \gamma \cdot \frac{X}{Q}\right)$</td><td>S.4
$(\bar{p} = \gamma)$</td><td>S.3
$\left(\bar{p} = \gamma \cdot \frac{X}{F}\right)$</td></tr>
<tr><td colspan="6"></td></tr>
<tr><td rowspan="3">$\boldsymbol{\frac{\tau^l}{1-\tau^l}}$</td><td>First Best
$(\lambda = 0)$</td><td colspan="3">0</td><td>$-\frac{u_p}{u_l}\frac{\bar{p}}{F}F_L < 0$</td></tr>
<tr><td>Second best add-on component</td><td colspan="3">$\lambda\frac{(\Delta_l - \Delta_c)}{(1+\lambda\Delta_c)} > 0$</td><td>$\lambda\frac{(\Delta_l-\Delta_c)}{(1+\lambda\Delta_c)} - \frac{u_p}{u_l}\lambda\frac{(\Delta_p-\Delta_c)}{(1+\lambda\Delta_c)}\frac{\bar{p}}{F}F_L > 0$</td></tr>
<tr><td></td><td colspan="3">τ^l is zero at first best and positive at second best</td><td>τ^l is negative at first best and ambiguous at second best</td></tr>
<tr><td colspan="6"></td></tr>
<tr><td rowspan="3">$\boldsymbol{\tau^a}$</td><td>First Best
$(\lambda = 0)$</td><td colspan="3">0</td><td>$\frac{u_p}{u_c}\frac{\bar{p}}{F}\frac{F_K}{r} < 0$</td></tr>
<tr><td>Second best add-on component</td><td colspan="3">0</td><td>$\frac{u_p}{u_c}\lambda\frac{(\Delta_p-\Delta_c)}{(1+\lambda\Delta_c)}\frac{\bar{p}}{F}\frac{F_K}{r} < 0$</td></tr>
<tr><td></td><td colspan="3">τ^a is zero both at first and second best</td><td>τ^a is negative both at first and second best</td></tr>
<tr><td colspan="6"></td></tr>
<tr><td rowspan="5">$\boldsymbol{\tau^x}$</td><td>First best without warm glow: τ^x_{FB}
$(\lambda = 0)$</td><td colspan="4">$\frac{\eta}{\rho+\mu}H\frac{u_Q}{u_c} > 0$</td></tr>
<tr><td>First best add-on component with warm-glow: $\tau^x_{FB}(p)$
$(\lambda = 0)$</td><td>0</td><td>$-\frac{u_p}{u_c}\left(\frac{\eta}{\rho+\mu}\frac{\bar{p}}{Q}\right) > 0$</td><td>0</td><td>$\left(\frac{u_p}{u_c}\right)^2\left(\frac{\bar{p}}{F^2}\frac{F_K}{\rho}\right) > 0$</td></tr>
<tr><td>Second best add-on component without warm-glow τ^x_{SB}</td><td colspan="4">$\frac{\eta}{\rho+\mu}H\frac{u_Q}{u_c}\lambda\frac{(\Delta_Q - \Delta_c)}{(1+\lambda\Delta_c)} > 0$</td></tr>
<tr><td>Second best add-on component with warm-glow $\tau^x_{SB}(p)$</td><td>$-\frac{u_p\lambda}{u_c}\frac{\bar{p}}{X}\frac{(\Delta_p - \Delta_c)}{(1+\lambda\Delta_c)} > 0$</td><td>$-\frac{u_p\lambda}{u_c}\frac{(\Delta_p - \Delta_c)}{(1+\lambda\Delta_c)}\frac{\bar{p}}{Q}\left(\frac{\eta}{\rho+\mu} + \frac{Q}{X}\right) > 0$</td><td>0</td><td>$-\frac{u_p\lambda}{u_c}\frac{(\Delta_p - \Delta_c)}{(1+\lambda\Delta_c)}\frac{\bar{p}}{X}\left(1 - \frac{u_p}{u_c}\frac{\bar{p}}{F}\frac{F_K}{\rho}\right) > 0$</td></tr>
<tr><td></td><td colspan="4">τ^x is positive both at first and second best</td></tr>
</table>