CESifo WORKING PAPERS

8749
2020

December 2020

Kantians Defy the Economists' Mantra of Uniform Pigovian Emissions Taxes

Thomas Eichner, Rüdiger Pethig

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8749*

Kantians Defy the Economists' Mantra of Uniform Pigovian Emissions Taxes

Abstract

This paper analyzes the efficient emissions taxation in economies with individuals who are morally motivated to reduce their emissions footprint. They are heterogenous with respect to their morality and their consumption preferences. We distinguish between the concepts of moral and conventional utilitarian (= material) welfare. The materially efficient tax rates turn out to be consumer-type specific; they are smaller than the Pigovian tax rate; and the smaller, the higher the individuals' propensity to act morally. Finally, we briefly characterize the second-best uniform emissions tax.

JEL-Codes: H210, Q580.

Keywords: Pigovian tax, material, moral, Kantian, consumer-type tax.

Thomas Eichner
Department of Economics
University of Hagen
Universitätsstr. 41
Germany – 58097 Hagen
thomas.eichner@fernuni-hagen.de

Rüdiger Pethig
Department of Economics
University of Siegen
Unteres Schloss 3
Germany – 57072 Siegen
pethig@vwl.wiwi.uni-siegen.de

We are grateful to Alistair Ulph and David Ulph for their helpful comments. Remaining errors are the authors' sole responsibility.

1 Introduction

The climate policy instruments applied in practice amount to imposing explicit or implicit (shadow) prices on emissions, and these prices often differ across sectors and regions (World Bank 2019). For reasons of cost-effectiveness and efficiency, economists call for a uniform (shadow) price on all emissions. The first-best corrective tax they unisonously recommend is the time-honored Pigovian tax (Pigou, 1947) that is designed to internalize all negative externalities caused by the emission of pollutants.

Central to the microeconomic rationale for the Pigovian tax is the behavior of *homo oeconomicus*. That agent's self-interested actions fail to explain, however, why a growing number of people make deliberate contributions to reduce their carbon emissions, e.g. by changing their diet or by making compensation payments for air travel. These people incur costs, but the benefits they receive in the form of reduced climate damage due to their lower emissions is negligible. We refrain from reviewing the explanations for such deviation from the behavior of homo oeconomicus put forward in recent years[1] and focus, instead, on the proposition that individuals deliberately reduce their emission footprint, since they are morally motivated "to do the right thing". Specifically, we consider a concept of moral utility inspired by Alger and Weibull (2013, 2016, 2020) which in turn goes back on Kant's ethical approach.

Alger and Weibull (2013, 2016) prove that a specific form of preferences with the same degree of morality for all individuals is evolutionary stable. We do adopt their concept, but drop the aspect of evolutionary stability and also deviate from their concept by allowing for consumers with alternative degrees of morality. In the context of international emissions cap competition, Alger and Weibull's morality concept has been applied by Eichner and Pethig (2020). Daube and Ulph (2016) consider a similar albeit not identical morality concept.[2] A different concept of Kantian behavior underlies the Kant equilibrium approach of Roemer (2010, 2015) that has been extended by Grafton et al. (2017) and Van Long (2020) to study the interaction of Kantian and Nashian agents in Kant-Nash equilibria.

This paper aims to characterize efficient emissions tax policies in economies with in-

---

[1] A review of factors that influence pro-environmental behavior is given by Kollmuss and Agyeman (2002) and Ertz and Sarigöllü (2016).

[2] For more information about the implication of Daube and Ulph's (2016) morality concept on the efficient environmental policy we refer to footnote 10.

dividuals who act morally in the sense of Alger and Weibull and who are heterogeneous with respect to their consumption preferences, in particular with respect to their degree of morality. We will show that analyzing economies with moral consumers makes it necessary to distinguish between the concepts of *moral* and *material* welfare, where the latter concept is that which economists use to apply in their models with conventional self-interested consumers (homines oeconomici). When individuals with moral preferences are considered, the important and philosophically non-trivial question arises, whether material or moral welfare should be the policymakers' relevant concept of welfare and efficiency (Alger and Weibull 2020). It is beyond the scope of the present paper to deal with this issue adequately.[3]

We show that if material efficiency is considered the relevant concept, the materially efficient tax rates are consumer-type specific, they differ from the morally efficient rates, in general, and all these tax rates are smaller than the Pigovian rate. As an implication of these unexpected and striking results, the implementation of the Pigovian tax would result in environmental damage that is lower than the materially efficient damage, and any other uniform emissions tax would also spell material inefficiency.

Consumer-type specific tax rates are also needed in other contexts, e.g. when externalities are heterogeneous, asymmetric or local (Diamond 1973, Green and Sheshinski 1976, Eckerstorfer and Wendner 2013, Knittel and Sandler 2018), or when emissions cannot be measured reliably enough to impose a Pigovian tax (Fullerton and West 2002). Since consumer-type specific tax rates are hard to implement, we follow the literature and briefly analyze second-best taxation. Due to the partial equilibrium nature of our model we focus on the second-best uniform emissions tax. It is shown that the second-best tax accounts for the society's average morality and that the welfare loss of second-best taxation is the larger, the larger the variance of the society's morality distribution.

The subsequent Section 2 describes, formalizes and analyzes the economy with moral consumers. First we introduce the building blocks of the model (Section 2.1), then we characterize efficient policies in economies with identical moral consumers (Section 2.2),

---

[3] In the concluding section, we put some arguments forward in favor of material welfare. Our procedure is to juxtapose and compare the tax policies that are efficient with regard to either welfare concept. But we think the characterization of materially efficient taxes is particularly interesting, because it is important to know what the consequences are, if economists and policymakers would wish to stick to the prevailing paradigm of the uniform (Pigovian) emissions tax in a world with moral consumers.

which prepares for the central Section 2.3, the analysis of economies with heterogeneous moral consumers. Section 3 characterizes second-best uniform emissions taxes. Section 4 concludes.

2 Efficient taxation when consumers act morally

We consider an economy with a continuum of individuals who consume the quantity z of a clean good and the quantity y of a dirty good. These goods are produced with the inputs $r_y = cy$, $c > 0$, and $r_z = z$, respectively. By means of that simple linear technology, the consumers' given endowments r of the input factor are readily transformed into alternative consumption bundles (y, z) that satisfy

$$z = r - cy. \tag{1}$$

Each unit of the dirty good releases one unit of emissions. Hence, aggregate emissions are equal to $m\bar{y}$, where $m \in \mathbb{R}_{++}$ measures the size of the population and $\bar{y}$ is the average consumption of the dirty good (to be formally defined below). All consumers suffer from environmental damage $H(m\bar{y})$ caused by aggregate emissions. The damage function H satisfies $H' > 0$ and $H'' \geq 0$.

Moral utility. Central to the model is how consumers value their consumption of the dirty and the clean good. The point of departure is a simple version of the conventional utility function in economic modeling

$$U(y, z, \theta) = B(y, \theta) + z - H(m\bar{y}), \tag{2}$$

where $B(y, \theta)$ is the private gross utility derived from consuming the dirty good. The benefit function B satisfies[4] $B_y > 0, B_{yy} < 0$, and $B_\theta \neq 0$. θ is a parameter that may vary across individuals in the set $\Theta \subset \mathbb{R}_+$. We denote the utility (2) as *material* utility and consumers who value their consumption by means of the utility function U as *self-interested* consumers. These consumers gain the marginal utility $U_y(y, z, \theta) = B_y(y, \theta)$ from consuming the dirty good, since they correctly take into account that any variation of their consumption leaves total emissions $m\bar{y}$ unchanged. Following Laffont (1975) and Alger and Weibull (2013, 2016,

---

[4] Subscripts attached to capital letters indicate partial derivatives.

2020), we also consider consumers (*Kantians*,[5] for short) who value their consumption with the *Kantian utility* function

$$V(y,z,\theta) = B(y,\theta) + z - H(my). \tag{3}$$

Kantians choose their consumption under the assumption that all others choose the same consumption as they do. They behave so without knowing whether the others really act like that. Their motivation is to choose that and only that dirty-good consumption they advocate all others choose as well. As a result, their marginal utility from consuming the dirty good is $V_y(y,z,\theta) = B_y(y,\theta) - mH'(my)$, which is smaller than that of self-interested consumers.

The propensity to act morally of real-world consumers is likely less rigorous than that of Kantians. To formalize this idea, we follow Alger and Weibull (2013, 2016, 2020) who introduce a morality parameter $\kappa \in K \equiv [0,1]$ and define the utility of a *moral consumer*[6] as the convex combination of (2) and (3), i.e. of material and Kantian utility,

$$\begin{aligned} W(y,z,\kappa,\theta) &:= (1-\kappa)U(y,z,\theta) + \kappa V(y,z,\theta) \\ &= B(y,\theta) + z - (1-\kappa)H(m\bar{y}) - \kappa H(my). \end{aligned} \tag{4}$$

The two polar cases of utility $W(y,z,\kappa,\theta)$ in (4) obviously are the material utility $W(y,z,\kappa = 0,\theta) = U(y,z,\theta)$ of a self-interested consumer and the moral utility $W(y,z,\kappa = 1,\theta) = V(y,z,\theta)$ of a Kantian. Note also that consumers with degree of morality $\kappa \in]0,1]$ can be characterized by both their *moral* utility $W(y,z,\kappa,\theta)$ and their *material* utility $U(y,z,\theta)$.

Consumer heterogeneity with respect to the degree of morality $\kappa \in K$ and with respect to the parameter $\theta \in \Theta$ is central to our subsequent analysis. We forego unnecessary complexity by varying one of these parameters at a time such that the distribution of types in the population will be represented either by the density function f or by the density function g, where

$$\begin{aligned} &f(\kappa) \geq 0, \quad \int_K f(\kappa)\mathrm{d}\kappa = 1, \quad \kappa \in K \equiv [0,1], \\ &g(\theta) \geq 0, \quad \int_\Theta g(\theta)\mathrm{d}\theta = 1, \quad \theta \in \Theta \subset \mathbb{R}_+. \end{aligned} \tag{5}$$

---

[5] In Alger and Weibull (2013, 2016, 2020), a self-interested consumer is called homo oeconomicus and a Kantian is called homo kantiensis.

[6] Alger and Weibull refer to a consumer with degree of morality $\kappa > 0$ as *homo moralis.*

When our focus is on heterogeneity in κ, we assume that $\theta_0 \in \Theta$ holds for all consumers and consider the economy $\{K, \theta_0\}$. Likewise, we assume that $\kappa_0 \in K$ holds for all consumers and consider the economy $\{\kappa_0, \Theta\}$, when we focus on heterogeneity in the preference parameter θ. We also briefly consider the economy $\{\kappa_0, \theta_0\}$ in which all consumers exhibit the same parameters κ_0 and θ_0, because that helps assessing the heterogeneity results.

Emissions in the market economy. Suppose next, there are perfectly competitive markets for the productive factor and the consumption goods and that the government imposes an emissions tax at rate t on the dirty good. Since production is linear, the equilibrium (producer) prices trivially are $p_r = 1, p_y = c$ and $p_z = 1$. The consumer's budget constraint takes the form

$$\underbrace{r + t\bar{y}}_{=\omega} = (p_y + t)y + p_z z = (c + t)y + z. \tag{6}$$

For the time being, we take the emissions tax rate $t \geq 0$ in (6) as given. The individuals' income $\omega = r + t\bar{y}$ consists of their given resource income r and the lumpsum transfer of tax revenues, $t\bar{y}$. Consumers take that transfer as given, they replace z in (4) by z from (6) and choose that consumption of the dirty good, which maximizes with respect to y the utility $W(y, \omega - (c+t)y, \kappa_0, \theta_0)$, $W(y, \omega - (c+t)y, \kappa_0, \theta)$ and $W(y, \omega - (c+t)y, \kappa, \theta_0)$ that are associated with the economies $\{\kappa_0, \theta_0\}$, $\{\kappa_0, \Theta\}$ and $\{K, \theta_0\}$, respectively. The pertaining first-order conditions are

$$\{\kappa_0, \theta_0\}: \qquad B_y(\tilde{y}, \theta_0) = c + t + \kappa_0 m H'(m\tilde{y}), \tag{7}$$

$$\{K, \theta_0\}: \qquad B_y(\tilde{y}, \theta_0) = c + t + \kappa m H'(m\tilde{y}), \tag{8}$$

$$\{\kappa_0, \Theta\}: \qquad B_y(\tilde{y}, \theta) = c + t + \kappa_0 m H'(m\tilde{y}). \tag{9}$$

The equations (7)-(9) determine a moral consumer's dirty-good consumption $\tilde{y}$ in the market economies under review, when the government imposes an emissions tax at rate $t \geq 0$. Total differentiation of (8) subject to $\mathrm{d}t = 0$ yields

$$\frac{\mathrm{d}y}{\mathrm{d}\kappa} = \frac{mH'}{B'' - \kappa m^2 H''} < 0. \tag{10}$$

(10) proves that the individuals' emissions in the market economy with predetermined emissions tax are the smaller, the higher their degree of morality.[7] That result conforms to

---

[7] That inverse relationship between emissions and degrees of morality is derived by Alger and Weibull (2020) in a short section on environmental economics.

intuition: The higher the propensity to act morally, the larger the voluntary emissions reductions compared to the emissions of self-interested consumers.

In the market economy, the price moral consumers pay for reducing their carbon footprint is a loss of material utility. To demonstrate that, consider a consumer with morality $\kappa \in]0,1[$ in the market economy $\{K, \theta_0\}$. Denote her dirty-good consumption in the market equilibrium by $\tilde{y}(\kappa) = \text{argmax}_y\, W\left(y, \omega - (c+t)y, \kappa, \theta_0\right)$ and her corresponding material utility by $u(\kappa) = B\left(\tilde{y}(\kappa), \theta_0\right) + \omega - (c+t)\tilde{y}(\kappa) - H(m\bar{y})$. If the moral consumer would act like a self-interested consumer ($\kappa = 0$), her dirty-good consumption would be $\tilde{y}(0) = \text{argmax}_y \left[B\left(y, \theta_0\right) + \omega - (c+t)y(\kappa) - H(m\bar{y})\right]$ and the corresponding material utility would be $u(0) = B\left(\tilde{y}(0), \theta_0\right) + \omega - (c+t)\tilde{y}(0) - H(m\bar{y})$. By definition of $\tilde{y}(0)$, we have $u(0) > B\left(y, \theta_0\right) + \omega - (c+t)y - H(m\bar{y})$ for all $y \neq y(0)$. Since $\tilde{y}(\kappa) \neq \tilde{y}(0)$ due to (10), it follows that $u(0) > u(\kappa)$.

Material and moral efficiency. Market allocations and policy outcomes use to be assessed by the concept of efficiency (or social optimum). An allocation of resources is said to be efficient, if it maximizes the country's (utilitarian) welfare defined as the aggregate individual utilities. Standard economic models without moral consumers apply the *material* welfare function, i.e. the integral of individual *material* utility functions (2). However, if consumers are moral, the social planner may choose between the welfare that is based on the individuals' *material* or on their *moral* utility. In the latter case, the planner maximizes the *moral* welfare function defined as the integral of the individual *moral* utility functions (4).

Specifically, we invoke (4), (5) and (6) and define the *moral* welfare functions for the economies $\{K, \theta_0\}$ and $\{\kappa_0, \Theta\}$ as

$$\Omega(y, K, \theta_0) = \int_K S(y, \kappa, \theta_0) f(\kappa) \mathrm{d}\kappa \qquad \text{and} \qquad \Omega(y, \kappa_0, \Theta) = \int_\Theta S(y, \kappa_0, \theta) g(\theta) \mathrm{d}\theta, \tag{11}$$

where

$$\begin{aligned} S(y, \kappa, \theta_0) &= B(y, \theta_0) + r - cy - (1-\kappa) H\left(m\bar{y}(K)\right) - \kappa H(my), & (12) \\ S(y, \kappa_0, \theta) &= B(y, \theta) + r - cy - (1-\kappa_0) H\left(m\bar{y}(\Theta)\right) - \kappa H(my). & (13) \end{aligned}$$

In (12) and (13), $\bar{y}(K) = \int_K y(\kappa) f(\kappa) \mathrm{d}\kappa$ and $\bar{y}(\Theta) = \int_\Theta y(\theta) g(\theta) \mathrm{d}\theta$ are the average consumption of the dirty good, when the consumers are heterogeneous with respect to κ and θ, respectively. The *material* welfare functions are special cases of the *moral* welfare functions (11) that result from replacing $S(y, \kappa, \theta_0)$ by

$$S(y, 0, \theta_0) = B(y, \theta_0) + r - cy - H\left(m\bar{y}(K)\right)$$

and $S(y, \kappa_0, \theta)$ by

$$S(y, 0, \theta) = B(y, \theta) + r - cy - H\left(m\bar{y}(\Theta)\right).$$

We characterize *morally* efficient allocations by maximizing the welfare functions (11) with respect to y. The maximizers $\hat{y}(K, \theta_0)$ and $\hat{y}(\kappa_0, \Theta)$ satisfy the first-order conditions

$$\begin{aligned} \{K, \theta_0\}: \qquad B_y(\hat{y}, \theta_0) &= c + (1-\bar{\kappa})mH'\left(m\hat{\bar{y}}(K)\right) + \kappa m H'(m\hat{y}), && (14) \\ \{\kappa_0, \Theta\}: \qquad B_y(\hat{y}, \theta) &= c + (1-\kappa_0)mH'\left(m\hat{\bar{y}}(\Theta)\right) + \kappa_0 m H'(m\hat{y}), && (15) \end{aligned}$$

where $\bar{\kappa} = \int_K \kappa f(\kappa) \mathrm{d}\kappa$. *Material* efficiency is characterized by the special cases of (14) and (15) for $\bar{\kappa} = \kappa = \kappa_0 = 0$,

$$\begin{aligned} \{K, \theta_0\}, \{\kappa_0, \theta_0\}: \qquad & B_y(\hat{y}, \theta_0) = c + mH'(m\hat{y}), && (16) \\ \{\kappa_0, \Theta\}: \qquad & B_y(\hat{y}, \theta) = c + mH'(m\hat{\bar{y}}(\Theta)). && (17) \end{aligned}$$

The straightforward implications of (14) - (17) are summarized in

Proposition 1. *(Divergence of efficiency concepts)*

(i) In the economies $\{\kappa_0, \theta_0\}$, in which consumers are identical, the concepts of moral and material efficiency coincide.

(ii) In the economies $\{\kappa_0, \Theta\}$ and $\{K, \theta_0\}$, the difference between morally and materially efficient allocations tends to be the greater, the more heterogeneous the consumers with respect to the preference parameters κ and θ.

Proposition 1 highlights that the coincidence of moral and material efficiency, and (hence) the independence of moral efficiency from the consumers' degree of morality, is a special feature of the symmetry assumption. Since consumers are heterogeneous in the real world, the relevant case is that the two efficiency concepts differ. We proceed with characterizing and juxtaposing the tax rates that restore either kind of efficiency in the economies $\{K, \theta_0\}$ and $\{\kappa_0, \Theta\}$, and we begin with a brief look at the economy $\{\kappa_0, \theta_0\}$.

Efficient taxation with identical consumers. In economy $\{\kappa_0, \theta_0\}$, the materially and morally efficient emissions tax rate is

$$t = (1 - \kappa_0) m H'(m\hat{y}(\theta_0)), \tag{18}$$

which makes the equation (7) coincide with equation (16). We summarize the implications of (7), (16) and (18) for boundary and intermediate morality parameters in

Proposition 2. *(Efficient taxation and identical consumers)*

(i) In the economy $\{\kappa_0, \theta_0\}$, moral as well as material efficiency requires

- *the (uniform) Pigovian tax,*[8] *if all consumers are self-interested ($\kappa_0 = 0$);*
- *no emissions tax, if all consumers are Kantians ($\kappa_0 = 1$).*

(ii) In the economy $\{\kappa_0 \in]0, 1[, \theta_0\}$, emissions are too high in the absence of taxation, and there exists a unique uniform emissions tax rate that restores both material and moral efficiency. The efficient tax rate is the smaller, the higher the consumers' morality κ_0; it declines towards zero, if κ_0 tends to one, and it increases towards the Pigovian tax, if κ_0 tends to zero.

The economies with consumers of intermediate morality (Proposition 2(ii)) are particular interesting and arguably more relevant than the polar economies of Proposition 2(i). The statement in Proposition 2(ii) that the efficient tax is the lower, the higher the consumers' degree of morality,[9] immediately follows from (18), because the efficient individual emissions $\hat{y}(\theta_0)$ are independent of κ_0. The larger the morality parameter κ_0, the larger the consumers' morally motivated deliberate contributions to reduce emissions and, as a consequence, the lower needs to be the efficiency-restoring emissions tax. That outcome is in line with the subsidiary principle. Given the challenge to cope with climate change, that principle says that the government should first account for the citizens' deliberate emissions reductions. Only if these reductions fall short of the efficient ones, it should step in with an emissions tax that closes the remaining inefficiency gap.

---

[8] We denote as Pigovian tax (rate) the tax (rate) which fully internalizes the pollution externalities in economies with identical or heterogeneous *self-interested* consumers. The Pigovian tax is always uniform across consumers.

[9] Alger and Weibull (2020) point out that policy advice based on models with self-interested consumers may exaggerate the emissions tax required to restore efficiency.

	Moral efficiency	**Material efficiency**
$\{\kappa_0, \Theta\}$	$t = (1-\kappa_0)mH'\left(m\hat{\hat{y}}(\kappa_0, \Theta)\right)$	$t(\theta) = mH'(m\hat{\hat{y}}(\Theta)) - \kappa_0 mH'(m\hat{y}(\theta))$
$\{\kappa_0 = 0, \Theta\}$	$t_{\text{Pigou}} = mH'(m\hat{\hat{y}}(\Theta))$	$t_{\text{Pigou}} = mH'(m\hat{\hat{y}}(\Theta))$
$\{\kappa_0 = 1, \Theta\}$	$t = 0$	$t = 0$
$\{K, \theta_0\}$	$t = (1-\bar{\kappa})mH'(m\hat{\hat{y}}(K, \theta_0))$	$t(\kappa) = (1-\kappa)mH'(m\hat{y}(\theta_0))$

Table 1: Rates of the emissions tax required to restore either moral or material efficiency in the economies $\{K, \theta_0\}$ and $\{\kappa_0, \Theta\}$

The straightforward conclusion is that in economies with identical moral consumers the efficiency-restoring tax rate is uniform and the same for both efficiency concepts. It is clear, however, that consumers are not alike in the real world, and since the efficiency concepts differ in that case (Proposition 1), we expect different efficiency-restoring tax rates.

Efficient taxation with heterogeneous consumers. Here we determine the tax rates that are necessary for attaining material and moral efficiency in the market economies $\{K, \theta_0\}$ and $\{\kappa_0, \Theta\}$. We proceed as in economy $\{\kappa_0, \theta_0\}$ above, i.e. in case of economy $\{K, \theta_0\}$, we compare (8) with (14) and (16), and in case of economy $\{\kappa_0, \Theta\}$, we compare (9) with (15) and (17). The results are listed in Table 1 and highlighted in

Proposition 3. *(Efficient taxation and heterogeneous consumers)*

(i) In the economy $\{\kappa_0, \Theta\}$, moral as well as material efficiency requires

- *the (uniform) Pigovian tax, if all consumers are self-interested ($\kappa_0 = 0$);*
- *no emissions tax, if all consumers are Kantians ($\kappa_0 = 1$).*

(ii) Consider a consumer of type $\kappa \in]0, 1[$ in the economy $\{K, \theta_0\}$ or a consumer of type θ in the economy $\{\kappa_0 \in]0, 1[, \Theta\}$.

(a) The tax rate that induces such a consumer to release the morally efficient emissions differs from the tax rate that induces her to release the materially efficient emissions. All of these tax rates are positive and smaller than the Pigovian tax; they increase towards the Pigovian tax, if κ tends to zero, and they decline towards

zero, if κ tends to one.

(b) Moral efficiency requires tax rates that are uniform across consumers, but material efficiency requires consumer-type specific tax rates.

Proposition 3(i) adds to the insight of Proposition 2(ii) that in the polar cases $\kappa_0 = 0$ and $\kappa_0 = 1$ the efficient taxation does not depend on whether consumer preferences differ in aspects other than morality. The findings in Proposition 3(iia) are in line with Proposition 2(ii) keeping in mind that despite their common properties, the tax rates in Proposition 3(iia) differ depending on which economy is under review and which efficiency concept is applied. The most unexpected and striking result is Proposition 3(iib). It challenges and defies the economists' time-honored and widely propagated insight that the uniformity of the emissions taxes/prices is a necessary condition for materially efficient anti-pollution policies. In economies with heterogeneous moral consumers, material efficiency can be attained neither via the Pigovian tax rate nor with any other (smaller) tax rate that is uniform across consumers.[10]

The reason why material efficiency requires consumer-type specific emissions tax rates in an economy $\{K, \theta_0\}$ is straightforward.[11] According to (16), the materially efficient emissions, $\hat{y}$, are independent of κ. We combine (8), Table 1 and (16) and observe that the efficient emissions $\hat{y}$ satisfy

$$B_y(\hat{y}, \theta_0) = c + \underbrace{\kappa m H'(m\hat{y})}_{=\tau(\kappa)} + \underbrace{(1-\kappa) m H'(m\hat{y})}_{=t(\kappa)} = c + \underbrace{m H'(m\hat{y})}_{=t_{\text{Pigou}}} \tag{19}$$

for all consumer types $\kappa \in]0, 1[$. Hence the term $\tau(\kappa) = \kappa m H'(m\hat{y})$ is increasing and the tax rate $t(\kappa)$ is decreasing in κ. It is interesting to interpret the term $\tau(\kappa)$ in (19) as the rate of an emissions tax which consumers with morality κ impose on themselves. The government

---

[10] The concept of moral consumers in Daube and Ulph (2016) is similar to that of Alger and Weibull, which we apply here. In stark contrast to our Proposition 3(ii), Daube and Ulph (2016, Proposition 5) obtain the remarkable result that the Pigovian tax is the materially efficient tax for all consumers in the economy $\{K, \theta_0\}$. This result is not general, however. It can be shown that if the moral utility of consumers differs in aspects other than the degree of morality, as e.g. in the economy $\{\kappa_0 \in]0, 1[, \Theta\}$, then the materially efficient emissions tax rates are also type-specific when the moral-utility concept of Daube and Ulph is applied. Since in the real world consumer preferences differ in many aspects other than their degree of morality, type-specificity of materially efficient taxation appears to be the relevant case not only in our moral-consumer concept, but also in that of Daube and Ulph (2016).

[11] Analogous arguments apply to the economy $\{\kappa_0, \Theta\}$.

then needs to choose the tax rate $t(\kappa)$ such that both rates add up to the Pigovian tax, which is the tax required to restore material efficiency in economies with self-interested consumers only. The fraction κ of the environmental damage is deliberately internalized by the consumer with morality κ. The remaining fraction $(1-\kappa)$ of the damage is internalized through the emissions tax $t(\kappa)$.

3 Second-best uniform emissions tax

Governments are not able to implement consumer-type specific tax rates, because they are neither able to reveal the individuals' morality nor to observe the material utility of moral individuals. In addition, consumer-type specific tax rates may be hard to implement in a political process. In this section, we analyze the second-best uniform emissions tax. Although the government does not know the morality of each individual, we assume that it has information about the society's distribution of moralities. We restrict our attention to economies $\{K, \theta_0\}$. Since the first-best is attainable with uniform emissions taxes, when the government maximizes moral welfare, in the sequel we assume that the government's objective is material welfare. More precisely, the government maximizes with respect to t the material welfare

$$\Omega = m \int_K \left[B(\tilde{y}, \theta_0) + r - c\tilde{y} - H\left(m \int_K \tilde{y} f(\kappa) \mathrm{d}\kappa \right) \right] f(\kappa) \mathrm{d}\kappa \tag{20}$$

subject to $\tilde{y} = \tilde{Y}(t, \kappa, \theta_0)$ which is implicitly determined by (8). The associated first-order condition can be rearranged to[12]

$$t = \mathbb{E}\left[mH'(m\tilde{Y}) \right] - \frac{\mathbb{E}\left[\kappa m H'(m\tilde{Y}) \tilde{Y}_t \right]}{\mathbb{E}\left[\tilde{Y}_t \right]}. \tag{21}$$

According to (21), the second-best uniform emissions tax is equal to the Pigovian tax $\mathbb{E}\left[mH'(m\tilde{Y}) \right]$ less the moral society's deliberate internalization of the environmental damage at the margin. More specifically, $\mathbb{E}\left[\kappa m H'(m\tilde{Y}) \tilde{Y}_t \right]$ is the tax rate effect on the environmental damage via the change of dirty-good consumption, that is internalized by a moral society, related to the tax rate effect on the society's aggregate dirty-good consumption $\left(\mathbb{E}\left[\tilde{Y}_t \right] \right)$.

---

[12] The derivation of (21) can be found in the Appendix. In (21), $\mathbb{E}$ is the expectation operator.

To make things more transparent, we specify

$$B(y,\theta) = ay - \frac{\theta}{2}y^2, \quad H(m\bar{y}) = hm\bar{y}, \tag{22}$$

where a, b and h are positive parameters. The specification of preferences in (22) is stylized, but at least the linearity of H is a reasonable approximation in case of climate change (Golosov et al. (2014, p. 78) and the quadratic B-function is the simplest form of modeling convexity. (22) yields the dirty-good consumption $\tilde{Y} = \frac{a-c-t-\kappa mh}{\theta_0}$ of a consumer with morality κ. Making use of $\tilde{Y}_t = -\frac{1}{\theta_0}$ in (21) yields

$$t = (1-\bar{\kappa})\, mh. \tag{23}$$

The second-best tax rate (23) is smaller than the Pigovian tax rate. The deduction equals the marginal environmental damage that society internalizes on average due to its morality. (23) is driven by $\tilde{Y}_{t\kappa} = 0$, i.e. the effect of reducing dirty-good consumption upon an increase in the tax rate is independent of the consumers' morality. We conjecture that (23) does not hold in economies with functional forms that differ from the parametric functions (22), but it may be a reasonable benchmark for second-best emissions taxes to be applied in practice.

Finally, we compare first-best and second-best emissions and welfares. Indicating first[second]-best levels by the superscript E [S], we show in the Appendix[13]

$$y^{\mathrm{E}} = \frac{a-c-mh}{\theta_0}, \quad y^{\mathrm{S}} = \frac{a-c-mh+(\bar{\kappa}-\kappa)mh}{\theta_0}, \tag{24}$$

$$\Omega^{\mathrm{E}} - \Omega^{\mathrm{S}} = \frac{m^3h^2}{\theta_0^2}\mathbb{Var}\,(\kappa)\,. \tag{25}$$

Whereas the consumers' first-best consumption y^{E} is independent of morality κ, the second-best consumption y^{S} is the lower, the larger the consumer's morality. Comparing y^{E} and y^{S} reveals that second-best consumption of individuals with morality below the average morality ($\kappa < \bar{\kappa}$) is inefficiently large, and that second-best consumption of individuals with morality above the average morality ($\kappa < \bar{\kappa}$) is inefficiently small. These deviations from efficiency cancel out when aggregating dirty-good consumption. Hence, first-best and second-best aggregate emissions are equal.[14] In view of (25), second-best taxation causes a welfare loss that is the larger, the larger the morality distribution spreads out from average morality, i.e. the larger $\mathbb{Var}\,(\kappa)$. We summarize these results in

---

[13] In (25), $\mathbb{Var}$ is the variance operator.

[14] The proof is given in the Appendix.

Proposition 4. *Suppose the economy satisfies (22) and the government sets a uniform second-best emissions tax. Then*

(i) the second-best tax rate is by the factor $(1-\bar{\kappa})$ smaller than the Pigovian tax rate, where $\bar{\kappa}$ is society's average degree of morality.

(ii) aggregate second-best emissions are equal to aggregate first-best emissions;

(iii) the welfare loss in the transition from first-best to second-best is the larger, the larger $\mathbb{Var}(\kappa)$.

4 Concluding remarks

In practice, environmental taxation often does not aim at allocative efficiency in the theoretical sense of reaching the maximum possible material or moral welfare. Instead, the policy goals of many countries take the form of reducing aggregate emissions by some specified percentage (during some specified time period). Such goals amount to implementing a cap on aggregate emissions, and that cap may or may not be set at the first-best level. Nevertheless, the policymakers should seek to implement the cap cost-effectively, i.e. at minimum material or moral welfare costs. In economies with self-interested consumers, material cost-effectiveness is ensured by implementing the cap by means of a uniform emissions tax or by means of the uniform price of emissions allowances in an emissions trading scheme (ETS). From arguments analogous to those used to establish Proposition 3(iib) it follows, however, that the materially cost-effective implementation of a given emissions cap requires differentiated emissions prices in the economies $\{\theta_0, K\}$ and $\{\kappa_0, \Theta\}$.[15]

Consequently, material cost-effectiveness cannot be achieved in ETSs with heterogeneous moral consumers, since the price of emissions allowances is uniform. In most ETSs that are in operation, traders are not consumers but firms or installations. If 'green' firms participate in an ETS that are morally motivated in the sense that they deliberately reduce their emissions beyond the level consistent with profit maximization under the ETS, the situation is qualitatively the same as in an ETS with moral consumers. Such an ETS fails to implement the given cap cost-ineffectively. It is an important item for future theoretical and

---

[15] To formally derive that result, we replace the first-order condition (14) [(15)] by the first-order condition of maximizing the moral welfare function $\Omega(y, \kappa, \theta_0)$ $[\Omega(y, \kappa_0, \Theta)]$ from (11) subject to the constraint that aggregate emissions are smaller than or equal to the predetermined emissions cap.

applied research to investigate how significant the deviation of green firms from conventional profit maximization is and how strong the resultant cost-ineffectiveness is.

As pointed out in the introduction, it is both difficult and important to answer the question whether the moral or material utility is the appropriate measure of well-being and social welfare.[16] As we see it, the principle arguments in favor or against the welfare concepts under review are the following.

From the theoretical point of view, one is inclined to argue that the concept we should apply is moral welfare. After all, the consumers maximize the moral utility function (4) subject to their budget constraint and thus reveal that the resulting consumption is their optimal choice given the prevailing constraints. In that view, the first column of Table 1 lists the emissions tax policies that implement moral efficiency. In all these policies, the tax rates are uniform across heterogeneous consumers. This approach may appear to be conclusive, at first sight, but it encounters the fundamental criticism of Sen (1977) that the concept of revealed preference can explain any action and behavior as being the individual's optimum choice.

Alternatively, one may consider the utilitarian material welfare as the relevant concept. A possible rationalization is the interpretation that the Kantians' moral motivation and goal is to promote the society towards material efficiency.[17] Ideally, the Kantians would like to make that emission footprint, which they would make, or would be assigned to make, in a materially efficient state of the economy - regardless of what other consumers do. Unfortunately, incomplete information prevents them from figuring out what exactly their materially efficient emissions are in a world of heterogeneous individuals. Therefore, they apply the informationally feasible auxiliary strategy of choosing those emissions, which maximize their material utility under the counterfactual assumption that all other consumers make the same

---

[16] The necessity of choosing among different welfare concepts arises in theoretical studies whenever the consumers' actions are assumed to differ from those of self-interested consumers with material utility functions. Various studies of that kind, for example studies, in which consumers are assumed to be altruistic (e.g. Eckerstorfer and Wender 2013) or where they compare their income with that of other income groups (e.g. Kanbur and Tuomala, 2006; Micheletto 2011) proceed as we do. They elaborate and compare the implications of either welfare concept.

[17] Alger and Weibull (2020) also endorse material welfare as the relevant concept, but they do so with reference to Harsanyi (1979, 1992) who distinguishes between 'personal' and 'moral' preferences and suggests that welfare should be based on 'personal' preferences.

choice. As shown above, in economies inhabited by identical Kantians that auxiliary strategy yields material efficiency, indeed. Yet, in economies with heterogeneous Kantians and/or other heterogeneous consumers, the resultant allocation is materially inefficient, as shown in the second column of Table 1, where we list the type-specific tax policies governments need to implement to achieve material efficiency.[18] In this narrative, a Kantian's deviation from her materially efficient carbon footprint is a consequence of her incomplete information. In the second column of Table 1 we list the type-specific tax policies governments need to implement to achieve material efficiency. Basically, the need of type-specific taxation to restore material efficiency arises, because moral consumers choose their consumption via the constrained maximization of their *moral* instead of their *material* utility, whereas governments aim at reaching the maximum *material* instead of the maximum *moral* welfare.

We argued that Kantians would like to make their materially efficient footprint, but fail to do so because they lack the relevant information. It is hardly conceivable that governments will be able to collect all information on preferences and moral motivation that it is needed to implement material efficiency by means of the correct consumer-type specific tax rates. To cope with that dilemma in practice, governments may have to consider setting second-best emissions taxes. A first pass has been provided in Section 4, but a more thorough investigation of the theoretical foundations and the practical design of such taxes, e.g. in general equilibrium models with taxes on both clean and dirty goods, is an important item for further research.

References

Alger, I. and J. W. Weibull (2020): Morality: evolutionary foundations and economic implications, in Basu, K., Rosenblatt, D. and C. Sepulveda (eds.), *The State of Economics, the State of the World*, Cambridge, MIT Press.

Alger, J. and J.W. Weibull (2016): Evolution and Kantian morality, *Games and Economic Behavior* 98, 56-67.

Alger, I. and J. W. Weibull (2013): Homo moralis - preference evolution under incomplete information and assortative matching, *Econometrica* 81, 2269 - 2302.

---

[18] Appropriate complementary taxation is also necessary to reach material efficiency, when consumers differ with respect to their degree of morality.

Daube, M. and D. Ulph (2016): Moral behaviour, altruism and environmental policy, *Environmental and Resource Economics* 63, 505-522.

Diamond, P. (1973): Consumption externalities and imperfect corrective pricing, *Bell Journal of Economics and Management Science* 4, 526-538.

Eckerstorfer, P. and R. Wendner (2013): Asymmetric and non-atmospheric consumption externalities, and efficient consumption taxation, *Journal of Public Economics* 106, 42-56.

Eichner, T. and R. Pethig (2020): Climate policy and moral consumers, *Scandinavian Journal of Economics*, forthcoming.

Ertz, M. and E. Sarigöllü (2016): Exploring pro-environmental behaviors of consumers: An analysis of contextual factors, attitude, and behaviors, *Journal of Business Research* 69, 3971-3980.

Fullerton, D. and S. E. West (2002): Can taxes on cars and on gasoline mimic an unavailable tax on emissions?, *Journal of Environmental Economics and Management* 43, 135-157.

Grafton, R.Q., Kompas, T. and N. van Long (2017): A brave new world? Kantian-Nashian interaction and the dynamics of global climate change mitigation, *European Economic Review* 99, 31-42.

Golosov, M., J. Hassler, P. Krusell, and A. Tsyvinski (2014): Optimal taxes on fossil fuel in general equilibrium, *Econometrica* 82, 41-88.

Green, J. and E. Sheshinski (1976): Direct versus indirect remedies for externalities, *Journal of Political Economy* 84, 797-808.

Harsanyi, J.C. (1992): Game and decision theoretic models in economics, in R.J. Aumann and S. Hart (eds.), *Handbook of Game Theory*, Amsterdam: North-Holland.

Harsanyi, J.C. (1979): Rule utilitarianism, rights, obligations and the theory of rational behavior, *Journal of Philosophy*, 76, 218-236.

Kanbur, R., Pirttillae, J. and M. Tuomala (2006): Non-welfarist optimal taxation and behavioral public economics, *Journal of Economic Surveys* 20, 849-868.

Knittel, C. and R. Sandler (2018): The welfare impact of second-best uniform-Pigouvian taxation: Evidence from tansportation, *American Economic Journal: Economic Policy* 10, 211-242.

Kollmuss, A. and J. Agyeman (2002): Mind the Gap: Why do people act environmentally and what are the barriers to pro-environmental behavior?, *Environmental Education Research* 8, 239-260.

Laffont, J.-J. (1975): Macroeconomic constraints, economic efficiency and ethics: An introduction to Kantian economics, *Economica* 42, 430-437.

Micheletto, L. (2011): Optimal nonlinear redistributive taxation and public good provision in an economy with Veblen effects, *Journal of Public Economic Theory* 13, 71-96.

Pigou, A.C. (1947): *A Study in Public Finance*, London, MacMillan.

Roemer, J.E. (2015): Kantian optimization. A microfoundation for cooperation, *Journal of Public Economics* 127, 45-57.

Roemer, J.E. (2010): Kantian equilibrium, *Scandinavian Journal of Economics* 112, 1-24.

Sen, A.K. (1977): Rational Fools: A Critique of the Behavioral Foundations of Economic Theory, *Philosophy & Public Affairs*, 6, 317-344.

Van Long, N. (2020): A dynamic game with interaction between Kantian players and Nashian players, in Pineau, P.-O., Sigué, S. and S. Taboubi (eds.), *Games in Management Science: Essays in Honor of Georges Zaccour*, Springer: Switzerland.

World Bank (2019): State and Trends of Carbon Pricing 2019, Washington, DC, Doi: 10.1596/978-1-4648-1435-8.

Appendix

Derivation of (21):

Differentiating (20) with respect to t we obtain

$$\Omega_t = m \int_K \left[(B' - c)\, \tilde{Y}_t - H'\left(m\tilde{Y}\right) \cdot \left(m \int_K \tilde{Y}_t\, f(\kappa)\mathrm{d}\kappa\right)\right] f(\kappa)\mathrm{d}\kappa = 0. \tag{26}$$

Accounting for (8) in (26) results in

$$\Omega_t = m \int_K \left[\left(t + \kappa m H'(m\tilde{Y})\right) \tilde{Y}_t - H'\left(m\tilde{Y}\right) \cdot \left(m \int_K \tilde{Y}_t\, f(\kappa)\mathrm{d}\kappa\right)\right] f(\kappa)\mathrm{d}\kappa = 0. \tag{27}$$

Further rearranging (27) leads to

$$
\begin{aligned}
& \int_K \left[t\tilde{Y}_t + mH'(m\tilde{Y}) \left(\kappa \tilde{Y}_t - \int_K \tilde{Y}_t \, f(\kappa)\mathrm{d}\kappa \right) \right] f(\kappa)\mathrm{d}\kappa = 0 \\
\Longleftrightarrow \qquad & t\mathbb{E}\left[\tilde{Y}_t\right] = \mathbb{E}\left[\tilde{Y}_t\right] \mathbb{E}\left[mH'(m\tilde{Y})\right] - \mathbb{E}\left[\kappa m H'(m\tilde{Y})\tilde{Y}_t\right]. \qquad (28)
\end{aligned}
$$

Dividing (28) by $\mathbb{E}\left[\tilde{Y}_t\right]$ establishes (21)

Derivation of (24) and (25):

The material first-best and second-best consumption, respectively, is characterized by

$$
\begin{aligned}
B'(y^{\mathrm{E}}, \theta_0) &= c + mh, && (29) \\
B'(y^{\mathrm{S}}, \theta_0) &= c + mh + (\kappa - \bar{\kappa})mh. && (30)
\end{aligned}
$$

Solving (29) and (30) for y^{E} and y^{S}, respectively, yields

$$
y^{\mathrm{E}} = \frac{a - c - mh}{\theta_0}, \quad y^{\mathrm{S}} = \frac{a - c - mh + (\bar{\kappa} - \kappa)mh}{\theta_0}. \qquad (31)
$$

Aggregate first-best emissions are

$$
m \cdot y^{\mathrm{E}} = m \cdot \frac{a - c - mh}{\theta_0}. \qquad (32)
$$

Integrating (31) we obtain aggregate second-best emissions

$$
m \cdot \int_K y^{\mathrm{S}} f(\kappa)\mathrm{d}\kappa = m \cdot \int_K \left(\frac{a - c - mh + (\bar{\kappa} - \kappa)mh}{\theta_0} \right) f(\kappa)\mathrm{d}\kappa = m \cdot \frac{a - c - mh}{\theta_0}. \qquad (33)
$$

Comparing (32) and (33) shows that aggregate first-best emissions are identical to aggregate second-best emissions.

Inserting (31) in (12) and (20), respectively, we obtain the *average* first-best and second-best material welfares

$$
\begin{aligned}
\frac{\Omega^{\mathrm{E}}}{m} &= \frac{(a - c - hm)^2}{2\theta_0}, && (34) \\
\frac{\Omega^{\mathrm{S}}}{m} &= \int_K \left[a y^{\mathrm{S}} - \frac{\theta_0}{2} \left(y^{\mathrm{S}}\right)^2 - c y^{\mathrm{S}} - hm y^{\mathrm{S}} \right] f(\kappa)\mathrm{d}\kappa \\
&= \frac{(a - c - hm)^2}{\theta_0} - \frac{\theta_0}{2} \int_K \left(y^{\mathrm{S}}\right)^2 f(\kappa)\mathrm{d}\kappa. && (35)
\end{aligned}
$$

Observe that

$$\begin{aligned}\int_K \left(y^{\mathrm{S}}\right)^2 f(\kappa)\mathrm{d}\kappa &= \int_K \left(\frac{a-c-mh+(\bar{\kappa}-\kappa)mh}{\theta_0}\right)^2 f(\kappa)\mathrm{d}\kappa \\ &= \int_K \left[\frac{(a-c-mh)^2+2(a-c-mh)(\bar{\kappa}-\kappa)mh+(\bar{\kappa}-\kappa)^2m^2h^2}{\theta_0^2}\right] f(\kappa)\mathrm{d}\kappa \\ &= \frac{(a-c-mh)^2}{\theta_0^2}+\frac{m^2h^2}{\theta_0^2}\int_K (\bar{\kappa}-\kappa)^2 f(\kappa)\mathrm{d}\kappa \\ &= \frac{(a-c-mh)^2}{\theta_0^2}+\frac{m^2h^2}{\theta_0^2}\mathbb{V}\mathrm{ar}\left(\kappa\right). \end{aligned} \tag{36}$$

Inserting (36) in (35) we get

$$\frac{\Omega^{\mathrm{S}}}{m} = \frac{(a-c-hm)^2}{2\theta_0} - \frac{m^2h^2}{\theta_0^2}\mathbb{V}\mathrm{ar}\left(\kappa\right). \tag{37}$$

Comparing (34) and (37) establishes (25).